

Introduction: The Chomskian Perspective on Language Study

Contents

Introduction

- 1 Linguistics: the science of language**
- 2 The native speaker: grammaticality and acceptability**
 - 2.1 Descriptive adequacy
 - 2.2 Grammaticality and acceptability
 - 2.3 The grammar as a system of principles
- 3 Knowledge of language**
 - 3.1 The poverty of the stimulus
 - 3.2 Universal grammar
 - 3.3 Parameters and universal grammar
 - 3.4 Language learning and language acquisition
 - 3.5 The generative linguist
- 4 The new comparative syntax**
 - 4.1 Principles and parameters: a recapitulation
 - 4.2 The *pro*-drop properties
 - 4.3 Relating the properties
 - 4.4 Agreement and *pro*-drop
- 5 Purpose and organization of the book**
 - 5.1 General purpose
 - 5.2 Organization
- 6 Exercises**

Introduction

The aim of this book is to offer an introduction to the version of generative syntax usually referred to as Government and Binding Theory.¹ I shall not dwell on this label here; its significance will become clear in later chapters of this book.

Government–Binding Theory is a natural development of earlier versions of generative grammar, initiated by Noam Chomsky some thirty years ago. The purpose of this introductory chapter is not to provide a historical survey of the Chomskian tradition. A full discussion of the history of the generative enterprise would in itself be the basis for a book.² What I shall do here is offer a short and informal sketch of the essential motivation for the line of enquiry to be pursued. Throughout the book the initial points will become more concrete and more precise.

By means of footnotes I shall also direct the reader to further reading related to the matter at hand. Much of the primary literature will be hard to follow for the reader who has not worked his³ way through the book, but I hope that the information will be useful for future reference.

¹ Chomsky (1991) himself expresses reservations about the label 'Government and Binding Theory' and refers to the theory we are concerned with here as the 'Principles and Parameters Theory'. The latter term is more comprehensive in that it covers work done in the Government and Binding tradition as developed in the present book, and also work done in a recent development in the generative framework usually referred to as the Minimalist Program (Chomsky, 1992). Since the label Government and Binding Theory or its abbreviation GB-theory is widespread we continue to use it here to refer to the generative work initiated by Chomsky's book *Lectures on Government and Binding* (1981a). The term allows us to distinguish the approach developed here from the more recent approach in the Minimalist Program.

² For a survey of the development of the theory see van Riemsdijk and Williams (1986). This work should be accessible once chapter 7 has been covered.

The reader will find a good introduction to generative grammar in general introductions to linguistics such as Akmajian, Demers and Harnish (1979), Fromkin and Rodman (1988, 1992), Lightfoot (1982), Smith and Wilson (1979), etc. These works should be accessible at this point. For more advanced introductions the reader is referred to Chomsky (1965, 1981a, b, c, 1982, 1986a, 1988, 1991), but reading them should be postponed until after chapter 7 of this book, at which point we shall have covered most of the technical issues that are discussed.

³ My use of the pronoun *his* for referents which may be either male and female follows the conventions of English grammar and I hope that the female readers of this book will not feel offended by it.

1 Linguistics: The Science of Language

When asked to indicate one prominent feature that distinguishes human beings from animals, many would probably say that this feature is 'language'. Even though animals may have communication systems, none of these systems is as rich or as versatile as the language used by humans. Language is human-specific.⁴ This means that an understanding of the mechanisms of human language may lead us to understand, at least partly, what it is that distinguishes human beings from animals. Linguistics, the study of language, gives us an insight into the human mind.

Leonard Bloomfield defined linguistics as the science of language (Bloomfield, 1935). Like all scientists, linguists will aim at formulating the general principles to account for the data with which they are faced. Linguists try to formulate generalizations about linguistic data, i.e. language.⁵

There are various ways of approaching the study of language. I assume the reader is familiar with the traditional view of language study, where the focus is often on the study of one specific language, say English. A linguist studying English will try to characterize the principles that determine the formation of English sentences. The goal will be to provide a systematic description of English sentence formation, the grammar of English. The description will have to account for data such as the following:

- 1a Agatha Christie has written many books.
 1b I don't like detective stories.

The sentences in (1) are well formed. They contrast with the sentences in (2), which are ill formed.

- 2a *Agatha Christie many books written has.
 2b *I detective stories like.

Well formed English sentences are constructed according to the grammar of English: they are **grammatical**. The sentences in (2) are not formed according

In their introduction to linguistics Akmajian, Demers and Harnish (1979) present a fairly comprehensive discussion of the differences between human language and animal language.

Robins (1967) and Newmeyer (1980, 1983) offer good surveys of the development of linguistics. These books will offer a broader background to situate the theory we are discussing here in its historical context.

to the grammar of English: they are **ungrammatical**, as indicated by the asterisks.

When writing a grammar, the linguist will not stop at merely listing examples with the appropriate grammaticality judgements. A simple catalogue of sentences may be an interesting basis for discussion but it cannot be the ultimate goal of scientific research. In addition to describing the data, the linguist will formulate general principles which will be applicable to further data. Informally, a linguist might account for the ungrammaticality of (2), for instance, by proposing that in English verbs precede their direct objects. A first hypothesis might be that English sentences are constructed according to the SVO pattern: subject precedes verb, verb precedes object. Let us call this the SVO hypothesis. Having formulated this hypothesis on the basis of a limited set of data, the linguist will test it on the basis of further data. The SVO hypothesis will lead him to predict, for instance, that (3a) and (3b) are grammatical; but as it stands, the hypothesis also predicts that (3c) and (3d) are ungrammatical: the objects, *detective stories* and *which stories* respectively, precede the subjects:

- 3a Jeeves is baking a cake.
- 3b John has bought a new car.
- 3c Detective stories, I don't like.
- 3d Which stories do you like?

Either the SVO hypothesis itself will have to be modified in the light of the data in (3c) and (3d) or one or more extra principles are needed which interact with the original hypothesis to account for the grammaticality of (3c) and (3d). We might, for example, formulate a rule of topicalization which moves a direct object to the beginning of the sentence to account for (3c). In addition we might formulate a rule for question formation which (i) moves the questioning element (*which stories*) to the initial position of the sentence, and (ii) inverts subject and auxiliary (*do*) (cf. (3d)).

The total of all the rules and principles that have been formulated with respect to a language constitutes the grammar of that language. A grammar of a language is a coherent system of rules and principles that are at the basis of the grammatical sentences of a language. We say that a grammar generates the sentences of a language.

A first requirement for any grammar is that it provides a characterization of the language it describes, i.e. the grammar must be able to distinguish those strings of words which are sentences of the language from those which are not sentences of the language in question. Such a grammar will be **observationally adequate**.

2 The Native Speaker: Grammaticality and Acceptability

2.1 Descriptive Adequacy

Not only linguists have the ability to judge English sentences. Every native speaker of English knows intuitively that the sentences in (1) and (3) are acceptable and that those in (2) are not. Moreover, every native speaker of English produces a large number of grammatical sentences and understands the English sentences that he comes across. The native speaker may not be able to formulate the general principles that underlie the sentences he produces, but he has an unconscious or tacit knowledge of such principles; he has internalized a grammar of the language. The native speaker's tacit knowledge of the grammar of his language is the focus of enquiry for the linguist working in the Chomskian tradition. We say that a grammar reaches descriptive adequacy if, in addition to describing the data, it provides an account for the native speaker's intuitions.

Let us consider some examples. We have proposed that (3c) and (3d) could be generated by a process that moves the direct object leftward to the beginning of the sentence. Now consider the examples in (4), which are not acceptable (hence the asterisk):

4a *Detective stories, I wonder if he likes.

4b *Where do you wonder if he lives?

To account for the unacceptability of (4a) we might propose that the process which moves the direct object in (3c) must be constrained: the direct object cannot move across *if*.

Similarly, when we consider (4b) we might propose that the rule of question formation must also be constrained: the questioning element (*where*) must not move across *if*. At this point we have reached **observational adequacy**: we provide a description of the facts. However, if we stop at this point we are missing a **significant generalization**. The ungrammaticality of (4a) and (4b) is due to the same constraint. A **descriptively adequate grammar** will not simply provide an analysis for (3c) and (3d) and for the deviance of (4a) and (4b), but it will try to capture the relation between (4a) and (4b) and formulate a general principle to explain why both (4a) and (4b) are felt to be unacceptable. Such a principle may be that no element in English must be moved across *if*. This general principle will also lead us to predict that the examples in (5) are ungrammatical, whereas those in (6) are grammatical:

- 5a *Where do you wonder if Emsworth has hidden the Empress?
 5b *Which detective do you wonder if Emsworth will invite for Sunday lunch?
 5c *To Bill, I wonder if he will give any money.
- 6a Where has Emsworth hidden the Empress?
 6b Which detective will Emsworth invite for Sunday lunch?
 6c To Bill, he won't give any money.

The general constraint which blocks movement of an element across *if* will be taken to be part of the native speaker's internal grammar.

A descriptively adequate grammar will not only describe the linguistic data, but it will contain the general principles and processes that enable the native speaker to produce and interpret sentences in his language and decide on the acceptability of sentences. Such a grammar is an explicit formulation of the tacit linguistic knowledge of the native speaker, his internal grammar.

The shift of focus from language itself to the native speaker's knowledge of language is the major feature of the Chomskian tradition. Both the generative linguist and the traditional linguist will be constructing grammars, i.e. general systems that underlie the sentences of a language. But the generative linguist conceives of his grammar as a reflex of the native speaker's competence. The grammar is a representation of the speaker's internal linguistic knowledge.

2.2 Grammaticality and Acceptability

At this point we turn to the notions of 'grammaticality' and 'acceptability'. 'Grammaticality' is a theoretical notion. A sentence is grammatical if it is formed according to the grammar of English as formulated by the linguist. 'Acceptability', on the other hand, is the term which characterizes the native speaker's intuitions about the linguistic data. Consider (7):

- 7a Bill had left. It was clear.
 7b [That Bill had left] was clear.
 7c It was clear [that Bill had left].
 7d Once that it was clear [that Bill had left], we gave up.
 7e Once that [that Bill had left] was clear, we gave up.

(7a) contains two independent sentences. In (7b) the bracketed sentence *Bill had left* is the subject of the complex sentence *that Bill had left was clear*.

We say that *Bill had left* is a subordinate clause. It is introduced by *that*, a subordinating conjunction. Similarly, in (7c) *that Bill had left* is a subordinate clause. In (7d) the sentence (7c) is a subordinate clause in a complex sentence. A grammar must generate complex sentences in which one clause is part of another one.

Let us turn to (7e). The sentence is odd for most native speakers: it is not acceptable. However, this sentence is formed according to the same principle that we posited to account for the formation of (7b)–(7d), i.e. that one sentence may become part of another sentence. Hence (7e) would be **grammatical**, though it is not acceptable.

Faced with intuitions such as that for (7e) the linguist might decide to modify the grammar he has formulated in such a way that sentence (7e) is considered to be ungrammatical. He may also decide, however, that (7e) is grammatical, and that the unacceptability of the sentence is due to independent reasons. For instance, (7e) may be argued to be unacceptable because the sentence is hard to process. In the latter case the unacceptability is not strictly due to linguistic factors but is due to the more general mechanisms used for processing information.

The native speaker who judges a sentence cannot decide whether it is grammatical. He only has intuitions about acceptability. It is for the linguist to determine whether the unacceptability of a sentence is due to grammatical principles or whether it may be due to other factors. It is the linguist's task to determine what it is that makes (7e) unacceptable. This entails that there may be disagreement between linguists as to whether certain unacceptable sentences are grammatical or not. The disagreement is not one of conflicting judgements of the sentence (although these may also exist), but it is one of analysis. The linguist will have to determine to what degree the unacceptability of a sentence is to be accounted for in terms of the grammar. All the linguist has to go by, though, is the native speaker's intuitions about language, and these, as argued above, are the result of the interaction between his internal grammar and other factors.

In this book we focus on the linguistic knowledge of the native speaker. We restrict our attention to his internal grammar. Obviously, the interaction between the grammar and other mental processes is also an interesting area of research, but it is not the topic of this book.

2.3 The Grammar as a System of Principles

One approach to formulating a grammar of a language would be to suppose that the speaker's internal knowledge of English, i.e. his internal grammar, is no more than a huge check-list of grammatical sentences. Speakers could be

thought to 'check' any sentence they come across against this internal inventory. Sentences which match a sentence in the list would be said to be grammatical, those that do not are ungrammatical. Depending on the degree of deviance of such ungrammatical sentences we could rank the sentences for ungrammaticality. A grammar of a language would then be simply a list of sentences. But it must be immediately obvious that listing all the grammatical sentences of a language is an impossible task and also that it misses the point.

Cataloguing all the grammatical sentences of English is first of all impossible because there is an infinite number of English sentences. In addition, there are other objections to such a listing enterprise. We stated above that linguistics is the scientific study of language. From such a perspective the listing of linguistic data is not enough. We expect general principles to explain the data.

For the generative linguist who tries to provide a representation of the native speaker's internal knowledge of a language a mere listing of sentences would never achieve descriptive adequacy: it could never account for the native speaker's knowledge of the language. Human beings – in our example speakers of English – have finite memories: we often forget things we have heard. Given that the capacity of our memories is finite, it would be absurd to claim that human beings are able to store all potential sentences of the language, an infinite set. It is thus inconceivable that the native speaker's internal linguistic knowledge is an inventory of sentences. We must assume that human beings are somehow equipped with a finite system of knowledge which enables them to construct and interpret an infinite number of sentences. This finite system of principles is what we referred to loosely above as the internal grammar of the language. The generative linguist will try to render explicit the finite system of principles that make up the native speaker's competence. In our example, the principle which prohibits moving elements across *if* will be able to account for the unacceptability of (4) and (5).

3 Knowledge of Language

3.1 *The Poverty of the Stimulus*

A speaker's knowledge of a language is largely unconscious. It is formally represented as a grammar. The grammar of a language generates the sentences of a language and assigns to each sentence a set of representations which provide the formal characterization of some of the properties of the sentence (semantic, syntactic, morphological, phonological, etc.). It is the linguist's task to render explicit the internal grammar of the speaker of a

language. In order to construct such an explicit grammar of a language, the linguist can rely to some extent on data taken from usage, the output of the speakers. However, usage data are inevitably an incomplete source of information. The sentences actually produced by a speaker are only a fragment of the sentences he *could* have produced. In order to arrive at a characterization of the speaker's potential, the linguist can also rely on the speaker's knowledge of the language, i.e. on his capacity to evaluate linguistic expressions in that particular language. For instance, speakers of English intuitively know that (8a) is an acceptable sentence and that (8b) is not:

- 8a She has invited Louise to her house.
 8b *Has invited Louise to her house.

Informally we will say that (8b) is unacceptable because the subject is missing. For some reason, to which we return in more detail in chapter 8, the grammar of English requires that finite sentences like (8a) have an overt subject. The grammar of Italian differs from that of English, as seen in (9). In (9a) the subject of *ha invitato* is expressed, in (9b) it is not realized:

- 9a Lei ha invitato Louisa a casa.
 she has invited Louisa at home
 9b Ha invitato Louisa a casa.

We will achieve descriptive adequacy if our grammar is able to provide an explicit characterization of the general principles of sentence formation in English. This grammar will, for instance, impose the overt realization of the subject pronoun in (8b).

Now another important and fascinating question arises: we would like to understand how native speakers of a language, in our example English, come to possess the knowledge of their language. We say that a theory reaches **explanatory** adequacy if it can account for the fact that the principles of the internal grammar can get to be known by the speakers, i.e. if it can account for language acquisition.

The problem of language acquisition has often been summarized in terms of the problem of the **poverty of the stimulus**. Our linguistic capacity, for instance our knowledge of English, goes beyond the evidence we have been exposed to in our childhood. The linguist wants to account for the fact that the linguistic competence is attained in spite of important inadequacies in the stimulus, the linguistic experience. Three types of inadequacies are standardly

referred to in the literature. First, we do not just come across grammatical sentences: everyday use of language contains slips of the tongue, hesitations, incomplete sentences, etc. Second, the experience, i.e. the stimulus, is finite, and we end up being able to produce and process an infinite number of sentences. Third, we acquire knowledge about our language for which we have no overt or positive evidence in the experience. For instance, consider the following sentences:

10a I think that Miss Marple will leave.

10b I think Miss Marple will leave.

11a This is the book that I bought in London.

11b This is the book I bought in London.

12a Who do you think that Miss Marple will question first?

12b Who do you think Miss Marple will question first?

On the basis of the examples in (10)–(12) the child learning English might well conclude that the conjunction *that* is optional; the data in (10)–(12) suggest that *that* can always be present and that it can always be absent. However, this conclusion would not be correct:

13a *Who do you think that will be questioned first?

13b Who do you think will be questioned first?

In the sentences in (13), the conjunction *that* must not be present. It is hard to see how the child can infer this from evidence to which he is exposed. Observe also that children are not explicitly taught that (13a) is ungrammatical. The problem can be summarized by saying that there is a gap between the data we are exposed to, the **input**, and our knowledge we achieve, the **output**; the stimulus underdetermines the knowledge we ultimately attain. This means that we cannot simply represent the acquisition of knowledge of language in terms of the schema (14a). The **triggering experience**, i.e. exposure to linguistic data, is not sufficient to allow a child to construct the grammar of his language.

14a Exposure

Triggering experience → Grammar of X

3.2 Universal Grammar

Given that neither formal teaching nor overt evidence seems to be the source of the native speaker's intuitions, it is proposed that a large part of the native speaker's knowledge of his language, i.e. the internal grammar, is innate. The idea is that human beings have a genetic endowment that enables them to learn language. It is this innate capacity for language learning common to all human beings that the generative linguist tries to characterize. Of course, it would be unreasonable to posit that some individuals – those that will become native speakers of English – are born with a specific grammar of English and that others – those that will end up speaking Italian as their first language – are born with the grammar of Italian readily stored in their minds. Human beings with normal mental faculties are able to learn any human language. The innate linguistic endowment must be geared to any human language and not to just one.

Let us discuss some examples informally in order to provide an outline of the proposal. We have introduced one generalization about English: the SVO hypothesis. The data in (7) lead us to formulate another hypothesis: any grammatical English sentence can apparently be embedded and become a subordinate clause in a complex sentence. Let us refer to this as the embedding principle.

15 Embedding principle⁶

A grammatical sentence can become a subordinate clause in a complex sentence.

The embedding principle tries to render explicit part of the tacit knowledge of the native speaker. This principle would be taken to be part of the grammar of English, hence available to the native speaker. But this principle is not one that is particular to the grammar of English, it is not language-specific. Rather, the embedding principle is part of the grammar of all human languages. Thus in French too we find sentences such as (16a) embedded in (16b):

16a Maigret a abandonné l'enquête.

Maigret has abandoned the enquiry.

16b Lucas a annoncé que Maigret a abandonné l'enquête.

Lucas has announced that . . .

As the reader will see later, the embedding principle is not in fact part of our grammar. The fact that sentences can be embedded can be deduced from the principles of sentence formation discussed in chapters 1 and 2.

Readers familiar with other languages will be able to check that the embedding principle applies in those languages too.

The embedding principle is a universal principle. Principles that hold of all languages are said to be part of universal grammar, or UG for short. Informally, UG is a system of all the principles that are common to all human languages, this means languages as different as English and Italian or Japanese.

A hypothesis adopted by generativists of the Chomskian tradition is precisely that universal grammar is innate to the human species. UG is a genetic endowment: we are born equipped with a set of universal linguistic principles. To quote Chomsky himself: 'Universal grammar may be thought of as some system of principles, common to the species and available to each individual prior to experience' (1981b: 7).

If we assume that there is such an innate linguistic endowment the task of attaining the knowledge of a specific grammar, say English, is facilitated. Someone learning English would not have to learn the embedding principle. It is innate; it is part of the genetic endowment.⁷

Universal grammar is the basis for acquiring language. It underlies all human languages. All and only human beings are equipped with UG and they are all able to learn languages. Other systems (say, dogs or television sets) are not equipped with UG and therefore will not be able to learn human languages. The linguistic endowment characterized as UG is species-specific.

3.3 Parameters and Universal Grammar

The innate linguistic endowment UG is not sufficient to enable us to speak a language. If all that is needed was UG then human beings would be able to speak any language wherever they were born and in whatever circumstances they grew up. The native language is that spoken by the child's immediate environment. It would be inconceivable, for instance, that a child growing up in a community where only English is spoken could become a native speaker of Japanese. Human beings usually master one language with native competence and they have a hard time learning other languages later in life. It is a well-known fact that achieving complete mastery of second or third languages in adulthood is exceptional.

While certain grammatical principles are universal, there is 'also a lot of variation between different languages. The grammar of English differs in

The reader may wonder why, if the principle is innate, children do not start using complex sentences straight away. However, it is conceivable that the development of the internal grammar interacts with a general maturation process. We leave this problem aside here.

important respects from that of, say, Japanese. Hence, if you 'know' the grammar of English, this will not entail that you 'know' the grammar of Japanese. In (1) we illustrated some simple English sentences and we saw that English sentences exhibit SVO word-order. In Japanese, on the other hand, the object precedes the verb; Japanese is SOV:

- 17 John-ga Mary-o but-ta.
 John-particle Mary-particle hit-past
 (Kuno, 1973: 3)

English and Japanese are similar in that sentences contain elements such as subjects, objects and verbs. But they differ in the way these elements are ordered linearly. The SVO hypothesis, which we postulated as part of English grammar, cannot be an absolute linguistic universal: it is part of the grammar of English (and of other languages) but not of that of Japanese. It is language-specific. How does a child learn that English has the SVO pattern? We could envisage the following scenario. The linguistic endowment UG makes available, among other things, the notions 'subject', 'object', 'verb'. Let us propose for the sake of the argument (cf. chapter 2, for a different view, though) that these are universal concepts, available in all human languages. Subject, verb and object will have to be linearly ordered. When learning a language the child will have to decide which is the word-order characteristic of his language. One option is to say that in fact word-order variation between languages is due to a primitive difference between these languages: it is a **parameter** along which English and Japanese vary. Languages could be said to vary with respect to the word-order parameter: UG provides the binary choice OV or VO, and individual languages opt for one **setting** of the parameter or another. We might say that the different word-orders of English and Japanese are directly correlated with the word-order parameter: English has the setting where the object follows the verb, Japanese has the opposite setting for the parameter. The child learning English will have to fix the parameter for the VO setting, the child learning Japanese will have to fix the parameter for the OV setting. For each case exposure to transitive sentences in the language should enable the child to perform the setting.

Other ways of accounting for word-order variation may come to mind. The reader may recall that we suggested that the sentence-initial position of the direct object in (3c) and in (3d) above were due to a fronting operation which moves the object leftward. It is then in fact conceivable that the same kind of leftward movement could be invoked to account for the word-order found in Japanese. Say, for instance, that we propose that UG initially makes only one order available for a verb and its objects, namely the VO order. It

could then be said that in Japanese a movement operation can shift the object to the left across the verb, resulting in the ordering OV. We have seen that we need such movement operations independently. The parameter distinguishing English and Japanese would then be expressed in terms of the availability of a particular leftward movement which can move the object to a position between the subject and the verb. Again the child who is learning Japanese will have to determine that the movement is available in Japanese, while the child learning English would assume that it is not.

Whichever option is chosen to account for word-order variation – and the debate is still very much open, we return to it in chapter 2 – the child learning a language must construct an internal grammar for that language. To achieve this task he uses, on the one hand, the universal notions and principles of UG and the choices that it makes available, and on the other hand he uses the data of his linguistic experience, in our example the English sentences he hears. Sentences such as those in (1) will provide evidence to the child that in English subject precedes verb and verb precedes object. A sentence such as that in (17) will enable the child exposed to Japanese data to decide that Japanese has SOV.

Exposure to linguistic material is an essential ingredient in the child's learning process. The child will need the linguistic experience to start constructing the internal grammar of his language and thus to attain the knowledge of a language. Without exposure the child would not be able to construct his internal grammar. UG is crucial in the organization of the primary linguistic experience. UG guides the way the child will interpret and organize the language he is exposed to. We have now postulated two properties of UG:

- (i) UG contains a set of absolute universals, notions and principles which do not vary from one language to the next.
- (ii) There are language-specific properties which are not fully determined by UG but which vary cross-linguistically. For these properties a range of choices is made available by UG.

Absolute universal principles are rigid and need not be learnt. But even with respect to the mastery of language-specific properties very little 'learning' is involved under the hypothesis outlined above. For those principles that are parametrized, the options available are determined by UG. Attaining linguistic knowledge consists in fixing the parameters.

From this point of view, we conclude that the mastery of a language is not really the result of learning. Rather, being equipped with UG (with its parameters) and exposed to a language, the child cannot but construct the grammar of the language he is exposed to. For this reason the term 'learning' is often replaced by the term 'acquisition'.

In addition, the exposure to language will also equip us with a vocabulary, the words of the language to which we are exposed. Even if we have an innate knowledge of the principles of language we must inevitably learn the lexicon of the language, the words and their meaning, in order to be able to put this knowledge into operation. Thus an English child will have to learn all the words in the sentences above, and indeed many more. And we go on learning new words throughout our lives. Similarly a French child will learn the French lexicon, etc.⁸

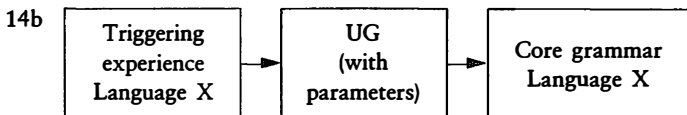
To sum up: human beings are born equipped with some internal unconscious knowledge of grammar: UG. UG is a set of universal principles of language, some of which are rigidly fixed, some of which parametrized. Via the input of the experience of one particular language this knowledge can be implemented. The acquisition process is 'triggered' by the exposure, the child's linguistic experience.

Exposure will also enable the child to learn the vocabulary of the language.⁹ The view of language acquisition in terms of parameter setting is the basis of current work in the generative tradition. The theory is sometimes referred to as the 'Principles and Parameters Theory' (cf. fn. 1).

3.4 *Language Learning and Language Acquisition*

Our ability to speak a language is based partly on the innate principles and parameters available in UG, partly on the triggering experience of exposure to a specific language. On the basis of these components we develop a grammar of one (or more) specific languages: the core grammar of such a language.

Schematically we can represent the generative view of language acquisition as follows:



⁸ The acquisition of the vocabulary of a language is also a matter of interest. For some introductory discussion the reader is referred to Lightfoot (1982: 121–2).

⁹ The reader will find interesting discussion of language acquisition for instance in Deprez and Pierce (1993), Hermon (1992), Lightfoot (1981, 1982, 1989, 1991, 1993), Radford (1990), Wexler and Manzini, (1987). For more general discussion see also Chomsky (1981a, b, c), and the literature cited there. Most of these references might be hard to read at this stage and the reader is advised to postpone reading these works until he has worked through chapters 1–7 of this book.

The exposure to some language, say English, will activate the innate principles of universal grammar. The child will fix the choices to be made for the language in question, for instance, that the object follows the verb. He will also learn the vocabulary of the language. To quote Chomsky:

Endowed with these principles, a system provided with adequate experience will develop a grammar of the peculiar and specific sort characteristic of human language . . . Lacking these principles, a system will develop no grammar or some different system. The telephone exchange, for example, has 'heard' much more English than any of us, but lacking the principles of universal grammar . . . it develops no grammar of English as part of its internal structure. (1981b: 8)

By the age of six a child exposed to English will have constructed the grammar of his language. This does not mean that no further development of his knowledge of language is possible. For instance, we go on learning new words throughout our lives. In addition we also learn certain less usual constructions of the language. These exceptional or marked patterns of the language are not taken to be part of the core grammar of the language, they belong to the **marked periphery** of the grammar and may be acquired later. The native speaker will also have to learn all of the social or cultural conventions associated with his language, for instance, that certain words belong to a very high style whereas others are informal. These conventions are not part of the grammar, they belong to the more general domain of human behaviour.

The aim of generative syntacticians is to develop a theory of language that is a model of the acquisition of language. Linguists want to provide an explicit formulation of the three components of (14b): (i) the principles of UG and the parametric variation across languages; (ii) the triggering experience needed to activate the principles of UG; and (iii) the core grammar of specific languages as it derives from these interacting components. A theory that can account for these three components will be said to have reached explanatory adequacy.

3.5 *The Generative Linguist*

The research programme as sketched here briefly and roughly is one that has been motivating linguistic research for the past thirty years and has given rise to many challenging results. The programme is indeed still developing.

It may be useful to repeat that the ultimate aim of generative linguistic

theory is not to describe the details of one specific language, but rather to formulate the underlying principles that determine the grammars of human languages. These grammars are seen as representations of the native speaker's knowledge. In the course of their enquiry, linguists will examine data drawn from individual languages, of course, but the investigator will always bear in mind the interacting components in (14b).

The generative linguist who tries to characterize knowledge of a language, say English, will wish to do two things: (i) he needs to determine what properties of English are universal; and (ii) what properties are English-specific and how these relate to the parameters of UG.

It must by now have become clear that by simply looking at English and only that, the generative linguist cannot hope to achieve his goal. All he can do is write a grammar of English that is observationally and descriptively adequate but he will not be able to provide a model of the knowledge of the native speaker and how it is attained. The generativist will have to compare English with other languages to discover to what extent the properties he has identified are universal and to what extent they are language-specific choices determined by universal grammar. Even when his main concern is some aspect of the grammar of English the linguist will have to go outside this one language and engage in contrastive work.

Work in generative linguistics is therefore by definition comparative. Generative linguists often do not focus on individual languages at all: they will use *any* human language to determine the general properties of UG and the choices it allows. Data from a dialect spoken by only a couple of hundred people are just as important as data from a language spoken by millions of people. Both languages are human languages and are learnt in the same way.

4 The New Comparative Syntax

4.1 *Principles and Parameters: A Recapitulation*

When we look at the development of generative syntax in the last twenty-five years one important tendency that can be isolated is a marked return to comparative approaches. The comparative approach is obviously not the creation of generative grammar: it finds a clear precedent in the nineteenth-century comparative approaches to language study (cf. Robins, 1967).

The main goal of nineteenth-century comparative grammar was historical, i.e. that of establishing relations of parenthood and kinship across languages. The goal of the comparative approach in the generative tradition is psychological, i.e. that of accounting for the knowledge of language. As we

have already seen, the following questions are asked: (i) What is knowledge of language? (ii) How is it acquired? The latter question focuses on the issue of how much of our linguistic knowledge is determined by experience and how much is due to a predetermined mental mechanism (cf. (14b)).

In order to determine how a specific language (say English) is acquired and how language in general is acquired we have to determine to what extent the properties of languages vary from one language to another, i.e. to what extent the properties are language-specific, and to what extent they are invariant across languages. Properties of language that vary cross-linguistically will be learnt by the speaker as a result of exposure to some specific linguistic environment: the fact that Italian allows the subject pronoun to be absent can be learnt through exposure to this language, for instance. Speakers who are repeatedly confronted with subjectless sentences such as (9b) will be able to infer that in the language they are exposed to the subject can be omitted. On the other hand, properties which are shared by all languages might well be taken to be part of UG, the predetermined linguistic competence of the human mind. Comparative studies of languages will play a crucial role towards providing us with answers to these questions, i.e. what is a universal and what is language-specific. In the present section we focus on the parametric variation between languages and try to clarify the notion of parameter.

Parameters are postulated to explain cross-linguistic variation. We should not assume, though, that each observed difference between one language and another corresponds to one parameter. The comparative study of languages has revealed that the properties with respect to which languages vary tend to organize themselves in clusters which are stable across languages and which allow us to arrive at a typology of languages. If a language has property X, it will also have property Y and property Z. The parametric approach will have to explain why certain properties co-occur.

4.2 *The Pro-drop Properties*

In order to illustrate this let us look at one of the better known parameters which has been postulated to account for the difference between English (8) and Italian (9). Recall that Italian differs from English in that the former, though not the latter, allows the subject of a finite clause to remain unexpressed. The parameter which distinguishes languages like English which do not allow a subject pronoun to be omitted and those like Italian which do is referred to as the *pro-drop* parameter. (For detailed discussion see, among others, Rizzi, 1982a, 1986a; Jaeggli and Safir, 1989.) Italian is a *pro-drop* language, English is not. That the subject pronoun can be omitted is not the only property to distinguish *pro-drop* languages like Italian from non-*pro-drop*

languages like English. In Italian, the overt subject can occupy a post-verbal position; this option is not available in English:

18a È arrivato Gianni.
is arrived Gianni
'Gianni has arrived.'

18b *Is arrived John.

19a Ha telefonato sua moglie.
has telephoned your wife
'Your wife has phoned.'

19b *Has telephoned your wife.

In Italian a subject of a subordinate clause can be moved to the main clause domain across the overt conjunction *che*, corresponding to *that*; in English this is not possible: if a subject is moved then the clause from which the subject has been moved cannot be introduced by a conjunction (cf. the discussion of (13) above). The correlation between the data in (20) and the *pro*-drop phenomenon is due to Perlmutter (1971).

20a Chi credi che abbia telefonato?
who believe (2sg) that have (subj) telephoned
'Who do you think has called?'

20b *Who do you think that has telephoned?

20c Who do you think has telephoned?

In Italian subjects of weather verbs such as *rain* are necessarily omitted, in English such subjects must be realized by a pronoun.

21a (*Ciò) piove.
(it) rains (3sg)
'It is raining.'

21b *(It) is raining.

Consider now the following:

22a Che Louisa non partirà è chiaro.
that Louise not will leave is clear
'That Louisa will not leave is clear.'

22b That Louise will not leave is clear.

In (22) the subordinate clauses *che Louisa non partirà* and *that Louise will not leave* function as the subjects of the sentences. The sentences can be paraphrased if we move the subordinate clause to a final position: in Italian the position vacated by the moved sentence remains empty, it cannot be blocked up by a pronominal element as illustrated in (23a). In contrast, in English we must stick in a pronoun *it* to fill the vacated subject position, as shown in (23b).

- 23a (*Ciò) è chiaro che Louisa non partirà.
 it is clear that Louisa will not leave
 23b *(It) is clear that Louisa will not leave.

These contrasts listed above are not autonomous properties of the languages in question, all can be related to the option which allows the subject pronoun to be omitted in Italian.

4.3 Relating the Properties

We started from the empirical observation that the subject pronoun can apparently be omitted. Observe that in Italian, the pronominal subject *can* also be overt; the overt realization of the subject pronoun has some semantic or pragmatic effect: for instance it signals contrast or it focuses on the subject:

- 24 *Lei parte e lui arriva.*
 she leaves and he arrives

When no contrast or no special focus on the subject is needed the pronoun is absent. This could be derived from some general consideration of economy: we might say that the non-expression of the subject pronoun requires less effort than when the pronoun is present, and that therefore the subject will only be present when the added effort of the overt expression has some yield. Subject pronouns appear only when it is impossible to leave them out.

The obligatory absence of the subject pronoun of weather verbs in (21a) can be related to the principle of economy suggested above: it is hard to see how a subject of a weather verb could have a contrastive function. This means that there will never be a reason to use the pronoun in Italian. A similar approach can be suggested for (22). When we move the subject clause in Italian the vacated position can be empty and it has to remain empty. Why should this be? We have already seen that the subject position in Italian need

not be filled, it can be empty. In English the subject position cannot be empty so we stick in a pronoun. *It* in the English example (23b) does not contribute anything to the meaning of the sentence, *it* cannot be contrasted or focused. But in Italian, subject pronouns are only used with a contrastive or emphatic function, so there will never be any motivation for inserting a pronoun in the Italian equivalent of (23b), (23a).

Let us turn to the examples with post-verbal subjects, (18) and (19). All English sentences must have subjects. This does not mean, though, that the subject must necessarily be a referential expression, as the following example illustrates:

25 There arrived three more students.

In (25) the subject position is occupied by the element *there*. *There* is related to an indefinite post-verbal subject. Let us say that *there* fills up the position vacated by an indefinite subject (we return to this in chapters 2 and 9). The essential point is that *there* cannot be contrastive or emphatic in (25). In the Italian examples in (18a) and (19a) we also have a post-verbal subject. Since in general Italian does not need a full pronoun to occupy the vacated subject position (23a), we do not need a filler for the subject position in such examples as (18) and (19).

The data in (20) might at first sight seem puzzling. It is generally accepted that one cannot move a subject from the position to the immediate right of the conjunction (*that* in English); (20b) suggests that this is possible in Italian. However, we cannot base our judgements on a superficial comparison of two sentences in two languages. We need to consider the way these sentences are formed, their derivation. On the basis of the data in (18) and (19) we are led to conclude that the subject NP in Italian may appear either pre-verbally or post-verbally. Hence (20b) has two possible derivations, schematically represented in (26):

26a Chi credi che — abbia telefonato?

26b Chi credi che abbia telefonato —?

In the representation (26a) *chi* originates in the position to the immediate right of *che*, in (26b) it originates in the post-verbal position, a position also available for subjects, as seen in (20b). Now it is known that in Italian, as in English, nothing bans the leftward movement of post-verbal material across a conjunction.

- 27a Who do you think that John will invite —?
 27b Chi credi che Gianni inviterà —?

The general principle which bans extracting material from a position to the immediate right of a conjunction can now be maintained for the grammar of English AND for the grammar of Italian. In the Italian sentences where this principle would appear to have been violated, the language uses the alternative derivation whereby the subject is moved from a post-verbal position.

The correlations established here for the contrast between a *pro*-drop language like Italian and a non-*pro*-drop language like English can extend straightforwardly to Spanish, for the first group (28), and French for the second (29);

28 *Spanish*

- 28a Baila bien.
 dances (3sg) well
 'He dances well.'
- 28b Llego Maria ayer a los doce.
 arrived Maria yesterday at noon
 'Mary arrived at noon yesterday.'
 (Jaeggli, 1981: 139)
- 28c ¿Quién dijiste que vino.
 who did you say that came
 'Who did you say came?'
 (Jaeggli, 1981: 145)
- 28d Me parece que Juan tiene hambre.
 me seems that Juan has hunger
 'It seems to me that Juan is hungry.'
 (Jaeggli, 1981: 146)

29 *French*

- 29a *(Elle) dance bien.
 (she) dances well
 'She dances well.'
- 29b *Arrivait Marie hier a midi.
 arrived Marie yesterday at noon
- 29c *Qui dis-tu que viendra?
 who say you that will come
- 29d *(Il) me semble que Jean a faim.
 (it) me seems that Jean has hunger
 'It seems to me that Juan is hungry.'

4.4 Agreement and Pro-drop

The reader may observe that the possibility of omitting a pronoun subject correlates with another property of the languages examined and which is particularly obvious when we compare English and Italian. If we look at the present tense paradigms for the verb inflection for these languages we observe a striking contrast:

30	<i>English</i>	<i>Italian</i>
	1sg I speak	io parlo
	2sg you speak	tu parli
	3sg she speaks	lei parla
	1pl we speak	noi parliamo
	2pl you speak	voi parlate
	3pl they speak	loro parlano
	2 forms	6 forms

In the case of Italian, every number/person combination has a different ending; as a result the inflectional paradigm distinguishes all six persons uniquely. There is no possibility of confusion: the ending of the verb immediately identifies the subject. One could say that such inflectional systems are rich. In contrast, the English system has only one distinctive form, that for the third person singular; all other persons are unmarked morphologically, the bare stem is used, which is also identical to the imperative and to the infinitive. In the literature, an attempt is made to correlate the inflectional paradigm of the language with the *pro*-drop parameter (cf. Jaeggli and Safir (1989), Rizzi (1986a), Taraldsen (1980)). Languages which have rich inflection are often *pro*-drop languages.¹⁰ Intuitively this correlation is expected: when the verb inflection is rich we can recover the content of the subject by virtue of the inflection and the pronoun would not add information. In languages with poor inflection the verb inflection does not suffice to recover the content of the subject and the pronoun is needed. We return to this issue in chapter 8. The inflectional system of French is relatively poor and French is not a *pro*-drop language; Spanish is a *pro*-drop language and has rich inflection.

The approach above suggests that a number of properties of languages and

¹⁰ Gilligan (1987) studies a sample of 100 languages from various language families and reports 76 languages with agreement which allow for the subject pronoun to be absent, against 17 languages without agreement and which allow the subject to be absent.

language types can be reduced to a unique elementary difference between their grammatical systems. The analysis of the *pro*-drop parameter, originally developed on the basis of Romance languages in the late 1970s, has led to what we can refer to as the new comparative syntax. A related development is the study of dialect variation, which has become strongly prominent in the 1980s; another promising line of research is that in the area of historical syntax. Diachronic developments of languages are interpreted again in terms of the Principles and Parameters model, diachronic changes consisting in re-settings of one or more parameters (cf. Lightfoot, 1979, 1991). In this book, the comparative approach is more prominent in chapter 8, which discusses non-overt elements; in chapter 9, where we discuss cross-linguistic variation in question formation, and in chapter 11, which concerns verb movement.

5 Purpose and Organization of the Book

5.1 *General Purpose*

In this book I provide a survey of some of the main results of generative research over the past thirty years. The book is not meant for the absolute beginner. The reader is expected to have some background in linguistics, specifically in syntax. He should, for instance, be able to parse sentences and be familiar with the tree diagram representation, and with the basic terminology of syntax. Notions such as sentence, clause, noun, verb, subject, object, etc., are presupposed. I assume therefore that the reader has had some introductory course to syntax or that he has read some introductory works.¹¹ However, in order to guarantee that we have a common starting-point, I shall often recapitulate the basic notions. It will also be shown how traditional concepts are used and reinterpreted within the generative framework.

The aim of the book is to offer a general introduction. I shall not go into all the complexities and details of ongoing research. Rather, I wish to familiarize the reader with the basic concepts used. I hope that the book will encourage the reader to turn to the primary literature himself and discover some of the more intricate problems. The references in the footnotes will provide indications for further reading.

Although the examples in the book will be taken primarily from English,

¹¹ I am thinking of works such as Akmajian and Heny (1975), Akmajian, Demers and Harnish (1978), Burton-Roberts (1986), Fromkin and Rodman (1988), Huddleston (1976), Jacobs and Rosenbaum (1970), Smith and Wilson (1979), Wekker and Haegeman (1985) to mention only a few.

this book is not a grammar of English. English is used as just one example of human language and we shall often discuss other languages. We shall try to decide what sort of internal grammar native speakers of English have at their disposal and to determine what it is that makes a sentence acceptable or unacceptable, what sort of grammatical principles can be advocated and to what extent these are universal or language-specific. In some sense we are like linguistic detectives. The linguistic data are like the clues a detective is given when starting his enquiry. He has to piece these data together, construct hypotheses, check these and ultimately he may discover the explanation for the evidence he has assembled. To remind the reader of this task I have chosen to illustrate the data with examples in which literary detectives play a prominent role. At the end of the book I hope that the reader will have become a competent linguistic detective himself.

5.2 Organization

The book is divided into twelve chapters. The first ten chapters provide the basic outline of the theory. The last two chapters highlight some recent developments of the theory. Each chapter is followed by a one-page summary and by a set of exercises. The exercises have a dual purpose. First, they will enable the reader to check if he has understood and assimilated the basic concepts introduced in the chapter. The empirical range of the discussion is broadened: many exercises will include a discussion of data drawn from languages other than English.

Second, the exercises will be used to draw the reader's attention to theoretical or empirical problems not touched upon in the chapter. Often a problem introduced by way of an exercise in an earlier chapter is then picked up in the discussion of a later chapter. Alternatively, the exercises will direct the reader to areas for further reading or for further research.

Footnotes will mainly be used to direct the reader to further reading. The footnotes will also indicate at which point in the book the reader should be able to tackle the literature in question.

6 Exercises

Exercise 1

Consider the following sentences. None of them is fully acceptable but they vary in their degree of deviance. If you are a native speaker of

English try to rank the sentences for acceptability. Wherever you can, try to construct an acceptable sentence modelled on the one you are judging. If you are not a native speaker of English you may attempt to carry out the task described above but it may be difficult. Another way of approaching this exercise is to ask some native speakers to do the exercise and compare their answers.

- 1 Which man do you know what John will give to?
- 2 Which man do you wonder when they will appoint?
- 3 Who do you wonder which present will give?
- 4 Which present do you wonder who will give?
- 5 Which man do you wonder whether John will invite?
- 6 Which man do you wonder whether will invite John?
- 7 Which man do you wonder what will give to John?
- 8 Which man do you wonder when will invite John?

Native English speakers are basically in agreement on the ranking of sentences 1–8. The judgements formulated are not the result of formal tuition. English grammar classes do not pay attention to sentences like 1–8. It is quite likely that speakers have never come across such sentences. In other words, they have not acquired the intuitions on the basis of overt evidence. On the contrary, given that the sentences above are judged as unacceptable, one does not expect them to be part of the linguistic data that we are exposed to.

On the basis of the judgements, try to classify the examples and formulate some principles that might account for the relative acceptability. You may find the discussion of examples (3), (4), (5) and (6) in the text of some help. In chapter 7 and following we shall discuss the sentences above and similar ones. We shall assume that they are ungrammatical and we shall attempt to formulate the rules and principles at work.

Exercise 2

If you are a native speaker of a language other than English translate the sentences in exercise 1 in your own language, keeping as close to the English models as you can, and rank them for acceptability. Try to formulate some principles to explain the degree of acceptability.

If you have access to judgements on the English data and on data

in other languages, try to check if the same degree of acceptability of the examples could be explained by the same principle(s).

Exercise 3

When reading section 3 the reader will have noted that there are certain uses of English which allow the omission of the subject and in which text example (8b) would have been grammatical. The following are attested examples.

- 1a A very sensible day yesterday. *Saw* no one. *Took* the bus to Southwark Bridge. *Walked* along Thames Street; *saw* a flight of steps down to the river. . . . *Found* the strand of the Thames, under the warehouses. . . . *Thought* of the refugees from Barcelona walking 40 miles, one with a baby in a parcel. . . . *Made* a circuit: *discovered* St Olave's Hart Street.
(Woolf, 1985: 203–4)
- 1b The poor little boy *wont* say whats the matter. He takes no interest in anything. *Wont* turn and wave to her . . . *drudges* on at Latin.
(Woolf, 1985: 117)
- 1c Brilliant *could* have stayed all day.
Brill – *must* come again.
Could see everything from wheelchair.
(Quotes from Visitors book 1991, The Green, Beaumaris, Anglesey North Wales)

Even a superficial glance at these examples shows us that all of the italicized verbs have one property in common: the subject is missing. In (1a) and (1c) the first person subject is omitted, in (1b) it is the third person. The omission of the subject in certain types of English is observed in traditional descriptions (Quirk, et al. 1985: 896–7). Such examples are relatively easy to come by in certain registers of English, which we could roughly characterize as belonging to abbreviated writing. We do not have to look for attested examples of usage to discuss such data; every native speaker of English will be able to think of relevant examples and even non-native speakers will quickly pick up this type of ellipsis in the appropriate register.

All the attested examples are instances where the subject of a root clause is omitted. By root clause we mean a clause which is not subordinate to another clause. The following variants on sentences drawn from Virginia Woolf's diary are unacceptable:

- 2a I must work, as *(I) told Sally G . . . (Woolf, 1985: 38)
- 2b I don't think *(I) need lie quaking at night . . . (Woolf, 1985: 38)
- 2c I find this morning that *(I) interrupted the crisis of that London Group meeting . . . (Woolf, 1985: 9)

Another property that we find is that attested examples never occur in questions. In the examples in (3) drawn from usage data, the subject pronoun cannot be omitted.

- 3a And what could *(we) do . . . (Woolf, 1985: 19)
- 3b What can *(I) say . . . (Woolf, 1985: 3)
- 3c Now who is *(she) . . . (Woolf, 1985: 15)
- 3d What shall *(I) write . . . (Woolf, 1985: 40)

The absence of such examples in subordinate questions is expected if the omission of the subject is a root phenomenon.

- 4 and this will show how hard *(I) work . . . (Woolf, 1985: 13)

The subject also never is and in fact cannot be omitted when it is preceded by a non-subject:

- 5a The next book *(I) think of calling Answers to Correspondents . . . (Woolf, 1985: 3)
- 5b Such twilight gossip *(it) seemed . . . (Woolf, 1985: 8)
- 5c This story *(I) repeated to Duncan last night . . . (Woolf, 1985: 9)
- 5d And there *(I) was in the rush of an end . . . (Woolf, 1985: 11)

When a negative constituent is preposed, resulting in a word-order where the auxiliary precedes the subject, the subject pronoun cannot be omitted.

- 6a Seldom have *(I) been more completely miserable than I was about 6.30 last night . . . (Woolf, 1985: 8)
- 6b Never have *(I) worked so hard at any book . . . (Woolf, 1985: 16)
- 6c Nor do *(I) wish even to write about it here . . . (Woolf, 1985: 44)

Finally observe that only subject pronouns are omitted: objects are not omitted. There is not a single example in Woolf's diary of the omission of an object and the omission of *me* in (7) leads to an unacceptable sentence.

- 7 This led *(me) to imagine any number of catastrophes . . . (Woolf, 1985: 9)

At first sight one might be tempted to conclude that this variety of English exhibits a manifestation of the *pro*-drop phenomenon discussed in section 4. Evaluate this proposal. You should draw on the English data given above, on the Italian data given in (8) and (9), and on your own intuitions. Using the argumentation introduced in section 4 try to state your argument as systematically and as explicitly as possible.

- 8 Credo che sia già partito.
I believe that he (subj) already left
'I think that he has already left.'

- 9a Dove è?
where is (3sg)
'Where is he?'

- 9b Che vuoi?
what want (2sg)
'What do you want?'

- 9c Questo libro non lo voglio.
this book *non* it want (1sg)
'This book, I don't want it.'

Readers whose first language is another non *pro*-drop language are encouraged to consider the question of the omissibility of the subject in abbreviated registers (diaries, informal notes) in their native language.

For a discussion of the omission of the subject in English the reader is referred to Haegeman (1990) and to Rizzi (1992a). The latter paper relates the phenomenon of omission of the subject in the diary register to data drawn from acquisition.