AN INTRODUCTION TO

ENGLISH SENTENCE STRUCTURE

ANDREW RADFORD

CAMBRIDGE www.cambridge.org/9780521516938

This page intentionally left blank

An Introduction to English Sentence Structure

This outstanding resource for students offers a step-by-step, practical introduction to English syntax and syntactic principles, as developed by Chomsky over the past 15 years. Assuming little or no prior background in syntax, Andrew Radford outlines the core concepts and how they can be used to describe various aspects of English sentence structure. This is an abridged version of Radford's major new textbook *Analysing English Sentences* (also published by Cambridge University Press), and will be welcomed as a handy introduction to current syntactic theory.

ANDREW RADFORD is Professor & Head of the Department of Language and Linguistics at the University of Essex. His recent publications include *Minimalist Syntax: Exploring the Structure of English* (Cambridge, 2004) and *English Syntax: An Introduction* (Cambridge, 2004).

An Introduction to English Sentence Structure

ANDREW RADFORD

University of Essex

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK
Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521516938

First published in print format 2009

ISBN-13 978-0-511-50090-9 eBook (Adobe Reader)

ISBN-13 978-0-521-51693-8 hardback

ISBN-13 978-0-521-73190-4 paperback

Contents

Preface

Aims

This book supercedes my *English Syntax* book, published in 2004. Although there is much in common between the two books, it should be noted that this book contains new material and new analyses (particularly in later chapters). It has two main aims. The first is to provide an intensive introduction to recent work in syntactic theory (more particularly to how the *syntactic component* operates within the model of grammar assumed in recent work within the framework of Chomsky's *Minimalist Program*). The second is to provide a description of a range of phenomena in English syntax, making use of Minimalist concepts and assumptions wherever possible.

Key features

The book is intended to be suitable both for people with only minimal grammatical knowledge, and for those who have already done quite a bit of syntax but want to know something (more) about Minimalism. It is not historicist or comparative in orientation, and does not presuppose knowledge of earlier or alternative models of grammar. It is written in an approachable style, avoiding unnecessary complexity and unexplained jargon. Each chapter contains:

- a core text (divided up into eight sections or so) focusing on a specific topic
- a summary recapitulating the main points in the chapter
- a list of key concepts/principles introduced in the chapter
- a bibliographical section providing extensive references to original source material
- a workbook section containing two different kinds of exercise
- a set of *model answers* accompanying the exercises, together with extensive *helpful hints* designed to eliminate common errors students make and to help students whose native language is not English
- an extensive glossary and integral list of abbreviations

The bibliographical background section often contains references to primary research works which are highly technical in nature, and so it would not be

appropriate for students to tackle them until they have read the whole book: they are intended to provide a useful source of bibliographical information for extended essays or research projects in particular areas, rather than being essential back-up reading: indeed, the exercises in the book are designed in such a way that they can be tackled on the basis of the coursebook material alone. The *glossary* at the end of the book provides simple illustrations of how key technical terms are used (both theory-specific terms like **EPP** and traditional terms like **subject**): technical terms are written in **bold** print when they are mentioned for the first time in the main text (*italics* being used for highlighting particular expressions – e.g. a key word appearing in an example sentence). The glossary also contains an integrated list of *abbreviations*.

The book is intensive and progressive in nature, which means that it starts at an elementary level but gets progressively harder as you delve further into the book. A group of students I taught an earlier version of the book to gave the following degree-of-difficulty score to each chapter on a 5-point scale ranging from 1 = *very easy* to 5 = *very hard*: ch.1 = 1.7; ch.2 = 2.2; ch.3 = 2.7; ch.4 = 2.9; ch.5 = 3.2; ch.6 = 3.4; ch.7 = 3.7; ch.8 = 4.2; ch.9 = 4.4.

Successive chapters become cumulatively more complex, in that each chapter presupposes material covered in previous chapters as well as introducing new material: hence it is helpful to go back and read material from earlier chapters every so often. In some cases, analyses presented in earlier chapters are subsequently refined or revised in the light of new assumptions made in later chapters.

Teaching materials

For teachers adopting the book, I have developed a series of web materials (in the form of Powerpoint transparencies) designed to provide two hours' worth of teaching material for each chapter. The relevant materials present detailed step-by-step analyses of those exercise examples which have the symbol **(w)** after them in the coursebook. They can be accessed at www.cambridge.org/radford

Companion volume

This book is being produced in parallel with a longer version entitled *Analysing English Sentences: A Minimalist Approach.* In this shorter version, the main text (particularly in the later chapters) is generally about a third shorter than the main text in the longer version (with the exception of chapters 1 and 6). This shorter version is aimed primarily at students whose native language is not English, and who are taking (English) syntax as a minor rather than a major course. The two books have an essentially parallel organisation into chapters and sections (though additional sections, technical discussion and bibliographial references are included in the longer version), and contain much the same exercise material.

In keeping the two books parallel in structure and organisation as far as possible, I am mindful of the comment made in a review of two earlier books which I produced in parallel longer and shorter versions (Radford 1997a and Radford 1997b) that some readers may wish to read the short version of a given chapter first, and then look at the longer version afterwards, and that this is 'not facilitated' if there is 'an annoyingly large number of non-correspondences' between the two (Ten Hacken 2001, p. 2). Accordingly, I have tried to maximise correspondence between the 'long' and 'short' versions of these two new books.

Acknowledgments

I am grateful to Neil Smith (of University College London) for his forebearance in patiently wading through an earlier draft of the manuscript and pointing out some of the imperfections in it, while managing to make his comments challenging and good-humoured at the same time. Thanks also go to my Essex colleague Bob Borsley for helpful comments, and to Michèle Vincent for preparing the index.

Dedication

This book is dedicated to my long-suffering wife Khadija (who has had to put up with extended periods of authorial autism) and to her family, who have always spoiled me shamefully (and done their best to indulge my every whim) whenever we visit Morocco.

1 Grammar

1.1 Overview

In broad terms, this book is concerned with aspects of grammar. Grammar is traditionally subdivided into two different but interrelated areas of study – **morphology** and **syntax**. Morphology is the study of how words are formed out of smaller units (called **morphemes**), and so addresses questions such as 'What are the component morphemes of a word like *antidisestablishmentarianism*, and what is the nature of the morphological operations by which they are combined together to form the overall word?' Syntax is the study of the way in which phrases and sentences are structured out of words, and so addresses questions like 'What is the structure of a sentence like *What's the president doing?* and what is the nature of the grammatical operations by which its component words are combined together to form the overall sentence structure?' In this chapter, we begin (in §1.2) by taking a brief look at the approach to the study of syntax taken in **traditional grammar**: this also provides an opportunity to introduce some useful grammatical terminology. In the remainder of the chapter, we look at the approach to syntax adopted within the theory of **Universal Grammar** developed by Chomsky.

1.2 Traditional grammar: Categories and functions

Within traditional grammar, the syntax of a language is described in terms of a taxonomy (i.e. classificatory list) of the range of different types of syntactic structures found in the language. The central assumption underpinning syntactic analysis in traditional grammar is that phrases and sentences are built up of a series of **constituents** (i.e. syntactic units), each of which belongs to a specific **grammatical category** and serves a specific **grammatical function**. Given this assumption, the task of the linguist in analysing the syntactic structure of any given type of sentence is to identify each of the constituents in the sentence, and (for each constituent) to say what category it belongs to and what function it serves. For example, in relation to the syntax of a simple sentence like:

(1) Students protested

it would traditionally be said that the sentence consists of two constituents (the word *students* and the word *protested*), that each of these constituents belongs to a specific grammatical category (*students* being a plural **noun** and *protested* a past tense **verb**) and that each serves a specific grammatical function (*students* being the **subject** of the sentence, and *protested* being the **predicate**). The overall sentence *Students protested* has the categorial status of a **clause** which is **finite** in nature (by virtue of denoting an event taking place at a specific time), and has the semantic function of expressing a **proposition** which is **declarative** in **force** (in that it is used to make a statement rather than e.g. ask a question). Accordingly, a traditional grammar of English would tell us that the simplest type of finite declarative clause found in English is a sentence like (1), in which a nominal subject is followed by a verbal predicate. Let's briefly look at some of the terminology used here.

In traditional grammar, words are assigned to grammatical categories (called **parts of speech**) on the basis of their **semantic** properties (i.e. meaning), **morphological** properties (i.e. the range of different forms they have) and **syntactic** properties (i.e. word-order properties relating to the positions they can occupy within sentences): a set of words which belong to the same category thus have a number of semantic, morphological and syntactic properties in common. There are traditionally said to be two different types of word, namely **content words/contentives** (= words which have substantive lexical content) on the one hand, and **function words/functors** (= words which essentially serve to mark grammatical properties) on the other. The differences between the two can be illustrated by comparing a contentive like *car* with a functor like *they*. A noun like *car* has substantive lexical content in that it denotes an object which typically has four wheels and an engine, and it would be easy enough to draw a picture of a typical *car*; by contrast, a pronoun such as *they* has no descriptive content (e.g. you can't draw a picture of *they*), but rather is a functor which simply marks grammatical (more specifically, person, number and case) properties in that it is a third person plural nominative pronoun. Because they have lexical semantic content, content words often (though not always) have antonyms (i.e. 'opposites') – e.g. the adjective *tall* has the antonym *short*, the verb *increase* has the antonym *decrease*, and the preposition *inside* has the antonym *outside*: by contrast, a typical function word like e.g. the pronoun *me* has no obvious antonym. Corresponding to these two different types of (content and function) word are two different kinds of grammatical category – namely **lexical/substantive categories** (= categories whose members are content words) on the one hand, and **functional categories** (= categories whose members are function words) on the other.

Let's begin by looking at the main **lexical/substantive categories** found in English – namely **noun, verb, adjective, adverb** and **preposition** (conventionally abbreviated to N, V, A, ADV and P in order to save space). **Nouns** (= N) are traditionally said to have the semantic property that they denote entities: so, *bottle* is a noun (since it denotes a type of object used to contain liquids),

water is a noun (since it denotes a type of liquid) and *John* is a noun (since it denotes a specific person). There are a number of distinct subtypes of noun: for example, a noun like *chair* is a **count noun** in that it can be counted (cf. *one chair, two chairs* . . .), whereas a noun like *furniture* is a **mass noun** in that it denotes an uncountable mass (hence the ungrammaticality of **one furniture,* **two furnitures* – a prefixed star/asterisk being used to indicate that an expression is ungrammatical). Likewise, a distinction is traditionally drawn between a **common noun** like *boy* (which can be modified by a determiner like *the* – as in *The boy is lying*) and a **proper noun** like *Andrew* (which cannot be used in the same way in English, as we see from the ungrammaticality of **The Andrew is lying*). Count nouns generally have the morphological property that they have two different forms: a **singular** form (like *horse* in *one horse*) used to denote a single entity, and a **plural** form (like *horses* in *two horses*) used to denote more than one entity. Common nouns have the syntactic property that only (an appropriate kind of) noun can be used to end a sentence such as *They have no* . . . In place of the dots here we could insert a singular count noun like *car,* or a plural count noun like *friends* or a mass noun like *money,* but not other types of word (e.g. not *see* or *slowly* or *up*, as these are not nouns).

A second lexical/substantive category is that of **verb** (= V). These are traditionally said to have the semantic property that they denote actions or events: so, *eat*, *sing*, *pull* and *resign* are all (action-denoting) verbs. From a syntactic point of view, verbs have the property that only an appropriate kind of verb (in its uninflected **infinitive** form) can be used to complete a sentence such as *They/It can* . . . So, words like *stay*, *leave*, *hide*, *die*, *starve* and *cry* are all verbs and hence can be used in place of the dots here (but words like *apple*, *under*, *pink* and *if* aren't). From a morphological point of view, regular verbs like *cry* in English have the property that they have four distinct forms: e.g. alongside the **bare** (i.e. uninflected) **form** *cry* we find the **present tense** form *cries*, the **past tense/perfect participle/passive participle** form *cried* and the **progressive participle** form *crying*. (See the Glossary of terminology at the end of this book if you are not familiar with these terms.)

A third lexical/substantive category is that of **adjective** (= A). These are traditionally said to have the semantic property of denoting states or attributes (cf. *ill*, *happy*, *tired*, *conscientious*, *red*, *cruel, old* etc.). They have the syntactic property that they can occur after *be* to complete a sentence like *They may be* . . . (as with *They may be tired/ill/happy* etc.), and the further syntactic property that (if they denote a **gradable** property which can exist in varying degrees) they can be modified by a degree word like *very/rather/somewhat* (cf. *She is very happy*). Many (but not all) adjectives have the morphological property that they have **comparative** forms ending in *-er* and **superlative** forms ending in *-est* (cf. *big/bigger/biggest*).

A fourth lexical/substantive category is that of **adverb** (= ADV). These often have the semantic property that they denote the manner in which an action is performed (as with *well* in *She sings well*). Regular adverbs have the morphological

property that they are formed from adjectives by the addition of the suffix *-ly* (so that corresponding to the adjective *sad* we have the adverb *sadly*). A syntactic property of adverbs is that an adverb (like e.g. *badly*) is the only kind of word which could be used to end sentences such as *She behaved . . .*, *He treats her . . .* or *He worded the statement . . .*

The fifth and final lexical/substantive category found in English is that of **preposition** (= P). Many of these have the semantic property of marking location (cf. *in/on/off/inside/outside/under/above/below*). They have the syntactic property that a preposition (with the appropriate kind of meaning) can be modified by *right* in the sense of 'completely', or by *straight* in the sense of 'directly' (as with the preposition *down* in *He fell right down the stairs* and the preposition *to* in *He went straight to bed*). Prepositions have the morphological property that they are invariable/uninflected forms (e.g. the preposition *off* has no past tense form **offed*, no superlative form **offest* and so on).

In addition to the five lexical/substantive categories identified above, English also has a number of functional categories. One such functional category is that of **determiner** (= D) – a category whose members are traditionally said to include the definite article *the* and the demonstrative determiners *this/that/these/those*. They are called determiners because they have the semantic property that they determine specific semantic properties of the noun expression that they introduce, marking it as a definite referring expression: for example, an expression like *the car* in a sentence such as *Shall we take the car?* is a definite referring expression in the sense that it refers to a definite (specific) car which is assumed to be familiar to the hearer/addressee. A related class of words are those which belong to the functional category **quantifier** (= Q), denoting expressions of quantity, such as *some/all/no/any/each/every/most/much/many*. (We shall also take the indefinite article *a* to be a quantifier – one which quantifies over a single entity.)

A further type of functional category found in English is that of **pronoun** (= PRN). Pronouns are items which are said to 'stand in place of' (the meaning of the prefix *pro-*) or 'refer back to' noun expressions. However, there are reasons to think that there are a number of different types of pronoun found in English and other languages. For example, in sentences such as *John has a red car and Jim has a blue one*, the word *one* is traditionally said to be a pronoun because it has no lexical semantic content of its own, but rather takes its content from its **antecedent** (i.e. *one* refers back to the noun *car* and so *one* is interpreted as having the same meaning as *car*). However, from a morphological perspective, the pronoun *one* behaves like a regular count noun in that it has a plural form ending in *-s* (as in *I'll take the green apples if you haven't got any red ones*). So, more accurately, we could say that *one* is an N-pronoun (or pronominal noun). By contrast, in a sentence like *Many miners were rescued, but some died*, the word *some* seems to function as a Q-pronoun (i.e. a pronominal quantifier). And in a sentence like *These apples are ripe, but those aren't*, the word *those* seems to be a D-pronoun (i.e. a pronominal determiner). Indeed, some linguists have argued that so-called **personal pronouns** like *I/me/we/us/you/he/him/she/her/it/they/them* are also

D-pronouns: the rationale for this is that some such pronouns can be used as determiners which modify a following noun (as in *We republicans don't trust you democrats*, where *we* could be argued to be a determiner modifying the noun *republicans*, and *you* could be seen as a determiner modifying the noun *democrats*). While, as noted here, pronouns can be argued to belong to a number of distinct types of category, in order to simplify discussion I shall simply refer to them as belonging to the category PRN throughout this book. (Because there are a number of different types of pronoun, some linguists prefer to refer to them by using the more general term **proform**.)

Another type of functional category found in English is that of **auxiliary (verb)**. They have the semantic property of marking grammatical properties such as **tense**, **aspect**, **voice** or **mood** (see the Glossary of terminology at the end of the book if you are not sure what these terms mean). Auxiliaries have the syntactic property that (unlike lexical/main verbs) they can be inverted with their subject in questions (so that corresponding to a statement like *It is raining* we have the question *Is it raining?* where the auxiliary *is* has moved in front of the subject *it* and is said to have been *inverted*). The items italicised in (2) below (in the use illustrated there) are traditionally categorised as auxiliaries taking a [bracketed] complement containing a bold-printed verb:

(2) (a) He *has*/*had* [**gone**]
(b) She *is*/*was* [**staying** at home]
(c) They *are*/*were* [**taken** away for questioning]
(d) He really *does*/*did* [**say** a lot]
(e) You *can*/*could* [**help** us]
(f) They *may*/*might* [**come** back]
(g) He *will*/*would* [**get** upset]
(h) I *shall*/*should* [**return**]

In the uses illustrated here, *have/be* in (2a,b) are (**perfect/progressive**) **aspect** auxiliaries, *be* in (2c) is a (**passive**) **voice** auxiliary, *do* in (2d) is an **expletive** or **dummy** auxiliary (i.e. one with no intrinsic lexical semantic content), and *can*/*could*/*may*/*might*/*will*/*would*/*shall*/*should* in (2e–h) are **modal** auxiliaries. What auxiliaries in sentences like those above have in common is the fact that they inflect for present/past tense. Hence, in work in syntax over the past ten years or so, they have been said to belong to the category T (= tense-marked auxiliary).

An interesting word which has been argued to be related to tense-marking auxiliaries in work over the past thirty years or so is the infinitive particle *to*, in sentences such as:

(3) They are now expecting the president *to* be impeached tomorrow

In a sentence like (3), infinitival *to* seems to have future time-reference (in that the act of impeachment will take place at some time in the future), and this is why we can use the word *tomorrow* in the *to*-clause. In this respect, infinitival *to* seems

to have much the same function as the auxiliary *will* in *They are now expecting that the president will be impeached tomorrow*, suggesting that infinitival *to* is an **infinitival tense marker**, and so belongs to the same category T as present/past tense auxiliaries such as *is/was*. The difference between auxiliaries and infinitival *to* is that most auxiliaries overtly inflect for present/past tense (though this is not true of the invariable auxiliaries *must* and *ought*), whereas infinitival *to* is invariable in form. We can thus say that an auxiliary like *will* is a finite T constituent, whereas infinitival *to* is a nonfinite T.

The last type of functional category which we will look at is a kind of word (like each of the words italicised in the examples below) which is traditionally termed a (subordinating) **conjunction:**

(4) (a) I think [*that* you may be right]
(b) I doubt [*if* you can help me]
(c) I'm anxious [*for* you to receive the best treatment possible]

Each of the bracketed clauses in (4) is a complement clause, in that it is the complement of the word immediately preceding it (*think*/*doubt*/*anxious*); for this reason, the italicised word which introduces each clause is known in work since the 1960s as a **complementiser** (= C), and this is the terminology which will be adopted throughout this book. Complementisers are functors in the sense that they encode particular sets of grammatical properties. For example, complementisers encode (non)finiteness by virtue of the fact that they are intrinsically finite or nonfinite. More specifically, the complementisers *that* and *if* are inherently finite in the sense that they can only be used to introduce a finite clause (i.e. a clause containing a present or past tense auxiliary or verb, like the present tense auxiliaries *may* and *can* in 4a and 4b); by contrast, *for* is an inherently infinitival complementiser, and so can be used to introduce a clause containing infinitival *to* (as in 4c). Moreover, *that* introduces a **declarative** clause (i.e. one which has the **force** of a statement), *if* introduces an **interrogative** clause (i.e. one which has the force of a question) and *for* introduces an **irrealis** clause (i.e. one relating to a hypothetical event which hasn't yet taken place and may or may not take place at some stage in the future). Hence, we can say *that* is a finite declarative complementiser, *if* is a finite interrogative complementiser and *for* is an infinitival irrealis complementiser.

Using the set of syntactic categories outlined above, we can employ the traditional **labelled bracketing** technique to *categorise* words (i.e. assign them to grammatical categories) in a way which describes how they are being used in a particular sentence. Using this technique, the words in sentence (5a) below can be categorised as in (5b):

(5) (a) The president is clearly feeling angry that Congress has refused to negotiate with him
(b) [$_{D}$ The] [$_{N}$ president] [$_{T}$ is] [$_{ADV}$ clearly] [$_{V}$ feeling] [$_{A}$ angry] [$_{C}$ that] [$_{N}$ Congress] [$_{T}$ has] [$_{V}$ refused] [$_{T}$ to] [$_{V}$ negotiate] [$_{P}$ with] [$_{PRN}$ him]

The labelled bracketing in (5b) tells us that *the* is a D/determiner, *president* a N/noun, *is* a T/present tense auxiliary, *clearly* an ADV/adverb, *feeling* a V/verb, *angry* an A/adjective, *that* a C/complementiser, *Congress* a N/noun, *has* a T/present tense auxiliary, *refused* a V/verb, *to* a T/infinitival tense particle, *negotiate* a V/verb, *with* a P/preposition and *him* a PRN/pronoun.

The discussion of grammatical categories presented above is merely a brief sketch: however, it suffices to illustrate the point that when traditional grammarians analyse the syntax of sentences, they begin by assigning each of the words in the sentence to a grammatical category which describes how it is being used in the sentence concerned. Grammatical differences between individual words belonging to the same category are traditionally described in terms of sets of **grammatical features**, and these features (by convention) are enclosed in square brackets. For example, both *she* and *us* are pronouns, but they differ in that *she* is a **third person** pronoun which is **feminine** in **gender**, **singular** in **number** and **nominative** in **case**, whereas *us* is a first person pronoun which is **plural** in number and **accusative** in case. Accordingly, we can describe the differences between these two pronouns by saying that the pronoun *she* carries the features [third-person, singular-number, feminine-gender, nominative-case], whereas *us* carries the features [first-person, plural-number, accusative-case].

As noted at the beginning of this section, traditional grammarians are also concerned to describe the **grammatical functions** which words and other expressions fulfil within the sentences containing them. We can illustrate this point in terms of the following set of sentences:

(6) (a) *John* smokes
(b) *The president* smokes
(c) *The president of Utopia* smokes
(d) *The former president of the island paradise of Utopia* smokes

Sentence (6a) comprises the noun *John* which serves the function of being the **subject** of the sentence (and denotes the person performing the act of smoking), and the verb *smokes* which serves the function of being the **predicate** of the sentence (and describes the act being performed). In (6a), the subject is the single noun *John*; but as the examples in (6b,c,d) show, the subject of a sentence can also be an (italicised) phrase like *the president*, or *the president of Utopia* or *the former president of the island paradise of Utopia*.

Now consider the following set of sentences:

(7) (a) John smokes *cigars*
(b) John smokes *Cuban cigars*
(c) John smokes *Cuban cigars imported from Havana*
(d) John smokes *a specific brand of Cuban cigars imported by a friend of his from Havana*

Sentence (7a) comprises the **subject** *John*, the **predicate** *smokes* and the **complement** (or **direct object**) *cigars*. (The complement *cigars* describes the entity on

which the act of smoking is being performed; as this example illustrates, subjects normally precede the verb with which they are associated in English, whereas complements typically follow the verb.) The complement in (7a) is the single noun *cigars*; but a complement can also be a **phrase**: in (7b), the complement of *smokes* is the phrase *Cuban cigars*; in (7c) the complement is the phrase *Cuban cigars imported from Havana*; and in (7d) the complement is the phrase *a specific brand of Cuban cigars imported by a friend of his from Havana*. A verb which has a noun or pronoun expression as its direct object complement is traditionally said to be **transitive**.

From a semantic perspective, subjects and complements share in common the fact that they generally represent entities directly involved in the particular action or event described by the predicate: to use the relevant semantic terminology, we can say that subjects and complements are **arguments** of the predicate with which they are associated. Predicates may have one or more arguments, as we see from sentences such as (8) below, where each of the bracketed nouns is a different argument of the italicised predicate:

(8) (a) [John] *resigned*
(b) [John] *felt* [remorse]
(c) [John] *sent* [Mary] [flowers]

A predicate like *resign* in (8a) which has a single argument is said to function as a **one-place predicate** (in the relevant use); one like *feel* in (8b) which has two arguments is a **two-place predicate**; and one like *send* in (8c) which has three arguments is a **three-place predicate**.

In addition to predicates and arguments, sentences can also contain **adjuncts**, as we can illustrate in relation to (9) below:

(9) (a) The president smokes a cigar *after dinner*
(b) The president smokes a cigar *in his office*

In both sentences in (9), *smokes* functions as a two-place predicate whose two arguments are its subject *the president* and its complement *a cigar*. But what is the function of the phrase *after dinner* which also occurs in (9a)? Since *after dinner* isn't one of the entities directly involved in the act of smoking (i.e. it isn't consuming or being consumed), it isn't an argument of the predicate *smoke*. On the contrary, *after dinner* simply serves to provide additional information about the time when the smoking activity takes place. In much the same way, the italicised expression *in his office* in (9b) provides additional information about the location of the smoking activity. An expression which serves to provide (optional) additional information about the time or place (or manner, or purpose etc.) of an activity or event is said to serve as an **adjunct**. So, *after dinner* and *in his office* in (9a,b) are both **adjuncts**.

So far, all the sentences we have looked at in (6–9) have been **simple sentences** which contain a single **clause**. However, alongside these we also find **complex sentences** which contain more than one clause, like (10) below:

(10) Mary knows John smokes

If we take the traditional definition of a clause as a predication structure (more precisely, a structure containing a predicate which has a subject, and which may or may not also contain one or more complements and adjuncts), it follows that since there are two predicates (*knows* and *smokes*) in (10), there are correspondingly two clauses – the *smokes* clause on the one hand, and the *knows* clause on the other. The *smokes* clause comprises the subject *John* and the predicate *smokes*; the *knows* clause comprises the subject *Mary*, the predicate *knows* and the complement *John smokes*. So, the complement of *knows* here is itself a clause – namely the clause *John smokes*. More precisely, the *smokes* clause is a **complement clause** (because it serves as the complement of *knows*), while the *knows* clause is the **main clause** (or **principal clause** or **independent clause** or **root clause**). The overall sentence (10) *Mary knows John smokes* is a **complex sentence** because it contains more than one clause. In much the same way, (11) below is also a complex sentence:

(11) The press clearly think the president deliberately lied to Congress

Once again, it comprises two clauses – one containing the predicate *think*, the other containing the predicate *lie*. The main clause comprises the subject *the press*, the adjunct *clearly*, the predicate *think* and the complement clause *the president deliberately lied to Congress*. The complement clause in turn comprises the subject *the president*, the adjunct *deliberately*, the predicate *lie* and the complement *to Congress*.

As was implicit in our earlier classification of (1) as a **finite** clause, traditional grammars draw a distinction between **finite** and **nonfinite** clauses. In this connection, consider the contrast between the italicised clauses below (all of which function as the complement of an underlined adjective or verb):

(12) (a) She was glad *that he apologised*
(b) She demanded *that he apologise*
(c) I can't imagine *him apologising*
(d) It would be sensible *for him to apologise*
(e) It's important to know *when to apologise*

The italicised clauses in (12a,b) are finite, and it is characteristic of finite clauses in English that they contain an (auxiliary or main) verb marked for tense/mood, and can have a nominative pronoun like *he* as their subject. In (12a), the verb *apologised* is finite by virtue of being inflected for past tense and **indicative mood**, and by virtue of having a nominative subject (*he*); in (12b), the verb *apologise* is finite by virtue of being inflected for **subjunctive mood** (and perhaps present tense, though this is far from clear), and by virtue of having a nominative subject (*he*). A clause containing a verb in the indicative mood denotes a real (or **realis**, to use the relevant grammatical term) event or state occurring at

a specific point in time; a subjunctive clause by contrast denotes a hypothetical or unreal (= **irrealis**) event or state which has not yet occurred and which may never occur. In contrast to the italicised clauses in (12a,b), the clauses italicised in (12c–e) are nonfinite, in that they contain no verb marked for tense or mood, and do not allow a nominative subject. For example, the verb *apologising* in (12c) is nonfinite because it is a tenseless and moodless **gerund** form, and has an **accusative** subject *him*. Likewise, the verb *apologise* in (12d,e) is a tenseless and moodless **infinitive** form (as we see from the fact that it follows the infinitive particle *to*), and has an accusative subject *him* in (12d), and a 'silent' (implicit) subject in (12e). (Excluded from our discussion here are gerund structures with genitive subjects like the italicised in 'I can't stand *his perpetual(ly) whining about syntax*', since these are more nominal than clausal in nature.)

As the examples in (12) illustrate, whether or not a clause is finite in turn determines the kind of subject it can have, in that finite clauses can have a **nominative** pronoun like *he* as their subject, but nonfinite clauses cannot. Accordingly, one way of telling whether a particular clause is finite or not is to see whether it can have a nominative pronoun (like *I/we/he/she/they*) as its subject. In this connection, consider whether the italicised clauses in the dialogues in (13a,b) below are finite or nonfinite:

(13) (a) SPEAKER A: I know you cheat on me
SPEAKER B: OK, I admit it. *I cheat on you*. But not with any of your friends
(b) SPEAKER A: I know you cheat on me
SPEAKER B: *Me cheat on you?* No way! I never would!

The fact that the italicised clause in speaker B's reply in (13a) has the nominative subject *I* suggests that it is finite, and hence that the verb *cheat* (as used in the italicised sentence in 13a) is a first person singular present tense form. By contrast, the fact that the italicised clause in speaker B's reply (13b) has the accusative subject *me* suggests that it is nonfinite, and that the verb *cheat* (as used in the italicised sentence in 13b) is an infinitive form (and indeed this is clear from sentences like *Me be a cheat? No way!* where we find the infinitive form *be*).

In addition to being finite or nonfinite, each clause within a sentence has a specific **force**. In this connection, consider the following simple (single-clause) sentences:

(14) (a) He went home
(b) Are you feeling OK?
(c) You be quiet!
(d) What a great idea that is!

A sentence like (14a) is traditionally said to be **declarative** in force, in that it is used to make a statement. (14b) is **interrogative** in force in that it is used to ask a question. (14c) is **imperative** in force, by virtue of being used to issue an order

or command. (14d) is **exclamative** in force, in that it is used to exclaim surprise or delight. In complex sentences, each clause has its own force, as we can see in relation to (15) below:

(15) (a) He asked where she had gone
(b) Did you know that he has retired?
(c) Tell her what a great time we had!

In (15a), the main (*asked*) clause is declarative, whereas the complement (*gone*) clause is interrogative; in (15b) the main (*know*) clause is interrogative, whereas the complement (*retired*) clause is declarative; and in (15c), the main (*tell*) clause is imperative, whereas the complement (*had*) clause is exclamative.

We can summarise this section as follows. From the perspective of traditional grammar, the syntax of a language is described in terms of a **taxonomy** (i.e. a classificatory list) of the range of different phrase-, clause- and sentence-types found in the language. So, for example, a typical traditional grammar of (say) English will include chapters on the syntax of negatives, interrogatives, exclamatives, imperatives and so on. The chapter on interrogatives will note (e.g.) that in main-clause questions in English like 'Is he winning?' the present tense **auxiliary** *is* **inverts** with (i.e. moves in front of) the subject *he*, but not in complement clause questions like the *if*-clause in 'I wonder if he *is* winning', and will typically not be concerned with trying to explain *why* **auxiliary inversion** applies in main clauses but not complement clauses: this reflects the fact that the primary goal of traditional grammar is *description* rather than *explanation*.

1.3 Universal Grammar

In contrast to the **taxonomic** approach adopted in traditional grammar, Chomsky takes a **cognitive** approach to the study of grammar. For Chomsky, the goal of the linguist is to determine what it is that native speakers *know* about their native language which enables them to speak and understand the language, and how this linguistic knowledge might be represented in the mind/brain: hence, in studying language, we are studying a specific kind of cognition (i.e. human knowledge). In a fairly obvious sense, any native speaker of a language can be said to *know* the grammar of his or her native language. For example, any native speaker of English can tell you that the negative counterpart of *I like syntax* is *I don't like syntax*, and not e.g. **I no like syntax*: in other words, native speakers know how to combine words together to form expressions (e.g. negative sentences) in their language. Likewise, any native speaker of English can tell you that a sentence like *She loves me more than you* is ambiguous and has two **interpretations** which can be paraphrased as 'She loves me more than she loves you' and 'She loves me more than you love me': in other words, native speakers also know how to **interpret** (i.e. assign meaning to) expressions in their language.

However, it is important to emphasise that this grammatical knowledge of how to form and interpret expressions in your native language is **tacit** (i.e. subconscious) rather than **explicit** (i.e. conscious): so, it's no good asking a native speaker of English a question such as 'How do you form negative sentences in English?' since human beings have no conscious awareness of the processes involved in speaking and understanding their native language. To introduce a technical term devised by Chomsky, we can say that native speakers have grammatical **competence** in their native language: by this, we mean that they have tacit knowledge of the grammar of their language – i.e. of how to form and interpret words, phrases and sentences in the language.

In work in the 1960s, Chomsky drew a distinction between **competence** (the native speaker's tacit knowledge of his or her language) and **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 1965, p. 4). Very often, performance is an imperfect reflection of competence: we all make occasional slips of the tongue, or occasionally misinterpret something which someone else says to us. However, this doesn't mean that we don't know our native language or that we don't have *competence* in it. Misproductions and misinterpretations are **performance errors**, attributable to a variety of performance factors like tiredness, boredom, drunkenness, drugs, external distractions and so forth. A grammar of a language tells you what you need to know in order to have native-like competence in the language (i.e. to be able to speak the language like a fluent native speaker): hence, it is clear that grammar is concerned with competence rather than performance. This is not to deny the interest of performance as a field of study, but merely to assert that performance is more properly studied within the different – though related – discipline of psycholinguistics, which studies the psychological processes underlying speech production and comprehension.

Thus, when we study the grammatical competence of a native speaker of a language like English we're studying a cognitive system internalised within the brain/mind of native speakers of English which is the product of a 'cognitive organ' which is 'shared among human beings and in crucial respects unique to them' (Chomsky 2006, p. 1). In the terminology adopted by Chomsky (1986a, pp. 19–56), our ultimate goal in studying competence is to characterise the nature of the internalised linguistic system (or **I-language**, as Chomsky terms it) which makes native speakers proficient in English. Such an approach has obvious implications for the descriptive linguist who is concerned to develop a grammar of a particular language like English. According to Chomsky (1986a, p. 22) a grammar of a language is 'a theory of the I-language . . . under investigation'. This means that in devising a grammar of English, we are attempting to uncover the internalised linguistic system (= I-language) possessed by native speakers of English – i.e. we are attempting to characterise a mental state (a state of competence, and thus linguistic knowledge).

Chomsky's ultimate goal is to devise a theory of **Universal Grammar/UG** which generalises from the grammars of particular I-languages to the grammars of all possible natural (i.e. human) I-languages. He defines UG (1986a, p. 23) as 'the theory of human I-languages . . . that identifies the I-languages that are humanly accessible under normal conditions'. (The expression 'are humanly accessible' means 'can be acquired by human beings'.) In other words, UG is a theory about the nature of possible grammars of human languages: hence, a theory of Universal Grammar answers the question: 'What are the defining characteristics of the grammars of human I-languages?'

There are a number of **criteria of adequacy** which a theory of Universal Grammar must satisfy. One such criterion (which is implicit in the use of the term *Universal Grammar*) is **universality**, in the sense that a theory of UG must provide us with the tools needed to provide a **descriptively adequate** grammar for any and every human I-language (i.e. a grammar which correctly describes how to form and interpret expressions in the relevant language). After all, a theory of UG would be of little interest if it enabled us to describe the grammar of English and French, but not that of Swahili or Chinese.

However, since the ultimate goal of any theory is explanation, it is not enough for a theory of Universal Grammar simply to list sets of universal properties of natural language grammars; on the contrary, a theory of UG must seek to *explain* the relevant properties. So, a key question for any adequate theory of UG to answer is: 'Why do grammars of human I-languages have the properties they do?' The requirement that a theory should explain why grammars have the properties they do is conventionally referred to as the criterion of **explanatory adequacy**.

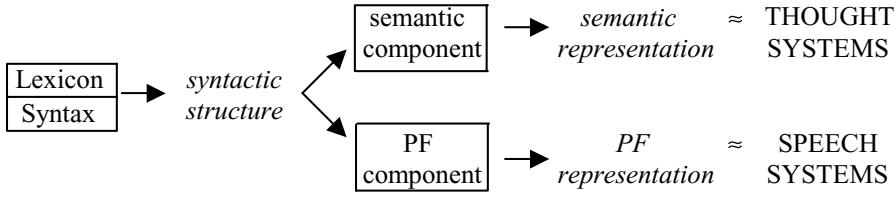
Since the theory of Universal Grammar is concerned with characterising the properties of natural (i.e. human) I-language grammars, an important question which we want our theory of UG to answer is: 'What are the defining characteristics of human I-languages which differentiate them from, for example, artificial languages like those used in mathematics and computing (e.g. Java, Prolog, C etc.), or from animal communication systems (e.g. the tail-wagging dance performed by bees to communicate the location of a food source to other bees)?' It therefore follows that the descriptive apparatus which our theory of Universal Grammar allows us to make use of in devising natural language grammars must not be so powerful that it can be used to describe not only natural languages, but also computer languages or animal communication systems (since any such excessively powerful theory wouldn't be able to pinpoint the criterial properties of natural languages which differentiate them from other types of communication system). In other words, a third condition which we have to impose on our theory of language is that it be maximally *constrained*: that is, we want our theory to provide us with technical devices which are so limited in their expressive power that they can only be used to describe natural languages, and are not appropriate for the description of other communication systems. A theory which is constrained in appropriate ways should enable us to provide a principled explanation for why

certain types of syntactic structure and syntactic operation simply aren't found in natural languages. One way of constraining grammars is to suppose that grammatical operations obey certain linguistic principles, and that any operation which violates the relevant principles leads to ungrammaticality: see the discussion in §1.5 below for a concrete example.

A related requirement is that linguistic theory should provide grammars which make use of the minimal theoretical apparatus required: in other words, grammars should be as simple as possible. Some earlier work in syntax involved the postulation of complex structures and principles: as a reaction to the excessive complexity of this kind of work, Chomsky in work over the past two decades has made the requirement to minimise the theoretical and descriptive apparatus used to describe language the cornerstone of the *Minimalist Program for Linguistic Theory* which he has been developing. He has suggested that language is a *perfect* system of *optimal design* in the sense that natural language grammars create structures which are designed to **interface** perfectly with other components of the mind – more specifically with speech and thought systems, so that (in the words of Chomsky 2005b, p. 2) 'Language is an optimal way to link sound and meaning.'

To make this discussion rather more concrete, let's look at the internal organisation of the grammar of a language. One component of a grammar is a **lexicon** (= dictionary = list of all the **lexical items/**words in the language and their linguistic properties), and in forming a given sentence out of a set of words, we first have to take the relevant words out of the lexicon. Our chosen words are then combined together by a series of syntactic computations in the **syntax** (i.e. in the **syntactic/computational component** of the grammar), thereby forming a **syntactic structure**. This syntactic structure serves as input into two other components of the grammar. One is the **semantic component** which **maps** (i.e. 'converts') the syntactic structure into a corresponding **semantic representation** (i.e. into a representation of linguistic aspects of its meaning): the other is a **PF component**, so called because it maps the syntactic structure into a **PF representation** (i.e. a representation of its **Phonetic Form**, giving us a phonetic **spellout** for each word, telling us how it is pronounced). The semantic representation interfaces with systems of thought, and the PF representation with systems of speech – as shown in diagrammatic form below:

(16)



Chomsky (2005b, p. 3) refers to the interface with thought systems as the 'conceptual-intentional interface (CI)', and to the interface with speech systems as the 'sensory-motor interface (SM)'. In terms of the model in (16), an

important consideration is that the (semantic and PF) representations which are 'handed over' to the (thought and speech) interface systems should contain only elements which are **legible** by the appropriate interface system – so that the semantic representations handed over to thought systems contain only elements contributing to meaning, and the PF representations handed over to speech systems contain only elements which contribute to phonetic form (i.e. to determining how the sentence is pronounced).

The neurophysiological mechanisms which underlie linguistic competence make it possible for young children to acquire language in a remarkably short period of time. Accordingly, a fourth condition which a linguistic theory must meet is that of **learnability**: it must provide grammars which are learnable by young children in a short period of time. The desire to maximise the **learnability** of natural language grammars provides an additional argument for minimising the theoretical apparatus used to describe languages, in the sense that the simpler grammars are, the simpler it is for children to acquire them.

1.4 The Language Faculty

Mention of *learnability* leads us to consider the related goal of developing a **theory of language acquisition**. An acquisition theory is concerned with the question of how children acquire grammars of their native languages. Children generally produce their first recognisable word (e.g. *Mama* or *Dada*) by around the age of twelve months (with considerable variation between individual children, however). For the next six months or so, there is little apparent evidence of grammatical development in their speech production, although the child's productive vocabulary typically increases by about five words a month until it reaches around thirty words at age eighteen months. Throughout this single-word stage, children's utterances comprise single words spoken in isolation: e.g. a child may say *Apple* when reaching for an apple, or *Up* when wanting to climb up onto someone's knee. During the single-word stage, it is difficult to find any immediately visible evidence of the acquisition of grammar, in that children do not make productive use of inflections (e.g. they don't productively add the plural *-s* ending to nouns, or the past tense *-d* ending to verbs), and don't productively combine words together to form two- and three-word utterances. (However, it should be noted that perception experiments have suggested that infants may acquire some syntactic knowledge even before one year of age.)

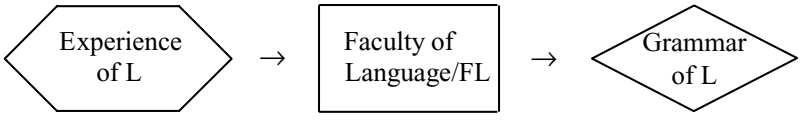
At around the age of eighteen months (though with considerable variation from one child to another), we find the first visible signs of the acquisition of grammar: children start to make productive use of inflections (e.g. using plural nouns like *doggies* alongside the singular form *doggy*, and inflected verb forms like *going/gone* alongside the uninflected verb form *go*), and similarly start to produce elementary two- and three-word utterances such as *Want Teddy*, *Eating cookie*, *Daddy gone office* etc. From this point on, there is a rapid expansion in

their grammatical development, until by the age of around thirty months they have typically acquired a wide variety of the inflections and core grammatical constructions used in English, and are able to produce adult-like sentences such as *Where's Mummy gone? What's Daddy doing? Can we go to the zoo, Daddy?* etc. (though occasional morphological and syntactic errors persist until the age of four years or so – e.g. *We goed there with Daddy*, *What we can do?* etc.).

So, the central phenomenon which any theory of language acquisition must seek to explain is this: how is it that after a long-drawn-out period of many months in which there is no obvious sign of grammatical development, at around the age of eighteen months there is a sudden spurt as multiword speech starts to emerge, and a phenomenal growth in grammatical development then takes place over the next twelve months? This *uniformity* and (once the spurt has started) *rapidity* in the pattern of children's linguistic development are the central facts which a theory of language acquisition must seek to explain. But how?

Chomsky maintains that the most plausible explanation for the uniformity and rapidity of first language acquisition is to posit that the course of acquisition is determined by a biologically endowed innate **Faculty of Language/FL** (or *language acquisition program*, to borrow a computer software metaphor) within the brain, which provides children with a genetically transmitted algorithm (i.e. set of procedures) for developing a grammar, on the basis of their linguistic **experience** (i.e. on the basis of the speech input they receive). The way in which Chomsky visualises the acquisition process can be represented schematically as in (17) below (where L is the language being acquired):

(17)



Children acquiring a language will observe people around them using the language, and the set of expressions in the language which a child hears (and the contexts in which they are used) in the course of acquiring the language constitute the child's linguistic **experience** of the language. This experience serves as input to the child's Faculty of Language/FL, which incorporates a set of UG principles (i.e. principles of Universal Grammar) which enable the child to use the experience to devise a grammar of the language being acquired. Thus, the input to the language faculty is the child's experience, and the output of the language faculty is a grammar of the language being acquired.

The claim that the course of language acquisition is determined by an innate language faculty is known popularly as the **innateness hypothesis**. Chomsky maintains that the ability to speak and acquire languages is unique to human beings, and that natural languages incorporate principles which are also unique to humans and which reflect the nature of the human mind:

> Whatever evidence we do have seems to me to support the view that the ability to acquire and use language is a species-specific human capacity, that there are very deep and restrictive principles that determine the nature of human language and are rooted in the specific character of the human mind. (Chomsky 1972, p. 102)

Moreover, he notes, language acquisition is an ability which all humans possess, entirely independently of their general intelligence:

> Even at low levels of intelligence, at pathological levels, we find a command of language that is totally unattainable by an ape that may, in other respects, surpass a human imbecile in problem-solving activity and other adaptive behaviour. (Chomsky 1972, p. 10)

In addition, the apparent uniformity in the types of grammars developed by different speakers of the same language suggests that children have genetic guidance in the task of constructing a grammar of their native language:

> We know that the grammars that are in fact constructed vary only slightly among speakers of the same language, despite wide variations not only in intelligence but also in the conditions under which language is acquired. (Chomsky 1972, p. 79)

Furthermore, the rapidity of acquisition (once the grammar spurt has started) also points to genetic guidance in grammar construction:

> Otherwise it is impossible to explain how children come to construct grammars . . . under the given conditions of time and access to data. (Chomsky 1972, p. 113)

(The sequence 'under . . . data' means simply 'in so short a time, and on the basis of such limited linguistic experience'.) What makes the uniformity and rapidity of acquisition even more remarkable is the fact that the child's linguistic experience is often degenerate (i.e. imperfect), since it is based on the linguistic performance of adult speakers, and this may be a poor reflection of their competence:

> A good deal of normal speech consists of false starts, disconnected phrases, and other deviations from idealised competence. (Chomsky 1972, p. 158)

If much of the speech input which children receive is ungrammatical (because of performance errors), how is it that they can use this degenerate experience to develop a (competence) grammar which specifies how to form grammatical sentences? Chomsky's answer is to draw the following analogy:

> Descartes asks: how is it when we see a sort of irregular figure drawn in front of us we see it as a triangle? He observes, quite correctly, that there's a disparity between the data presented to us and the percept that we construct. And he argues, I think quite plausibly, that we see the figure as a triangle because there's something about the nature of our minds which makes the image of a triangle easily constructible by the mind. (Chomsky 1968, p. 687)

The obvious implication is that in much the same way as we are genetically predisposed to analyse shapes (however irregular) as having specific geometric properties, so too we are genetically predisposed to analyse sentences (however ungrammatical) as having specific grammatical properties.

A further argument Chomsky uses in support of the innateness hypothesis relates to the fact that language acquisition is an entirely subconscious and involuntary activity (in the sense that you can't consciously choose whether or not to acquire your native language – though you can choose whether or not you wish to learn chess); it is also an activity which is largely unguided (in the sense that parents don't teach children to talk):

> Children acquire . . . languages quite successfully even though no special care is taken to teach them and no special attention is given to their progress. (Chomsky 1965, pp. 200–1)

The implication is that we don't learn to have a native language, any more than we learn to have arms or legs; the ability to acquire a native language is part of our genetic endowment – just like the ability to learn to walk.

Studies of language acquisition lend empirical support to the innateness hypothesis. Research has suggested that there is a **critical period** for the acquisition of syntax, in the sense that children who learn a given language before puberty generally achieve native competence in it, whereas those who acquire a (first or second) language after the age of nine or ten years rarely manage to achieve native-like syntactic competence. A particularly poignant example of this is a child called Genie, who was deprived of speech input and kept locked up on her own in a room until age thirteen. When eventually taken into care and exposed to intensive language input, her vocabulary grew enormously, but her syntax never developed. This suggests that the acquisition of syntax is determined by an innate 'language acquisition program' which is in effect switched off (or gradually atrophies) around the onset of puberty.

Further support for the key claim in the *innateness hypothesis* that the human Language Faculty comprises a modular cognitive system autonomous of non-linguistic cognitive systems such as vision, hearing, reasoning or memory comes from the study of language disorders. Some disorders (such as *Specific Language Impairment*) involve impairment of linguistic abilities without concomitant impairment of other cognitive systems. By contrast, other types of disorder (such as *Williams Syndrome*) involve impairment of cognitive abilities in the absence of any major impairment of linguistic abilities. This double dissociation between linguistic and cognitive abilities lends additional plausibility to the claim that linguistic competence is the product of an autonomous Language Faculty.

Given the assumption that human beings are endowed with an innate language faculty, the overall goal of linguistic theory is to attempt to uncover:

> the properties that are specific to human language, that is, to the 'faculty of language' FL. To borrow Jespersen's formulation eighty years ago, the goal is to unearth 'the great principles underlying the grammars of all languages' with the goal of 'gaining a deeper insight into the innermost nature of human language and of human thought.' The biolinguistic perspective views FL as an 'organ of the body,' one of many subcomponents of an organism that interact in its normal life. (Chomsky 2005b, p. 1)

However, Chomsky (2006, p. 1) notes that some properties of human language may reflect 'principles of biology more generally, and perhaps even more fundamental principles about the natural world'. Accordingly:

> development of language in the individual must involve three factors: (1) genetic endowment, which sets limits on the attainable languages, thereby making language acquisition possible; (2) external data, converted to the experience that selects one or another language within a narrow range; (3) principles not specific to FL. (Chomsky 2006, p. 2: FL = Faculty of Language)

The 'third factor principles' referred to under (3) 'enter into all facets of growth and evolution' and include 'principles of efficient computation' (Chomsky 2006, p. 2) and – more generally – 'properties of the human brain that determine what cognitive systems can exist, though too little is yet known about these to draw specific conclusions about the design of FL' (Chomsky 2006, fn. 6)

1.5 Principles of Universal Grammar

If (as Chomsky claims) human beings are biologically endowed with an innate language faculty, an obvious question to ask is what the nature of the language faculty is. An important point to note in this regard is that children can in principle acquire *any* natural language as their native language (e.g. Afghan orphans brought up by English-speaking foster parents in an English-speaking community acquire English as their first language). It therefore follows that the language faculty must incorporate a theory of **Universal Grammar/UG** which enables the child to develop a grammar of *any* natural language on the basis of suitable linguistic experience of the language (i.e. sufficient speech input). Experience of a particular language L (examples of words, phrases and sentences in L which the child hears produced by native speakers of L in particular contexts) serves as input to the child's language faculty which incorporates a theory of Universal Grammar providing the child with a procedure for developing a grammar of L.

If the acquisition of grammatical competence is indeed controlled by a genetically endowed language faculty incorporating a theory of UG, then it follows that certain aspects of child (and adult) competence are known without

experience, and hence must be part of the genetic information about language with which we are biologically endowed at birth. Such aspects of language would not have to be learned, precisely because they form part of the child's genetic inheritance. If we make the (plausible) assumption that the language faculty does not vary significantly from one (normal) human being to another, those aspects of language which are innately determined will also be universal. Thus, in seeking to determine the nature of the language faculty, we are in effect looking for **UG principles** (i.e. principles of Universal Grammar) which determine the very nature of language.

But how can we uncover such principles? The answer is that since the relevant principles are posited to be universal, it follows that they will affect the application of every relevant type of grammatical operation in every language. Thus, detailed analysis of one grammatical construction in one language could reveal evidence of the operation of principles of Universal Grammar. By way of illustration, let's look at question-formation in English. In this connection, consider the following dialogue:

(18) SPEAKER A: He had said someone would do something
SPEAKER B: He had said who would do what?

In (18), speaker B largely echoes what speaker A says, except for replacing *someone* by *who* and *something* by *what.* For obvious reasons, the type of question produced by speaker B in (18) is called an **echo question**. However, speaker B could alternatively have replied with a **non-echo question** like that below:

(19) Who had he said would do what?

If we compare the echo question *He had said who would do what?* in (18) with the corresponding non-echo question *Who had he said would do what?* in (19), we find that (19) involves two movement operations which are not found in (18). One is an **auxiliary inversion** operation by which the past tense auxiliary *had* is moved in front of its subject *he*. The other is a **wh-movement** operation by which the **wh-word** *who* is moved to the front of the overall sentence, and positioned in front of *had*. (A wh-word is a question word like *who/what/where/when* etc. beginning with *wh*.)

A closer look at questions like (19) provides evidence that there are UG principles which constrain the way in which movement operations may apply. An interesting property of the questions in (18B, 19) is that they contain two auxiliaries (*had* and *would*) and two wh-words (*who* and *what*). Now, if we compare (19) with the corresponding echo-question in (18), we find that the *first* of the two auxiliaries (*had*) and the *first* of the wh-words (*who*) is moved to the front of the sentence in (19). If we try inverting the second auxiliary (*would*) and fronting the second wh-word (*what*), we end up with ungrammatical sentences, as we see from (20c–e) below (key items are bold-printed/italicised, and the corresponding echo question is given in parentheses; 20a is repeated from the echo question in 18B, and 20b from 19):

(20) (a) He **had** said *who* **would** do *what*? (= echo question)
(b) *Who* **had** he said would do what? (cf. He **had** said *who* would do what?)
(c) **Who* **would** he had said do what? (cf. He had said *who* **would** do what?)
(d) **What* **had** he said who would do? (cf. He **had** said who would do *what*?)
(e) **What* **would** he had said who do? (cf. He had said who **would** do *what*?)

If we compare (20b) with its echo-question counterpart (20a) *He had said who would do what?* we see that (20b) involves preposing the first wh-word *who* and the first auxiliary *had*, and that this results in a grammatical sentence. By contrast, (20c) involves preposing the first wh-word *who* and the second auxiliary *would*; (20d) involves preposing the second wh-word *what* and the first auxiliary *had*; and (20e) involves preposing the second wh-word *what* and the second auxiliary *would*. The generalisation which emerges from the data in (20) is that auxiliary inversion preposes the *closest* auxiliary *had* (i.e. the one nearest the beginning of the sentence in (20a) above) and likewise wh-fronting preposes the *closest* wh-expression *who*. The fact that two quite distinct movement operations (auxiliary inversion and wh-movement) are subject to the same locality condition (which requires preposing of the *most local* – i.e. closest – expression of the relevant type) suggests that one of the principles of Universal Grammar incorporated into the language faculty is a **Locality Principle** which can be outlined informally as:

(21) **Locality Principle**
Grammatical operations are local

In consequence of (21), auxiliary inversion preposes the closest auxiliary, and wh-movement preposes the closest wh-expression. It seems reasonable to suppose that (21) is a principle of Universal Grammar (rather than an idiosyncratic property of question-formation in English). In fact, the strongest possible hypothesis we could put forward is that (21) holds of all grammatical operations in all natural languages, not just of movement operations; and indeed we shall see in later chapters that other types of grammatical operation (including **agreement** and **case assignment**) are subject to a similar locality condition. If so, and if we assume that abstract grammatical principles which are universal are part of our biological endowment, then the natural conclusion to reach is that (21) is a principle which is biologically wired into the language faculty, and which thus forms part of our genetic make-up.

A theory of grammar which posits that grammatical operations are constrained by innate principles of UG offers the important advantage that it minimises the burden of grammatical learning imposed on the child (in the sense that children do not have to learn e.g. that auxiliary inversion affects the first auxiliary in a sentence, or that wh-movement likewise affects the first wh-expression). This is an important consideration, since we saw earlier that learnability is a criterion of adequacy for any theory of grammar – i.e. any adequate theory of grammar must be able to explain how children come to learn the grammar of their native language(s) in such a rapid and uniform fashion. The UG theory developed by Chomsky provides a straightforward account of the rapidity of the child's

grammatical development, since it posits that there is a universal set of innately endowed grammatical principles which determine how grammatical operations apply in natural language grammars. Since UG principles which are innately endowed are wired into the language faculty and so do not have to be learned by the child, this minimises the learning load placed on the child, and thereby maximises the learnability of natural language grammars. It also (correctly) predicts that there are certain types of error which children will not make – e.g. producing sentences such as (20c–e).

1.6 Parameters

Thus far, we have argued that the language faculty incorporates a set of universal principles which guide the child in acquiring a grammar. However, it clearly cannot be the case that all aspects of the grammar of languages are universal; if this were so, all natural languages would have the same grammar and there would be no **grammatical learning** involved in language acquisition (i.e. no need for children to learn anything about the grammar of the language they are acquiring), only **lexical learning** (viz. learning the lexical items/words in the language and their idiosyncratic linguistic properties, e.g. whether a given item has an irregular plural or past tense form). But although there are universal principles which determine the broad outlines of the grammar of natural languages, there also seem to be language-particular aspects of grammar which children have to learn as part of the task of acquiring their native language. Thus, language acquisition involves not only lexical learning but also some grammatical learning. Let's take a closer look at the grammatical learning involved, and what it tells us about the language acquisition process.

Clearly, grammatical learning is not going to involve learning those aspects of grammar which are determined by universal (hence innate) grammatical operations and principles. Rather, grammatical learning will be limited to those **parameters** (i.e. dimensions or aspects) of grammar which are subject to language-particular variation (and hence vary from one language to another). In other words, grammatical learning will be limited to parametrised aspects of grammar (i.e. those aspects of grammar which are subject to parametric variation from one language to another). The obvious way to determine just what aspects of the grammar of their native language children have to learn is to examine the range of **parametric variation** found in the grammars of different (adult) natural languages.

We can illustrate one type of parametric variation across languages in terms of the following contrast between the English example in (22a) below and its Italian counterpart in (22b):

(22) (a) Maria thinks that *(they) speak French
(b) Maria pensa che parlano francese
'Maria thinks that speak French'

(The notation *(*they*) in 22a means that the sentence is ungrammatical if *they* is omitted – i.e. that the sentence **Maria thinks that speak French* is ungrammatical.) The finite (present tense) verb *speak* in the English sentence (22a) requires an overt subject like *they*, but its Italian counterpart $parlano_{speak}$ in (22b) has no overt subject. However, there are two pieces of evidence suggesting that the Italian verb $parlano_{speak}$ must have a 'silent' subject of some kind. One is *semantic* in nature, in that the verb $parlano_{speak}$ is understood as having a third person plural subject, and this understood subject is translated into English as *they*; in more technical terms, this amounts to saying that in the relevant use, the verb $parlano_{speak}$ is a two-place predicate which requires both a subject argument and an object argument, and so it must have an 'understood' silent subject of some kind in (22b). The second piece of evidence is grammatical in nature. Finite verbs agree with their subjects in Italian: hence, in order to account for the fact that the verb $parlano_{speak}$ is in the third person plural form in (22b), we need to posit that it has a third person plural subject to agree with. Since the verb $parlano_{speak}$ has no overt subject, it must have a **null subject** which can be thought of as a silent or invisible counterpart of the pronoun *they* which appears in the corresponding English sentence (22a). This null subject is conventionally designated as **pro**, so that (22b) has the fuller structure *Maria pensa che **pro** parlano francese* 'Maria thinks that *pro* speak French,' where *pro* is a null subject pronoun.

The more general conclusion to be drawn from our discussion here is that in languages like Italian, any finite verb can have either an overt subject like *Maria* or a null *pro* subject. But things are very different in English. Although finite verbs can have an overt subject like *Maria* in English, they cannot normally have a null *pro* subject – hence the ungrammaticality of **Maria thinks that speak French* (where the verb *speak* has a null subject). So, finite verbs in a language like Italian can have either overt or null subjects, but in a language like English, finite verbs can generally have only overt subjects, not null subjects. We can describe the differences between the two types of language by saying that Italian is a **null-subject language**, whereas English is a **non-null-subject language**. More generally, there appears to be parametric variation between languages as to whether or not they allow finite verbs to have null subjects. The relevant parameter (termed the **Null Subject Parameter**) would appear to be a binary one, with only two possible settings for any given language L, viz. *L either does or doesn't allow any finite verb to have a null subject*. There appears to be no language which allows the subjects of some finite verbs to be null, but not others – e.g. no language in which it is OK to say *Drinks wine* (meaning 'He/she drinks wine') but not OK to say *Eats pasta* (meaning 'He/she eats pasta'). The range of grammatical variation found across languages appears to be strictly limited to just two possibilities – languages either do or don't systematically allow finite verbs to have null subjects.

A more familiar aspect of grammar which appears to be parametrised relates to word order, in that different types of language have different word orders in specific types of construction. One type of word-order variation can be

illustrated in relation to the following contrast between English and Chinese questions:

(23) (a) What do you think he will say?

(b) Ni xiang ta hui shuo shenme
You think he will say what?

In simple wh-questions in English (i.e. questions containing a single word beginning with *wh-* like *what/where/when/why*) the wh-expression is moved to the beginning of the sentence, as is the case with *what* in (23a). By contrast, in Chinese, the wh-word does not move to the front of the sentence, but rather remains **in situ** (i.e. in the same place as would be occupied by a corresponding non-interrogative expression), so that *shenme* 'what' is positioned after the verb *shuo* 'say' because it is the (direct object) complement of the verb, and complements of the relevant type are normally positioned after their verbs in Chinese. Thus, another parameter of variation between languages is the **Wh-Parameter** – a parameter which determines whether wh-expressions are fronted (i.e. moved to the front of the overall interrogative structure containing them) or not. Significantly, this parameter again appears to be one which is binary in nature, in that it allows for only two possibilities – viz. a language either does or doesn't allow **wh-movement** (i.e. movement of wh-expressions to the front of the sentence). Many other possibilities for wh-movement just don't seem to occur in natural language: for example, there is no language in which the counterpart of *who* undergoes wh-fronting but not the counterpart of *what* (e.g. no language in which it is OK to say *Who did you see?* but not *What did you see?*). Likewise, there is no language in which wh-complements of some verbs can undergo fronting, but not wh-complements of other verbs (e.g. no language in which it is OK to say *What did he drink?* but not *What did he eat?*). It would seem that the range of parametric variation found with respect to wh-fronting is limited to just two possibilities: viz. a language either does or doesn't allow wh-expressions to be systematically fronted.

Let's now turn to look at a rather different type of word-order variation, concerning the relative position of **heads** and **complements** within phrases. It is a general (indeed, universal) property of phrases that every phrase has a head word which determines the nature of the overall phrase. For example, an expression such as *students of philosophy* is a plural Noun Phrase because its head word (i.e. the key word in the phrase whose nature determines the properties of the overall phrase) is the plural noun *students*: the noun *students* (and not the noun *philosophy*) is the head word because the phrase *students of philosophy* denotes kinds of student, not kinds of philosophy. The following expression *of philosophy* which combines with the head noun *students* to form the Noun Phrase *students of philosophy* functions as the **complement** of the noun *students*. In much the same way, an expression such as *in the kitchen* is a Prepositional Phrase which comprises the head preposition *in* and its complement *the kitchen*. Likewise, an expression such as *stay with me* is a Verb Phrase which comprises the head verb

stay and its complement *with me*. And similarly, an expression such as *fond of fast food* is an Adjectival Phrase formed by combining the head adjective *fond* with its complement *of fast food.*

In English all heads (whether nouns, verbs, prepositions or adjectives etc.) immediately precede their complements; however, there are also languages like Korean in which all heads immediately follow their complements. In informal terms, we can say that English is a **head-first language**, whereas Korean is a **head-last language**. The differences between the two languages can be illustrated by comparing the English examples in (24) below with their Korean counterparts in (25):

(24) (a) Close the door (b) desire for change

(25) (a) Muneul dadara (b) byunhwa-edaehan galmang
Door close change-for desire

In the English Verb Phrase *close the door* in (24a), the head verb *close* immediately precedes its complement *the door*; if we suppose that *the door* is a Determiner Phrase, then the head of the phrase (= the determiner *the*) immediately precedes its complement (= the noun *door*). Likewise, in the English Noun Phrase *desire for change* in (24b), the head noun *desire* immediately precedes its complement *for change*; the complement *for change* is in turn a Prepositional Phrase in which the head preposition *for* likewise immediately precedes its complement *change*. Since English consistently positions heads before complements, it is a head-first language. By contrast, we find precisely the opposite ordering in Korean. In the Verb Phrase *muneul dadara* (literally 'door close') in (25a), the head verb *dadara* 'close' immediately follows its complement *muneul* 'door'; likewise, in the Noun Phrase *byunhwa-edaehan galmang* (literally 'change-for desire') in (25b) the head noun *galmang* 'desire' immediately follows its complement *byunhwa-edaehan* 'change-for'; the expression *byunhwa-edaehan* 'change-for' is in turn a Prepositional Phrase whose head preposition *edaehan* 'for/about' immediately follows its complement *byunhwa* 'change' (so that *edaehan* might more appropriately be called a **postposition**; prepositions and postpositions are differents kinds of **adposition**). Since Korean consistently positions heads immediately after their complements, it is a head-last language. Given that English is head-first and Korean head-last, it is clear that the relative positioning of heads with respect to their complements is one word-order parameter along which languages differ; the relevant parameter is termed the **Head Position Parameter**.

It should be noted, however, that word-order variation in respect of the relative positioning of heads and complements falls within narrowly circumscribed limits. There are many logically possible types of word-order variation which just don't seem to occur in natural languages. For example, we might imagine that in a given language some verbs would precede and others follow their complements, so that (e.g.) if two new hypothetical verbs like *scrunge* and *plurg* were coined in English, then *scrunge* might take a following complement, and *plurg* a preceding

complement. And yet, this doesn't ever seem to happen: rather, all verbs typically occupy the same position in a given language with respect to a given type of complement.

What this suggests is that there are universal **constraints** (i.e. restrictions) on the range of parametric variation found across languages in respect of the relative ordering of heads and complements. It would seem that there are only two different possibilities which the theory of Universal Grammar allows for: a given type of structure in a given language must either be **head-first** (with the relevant heads positioned immediately before their complements), or **head-last** (with the relevant heads positioned immediately after their complements). Many other logically possible orderings of heads with respect to complements appear not to be found in natural language grammars. The obvious question to ask is why this should be. The answer given by the theory of parameters is that the language faculty imposes genetic constraints on the range of parametric variation permitted in natural language grammars. In the case of the **Head Position Parameter** (i.e. the parameter which determines the relative positioning of heads with respect to their complements), the language faculty allows only a binary set of possibilities – namely that a given kind of structure in a given language is either consistently head-first or consistently head-last.

We can generalise our discussion in this section in the following terms. If the **Head Position Parameter** reduces to a simple binary choice, and if the **Wh-Parameter** and the **Null Subject Parameter** also involve binary choices, it seems implausible that **binarity** could be an accidental property of these particular parameters. Rather, it seems much more likely that it is an inherent property of parameters that they constrain the range of structural variation between languages, and limit it to a simple binary choice. Generalising still further, it seems possible that all grammatical variation between languages can be characterised in terms of a set of parameters, and that for each parameter, the language faculty specifies a binary choice of possible values for the parameter.

1.7 Parameter-setting

The theory of parameters outlined in the previous section has important implications for a theory of language acquisition. If all grammatical variation can be characterised in terms of a series of parameters with binary settings, it follows that the only grammatical learning which children have to undertake in relation to the syntactic properties of the relevant class of constructions is to determine (on the basis of their linguistic experience) which of the two alternative settings for each parameter is the appropriate one for the language being acquired. So, for example, children have to learn whether the native language they are acquiring is a null subject language or not, whether it is a wh-movement language or not, and whether it is a head-first language or not . . . and so on for all the other parameters along which languages vary. Of course, children also face

the formidable task of **lexical learning** – i.e. building up their vocabulary in the relevant language, learning what words mean and what range of forms they have (e.g. whether they are regular or irregular in respect of their morphology), what kinds of structures they can be used in and so on. On this view, the acquisition of grammar involves the twin tasks of **lexical learning** and **structural learning** (with the latter involving **parameter-setting**).

This leads us to the following view of the language acquisition process. The central task which the child faces in acquiring a language is to construct a grammar of the language. The innate Language Faculty incorporates (i) a set of universal grammatical principles, and (ii) a set of grammatical parameters which impose severe constraints on the range of grammatical variation permitted in natural languages (perhaps limiting variation to binary choices). Since universal principles don't have to be learned, the child's syntactic learning task is limited to that of **parameter-setting** (i.e. determining an appropriate setting for each of the relevant grammatical parameters). For obvious reasons, the theory outlined here (developed by Chomsky at the beginning of the 1980s) is known as **Principles-and-Parameters Theory/PPT**.

The PPT model clearly has important implications for the nature of the language acquisition process, since it vastly reduces the complexity of the acquisition task which children face. PPT hypothesises that grammatical properties which are universal will not have to be learned by the child, since they are wired into the language faculty and hence part of the child's genetic endowment: on the contrary, all the child has to learn are those grammatical properties which are subject to parametric variation across languages. Moreover, the child's learning task will be further simplified if it turns out (as research since 1980 has suggested) that the values which a parameter can have fall within a narrowly specified range, perhaps characterisable in terms of a series of binary choices. This simplified **parameter-setting model** of the acquisition of grammar has given rise to a metaphorical acquisition model in which the child is visualised as having to set a series of switches in one of two positions (*up*/*down*) – each such switch representing a different parameter. In the case of the **Head Position Parameter**, we can imagine that if the switch is set in the *up* position (for particular types of head), the language will show head-first word order in relevant kinds of structure, whereas if it is set in the *down* position, the order will be head-last. Of course, an obvious implication of the switch metaphor is that the switch must be set in either one position or the other, and cannot be set in both positions. (This would preclude e.g. the possibility of a language having both head-first and head-last word order in a given type of structure.)

The assumption that acquiring the grammar of a language involves the relatively simple task of setting a number of grammatical parameters provides a natural way of accounting for the fact that the acquisition of specific parameters appears to be a remarkably rapid and error-free process in young children. For example, young children acquiring English as their native language seem to set the Head Position Parameter at its appropriate head-first setting from the very

earliest multiword utterances they produce (at around eighteen months of age), and seem to know (tacitly, not explicitly, of course) that English is a head-first language. Accordingly, the earliest verb phrases and Prepositional Phrases produced by young children acquiring English consistently show verbs and prepositions positioned before their complements, as structures such as the following indicate (produced by a young boy called Jem/James at age twenty months; head verbs are italicised in (26a) and head prepositions in (26b), and their complements are in non-italic print):

(26) (a) *Touch* heads. *Cuddle* book. *Want* crayons. *Want* malteser. *Open* door. *Want* biscuit. *Bang* bottom. *See* cats. *Sit* down

(b) *On* Mummy. *To* lady. *Without* shoe. *With* potty. *In* keyhole. *In* school. *On* carpet. *On* box. *With* crayons. *To* mummy

The obvious conclusion to be drawn from structures like (26) is that children like Jem consistently position heads before their complements from the very earliest multiword utterances they produce. They do not use different orders for different words of the same type (e.g. they don't position the verb *see* after its complement but the verb *want* before its complement), or for different types of words (e.g. they don't position verbs before and prepositions after their complements).

A natural question to ask at this point is how we can provide a principled explanation for the fact that from the very onset of multiword speech we find English children correctly positioning heads before their complements. The **Principles-and-Parameters** model enables us to provide an explanation for why children manage to learn the relative ordering of heads and complements in such a rapid and error-free fashion. The answer provided by the model is that learning this aspect of word order involves the comparatively simple task of setting a binary parameter at its appropriate value. This task will be a relatively straightforward one if the language faculty tells the child that the only possible choice is for a given type of structure in a given language to be uniformly head-first or uniformly head-last. Given such an assumption, the child could set the parameter correctly on the basis of minimal linguistic experience. For example, once the child is able to analyse the structure of an adult utterance such as *Help Daddy* and knows that it contains a Verb Phrase comprising the head verb *help* and its complement *Daddy*, then (on the assumption that the language faculty specifies that all heads of a given type behave uniformly with regard to whether they are positioned before or after their complements), the child will automatically know that all verbs in English are canonically (i.e. normally) positioned before their complements.

One of the questions posed by the parameter-setting model of acquisition outlined here is just how children come to arrive at the appropriate setting for a given parameter, and what kind(s) of evidence they make use of in setting parameters. There are two types of evidence which we might expect to be available to the language learner in principle, namely **positive evidence** and **negative evidence**. Positive evidence comprises a set of observed expressions

illustrating a particular phenomenon: for example, if children's speech input is made up of structures in which heads precede their complements, this provides them with positive evidence which enables them to set the Head Position Parameter at the head-first setting appropriate to English. Negative evidence might be of two kinds – **direct** or **indirect**. Direct negative evidence could come from the correction of children's errors by other speakers of the language. However, (contrary to what is often imagined) correction plays a fairly insignificant role in language acquisition, for two reasons. Firstly, correction is relatively infrequent: adults simply don't correct all the errors children make (if they did, children would soon become inhibited and discouraged from speaking). Secondly, children are notoriously unresponsive to correction, as the following dialogue (from McNeill 1966, p. 69) illustrates:

(27) CHILD: Nobody don't like me
ADULT: No, say: 'Nobody likes me'
CHILD: Nobody don't like me
(*8 repetitions of this dialogue*)
ADULT: No, now listen carefully. Say 'Nobody likes me'
CHILD: Oh, nobody don't likes me

As Hyams (1986, p. 91) notes: 'Negative evidence in the form of parental disapproval or overt corrections has no discernible effect on the child's developing syntactic ability.'

Direct negative evidence might also take the form of self-correction by other speakers. Such self-corrections tend to have a characteristic intonation and rhythm of their own, and may be signalled by a variety of fillers (such as those italicised in (28) below):

(28) (a) The picture was hanged . . . *or rather* hung . . . in the Tate Gallery
(b) The picture was hanged . . . *sorry* hung . . . in the Tate Gallery
(c) The picture was hanged . . . *I mean* hung . . . in the Tate Gallery

However, self-correction is arguably too infrequent a phenomenon to play a major role in the acquisition process.

Rather than say that children rely on direct negative evidence, we might instead imagine that they learn from **indirect negative evidence** (i.e. evidence relating to the non-occurrence of certain types of structure). Suppose that a child's experience includes no examples of structures in which heads follow their complements (e.g. no Prepositional Phrases like **dinner after* in which the head preposition *after* follows its complement *dinner*, and no Verb Phrases such as **cake eat* in which the head verb *eat* follows its complement *cake*). On the basis of such indirect negative evidence (i.e. observing that such structures never occur in English), the child might infer that English is not a head-last language.

Although it might seem natural to suppose that indirect negative evidence plays some role in the acquisition process, there are potential **learnability** problems posed by any such claim. After all, the fact that a given construction does not occur

in a given chunk of the child's experience does not provide conclusive evidence that the structure is ungrammatical, since it may well be that the non-occurrence of the relevant structure in the relevant chunk of experience is an accidental (rather than a systematic) gap. Thus, the child would need to process a very large (in principle, infinite) chunk of experience in order to be sure that non-occurrence reflects ungrammaticality. It is implausible that young children process massive chunks of experience in this way and search through it for negative evidence about the non-occurrence of certain types of structure, since this would impose an unrealistic memory load on them. In any case, given the assumption that parameters are binary and single-valued, negative evidence becomes entirely unnecessary: after all, once the child hears a Prepositional Phrase like *with Daddy* in which the head preposition *with* precedes its complement *Daddy*, the child will have positive evidence that English allows head-first order in prepositional phrases; and given the assumption that the Head Position Parameter is a binary one and the further assumption that each parameter allows only a single setting, then it follows (as a matter of logical necessity) that if English allows head-first Prepositional Phrases, it will not allow head-last Prepositional Phrases. Thus, in order for the child to know that English doesn't allow head-last Prepositional Phrases, the child does not need negative evidence from the non-occurrence of such structures, but rather can rely on positive evidence from the occurrence of the converse order in head-first structures (on the assumption that if a given structure is head-first, UG specifies that it cannot be head-last). And, as we have already noted, a minimal amount of positive evidence is required in order to identify English as a uniformly head-first language (i.e. a language in which *all* heads precede their complements). Learnability considerations such as these have led Chomsky (1986a, p. 55) to conclude that 'There is good reason to believe that children learn language from positive evidence only.' The claim that children do not make use of negative evidence in setting parameters is known as the **No-Negative-Evidence Hypothesis**; it is a hypothesis which is widely assumed in current acquisition research.

1.8 Summary

We began this chapter in §1.2 with a brief look at traditional grammar, noting that this is a **taxonomic** (i.e. classificatory) system in which the syntax of a given sentence is described by assigning each of the constituents in the sentence to a grammatical category, and saying what grammatical function it has. In §1.3, we noted that Chomsky takes a very different **cognitive** approach to the study of language in which a grammar of a language is a model of the grammatical knowledge (or **competence**) internalised in the mind/brain of a native speaker (hence a model of the speaker's **I-language**). We saw that Chomsky's ultimate goal is to develop a theory of **Universal Grammar/UG** which characterises the defining properties of the grammars of natural languages – a theory which is universal, explanatory and constrained, and which provides descriptively

adequate grammars that are minimally complex and hence learnable. In §1.4, we went on to look at the nature of language acquisition, and argued that the most fundamental question for a theory of language acquisition to answer is why it should be that after a period of a year and a half during which there is little evidence of grammatical development visible in the child's speech output, most of the grammar of the language is acquired by children during the course of the following year. We outlined the **Innateness Hypothesis** put forward by Chomsky, under which the course of language acquisition is genetically predetermined by an innate **Language Faculty**. In §1.5, we noted Chomsky's claim that the Language Faculty incorporates a theory of **Universal Grammar/UG** which embodies a set of universal grammatical principles that determine the ways in which grammatical operations work; and we saw that the syntax of questions in English provides evidence for postulating that syntactic operations are constrained by the following principle:

Locality Principle: Every grammatical operation is *local* in the sense that it affects the closest constituent of the relevant type

In §1.6, we went on to argue that the grammars of natural languages vary along a number of **parameters**. We looked at three such parameters, namely:

Wh-Parameter: Some languages (like English) require movement of an interrogative wh-expression to the front of an interrogative clause, whereas others (like Chinese) leave interrogative wh-expressions in situ

Null Subject Parameter: Some languages (like Italian) allow a null pronoun (= *pro*) to be used as the subject of any finite (auxiliary or main) verb, whereas other languages (like English) do not

Head Position Parameter: Some languages (like English) position head words immediately before their complements, whereas others (like Korean) position them immediately after their complements.

We hypothesised that each such parameter has a binary choice of settings. In §1.7, we argued that the syntactic learning task which children face involves **parameter-setting** – i.e. determining which of two possible settings is the appropriate one for each parameter in the language being acquired. We further argued that if parameters have binary settings (e.g. so that a given kind of structure in a given language is either head-first or head-last), we should expect to find evidence that children correctly set parameters from the very onset of multiword speech: and we presented evidence to suggest that from their very earliest multiword utterances, children acquiring English as their mother tongue correctly set the Head Position Parameter at the head-first value appropriate for English. We concluded that the acquisition of grammar involves the twin tasks of lexical learning (i.e. acquiring a **lexicon/**vocabulary) and parameter-setting. We went on to ask what kind of evidence children use in setting parameters, and concluded that they use **positive evidence** from their experience of the occurrence of specific types of structure (e.g. head-first structures, or null subject structures or wh-movement structures).

1.9 Bibliographical background

For a fuller account of the grammatical categories discussed in §1.2, see chapter 2 of Radford (2004a) or (2004b). On the nature of determiners, see Giusti (1997), Spinillo (2004) and Isac (2006). On different types of pronoun, see Cardinaletti and Starke (1999), Wiltschko (2001) and Déchaine and Wiltschko (2002). On the claim that personal pronouns are D constituents, see Postal (1966) and Abney (1987). On the claim that infinitival *to* is a tense particle, see Freidin (2004, p. 117, fn. 32). For a technical discussion of tense, see Julien (2001) and Ishii (2006a). The term *complementiser* dates back to Rosenbaum (1965, 1967) and Bresnan (1970). For more extensive discussion of the notion of I-language introduced in §1.3, see Smith (2004). Chomsky's *Minimalist Program* is developed in Chomsky (1993, 1995, 1998, 1999, 2001, 2002, 2005a, 2005b, 2006). For discussion of Chomsky's idea that language is a perfect system of optimal design, see Lappin, Levine and Johnson (2000a, 2000b, 2001), Holmberg (2000a), Piattelli-Palmarini (2000), Reuland (2000, 2001b), Roberts (2000, 2001a), Uriagereka (2000, 2001) and Freidin and Vergnaud (2001). For further discussion of the innateness hypothesis outlined in §1.4, see Lightfoot (1999), Anderson and Lightfoot (2002), Antony and Hornstein (2003), Givón (2002), Hauser, Chomsky and Fitch (2002) and Fitch, Hauser and Chomsky (2005); for a more critical view, see Everett (2005, 2006) and Sampson (2005), and for a reply to such criticism, see Chomsky's contributions to Antony and Hornstein (2003). For a textbook summary of perceptual evidence that very young infants may be sensitive to syntactic structure, see Lust (2006, §9.2.1). For evaluation of the idea that children learn languages in spite of receiving *degenerate input*, see Pullum and Scholz (2002), Thomas (2002), Sampson (2002), Fodor and Crowther (2002), Lasnik and Uriagereka (2002), Legate and Yang (2002), Crain and Pietroski (2002), Scholz and Pullum (2002), Lewis and Elman (2002) and Gualmini and Crain (2005). For discussion of the *critical period* in language acquisition, see Lenneberg (1967), Hurford (1991) and Smith (1998, 2004); on Genie, see Curtiss (1977) and Rymer (1993). On evidence of a double dissociation between linguistic and cognitive abilities, see Clahsen (2008). The idea outlined in §1.5 that grammars incorporate a set of UG principles is developed in Chomsky (1981). The Locality Principle sketched in the same section has its historical roots in a number of related principles, including the *Relativised Minimality Principle* of Rizzi (1990), the *Shortest Move* principle of Chomsky (1995) and the *Attract Closest Principle* of Richards (1997). The idea that grammatical differences between languages can be reduced to a small number of parameters is developed in Chomsky (1981). A complication glossed over in the text discussion of the Null Subject Parameter is posed by languages in which only some finite verb forms can have null subjects: see Vainikka and Levy (1999) and the collection of papers in Jaeggli and Safir (1989) for illustration and discussion. The discussion of the Wh-Parameter in the main text is simplified by ignoring the complication that some languages allow more than one wh-expression to

be fronted in wh-questions (see Bošković 2002a, Grohmann 2006 and Surányi 2006), and the additional complication that wh-movement appears to be optional in some languages, either in main clauses, or in main and complement clauses alike (see Denham 2000, and Cheng and Rooryck 2000); on *wh-in-situ* structures, see Pesetsky (1987), Cheng (1997), Cole and Hermon (1998), Reinhart (1998) and Bruening (2007). The claim made in the outline of the Head Position Parameter that all heads of a given type occupy a uniform position with respect to their complements is called into question by the behaviour of prepositions in German, most of which precede their complements, but a few of which (e.g. *entlang* 'along') follow them. Although we assumed in the text that parameters have binary settings, it should be noted that some researchers have assumed that parameters can have more than two alternative settings (e.g. Manzini and Wexler 1987). For discussion of a wide range of parametric variation between languages, see Cinque and Kayne (2005). For a critique of the idea that cross-linguistic variation is reducible to a small number of structural *parameters*, see Culicover and Nowak (2003), Newmeyer (2004, 2006) and Abeillé and Borsley (2006): for a defence of parameters, see Roberts and Holmberg (2006). For a defence of the claim made in §1.7 that parameters are correctly set by children at a very early stage in their development, see Wexler (1998). The claim that no negative evidence is used in setting parameters is made in Chomsky (1981, pp. 8–9); supporting evidence can be found in McNeill (1966), Brown, Cazden and Bellugi (1968), Brown and Hanlon (1970), Braine (1971), Bowerman (1988), Morgan and Travis (1989) and Marcus (1993) – but for potential counterevidence, see Lappin and Shieber (2007). On how children set parameters, see Fodor (2001) and Fodor and Sakas (2005). For a technical account of language acquisition within the framework used here, see Guasti (2002) and Lust (2006).

Workbook section

Exercise 1.1

Word-order parameters like the **Head Position Parameter** determine the canonical (i.e. 'basic', 'normal' or 'underlying') word order found in particular types of structure in a given language. However (as we will see in subsequent chapters), languages may have a variety of movement operations which allow particular types of expression to be fronted (i.e. preposed) and thereby be moved out of their canonical position into some new position at the front of a particular phrase, clause or sentence. For example, in a head-first language like English, both main and auxiliary verbs immediately precede their complements. Accordingly, in a sentence like *John has gone home*, the verb *gone* immediately precedes its complement *home*, and the auxiliary *has* immediately precedes its complement *gone home*. But in a question like *Where has John gone?* wh-movement means that the complement *where* of the verb *gone* is moved to the front of the overall sentence, and so the verb *gone* no longer immediately precedes its complement *where*. Likewise, the auxiliary *has* undergoes auxiliary inversion (thereby moving in front of its subject

John) and consequently no longer immediately precedes its complement *gone where*.

Below are a number of sentences taken from various plays written by Shakespeare, representing a variety of English sometimes referred to as *Elizabethan English* (because it was spoken during the reign of Queen Elizabeth the First). Elizabethan English (like present-day English) was a head-first language in which heads were canonically positioned in front of their complements. In relation to the sentences below, show how movement operations which fronted various types of expression could mask the underlying head-first setting of the Head Position Parameter in Elizabethan English.

1 Seawater shalt thou drink (Prospero, *The Tempest*, I.ii)
2 That letter hath she delivered (Speed, *Two Gentlemen of Verona*, II.i)
3 Friend hast thou none (Duke, *Measure for Measure*, III.i)
4 True is it that we have seen better days (Duke Senior, *As You Like It*, II.vii) **(w)**
5 She may more suitors have (Tranio, *The Taming of the Shrew*, I.ii)
6 Run you to the citadel! (Iago, *Othello*, V.i)
7 Came you from the church? (Tranio, *Taming of the Shrew*, III.ii)
8 What think you he hath confessed? (First Lord, *All's Well That Ends Well*, IV.iii) **(w)**
9 What will this come to? (Flavius, *Timon of Athens*, I.ii)
10 What visions have I seen! (Titania, *Midsummer Night's Dream*, V.i)

Helpful hints

Take *none* in 3, *more* in 5 and *what* in 10 to be quantifiers with a noun as their complement (and assume that the negative quantifier is spelled out as *no* if immediately followed by its complement, but as *none* otherwise). Note that 1–5 are declarative sentences (used to make a statement), 6 is an imperative sentence (used to issue an order), 7–9 are interrogative sentences (used to ask a question) and 10 is an exclamative sentence (used to exclaim amazement).

Model answer for 1

The auxiliary verb *shalt* 'shall' has the subject *thou* '$\text{you}_{\text{singular}}$' and the complement *drink seawater*.

The main verb *drink* has the complement *seawater.* If no movement operations took place in the relevant sentence, we should expect to find the word order *Thou shalt drink seawater*, with the auxiliary *shalt* immediately preceding its complement *drink seawater*, and the verb *drink* immediately preceding its complement *seawater*, in keeping with the assumption that Elizabethan English has a head-first setting for the Head Position Parameter. However, the noun *seawater* undergoes a fronting/preposing operation in order to highlight it, and this means that instead of occupying its canonical position immediately after the verb *drink*, it is instead moved to a new position at the front of the overall sentence. Likewise, the auxiliary *shalt* undergoes a separate (subject–auxiliary) inversion operation which means that instead of occupying its canonical position immediately preceding its complement *drink seawater*, it is instead moved to a new position immediately preceding its subject *thou*. The effect of these two movement operations is shown schematically below:

UNDERLYING ORDER: Thou shalt drink *seawater*
SUPERFICIAL ORDER: *Seawater* shalt thou drink

In the underlying order, the auxiliary *shalt* immediately precedes its complement *drink seawater*, and the verb *drink* immediately precedes its complement *seawater.* But preposing *seawater* and