1

Language change and grammar change

1.1   Introduction

> Ða com of more under misthleoþum
> Grendel gongan, Godes yrre bær;
> mynte se manscaða manna cynnes
> sumne besyrwan in sele þam hean.   (*Beo* 710–13)

These are four lines from one of the earliest Old English texts, the famous heroic poem *Beowulf*, which was composed over one thousand years ago. This piece of language, indeed Old English in general, is almost completely unreadable without specialized training; the most an unskilled reader will recognize is a few words still around in the language, like *of* and *under*. A word-by-word translation is: then came from moor under misty cliffs / Grendel go, God's anger bore / meant the foul-foe of-the-men / one trap in hall the high. An idiomatic translation into Modern English is: 'Then from the moor under the misty cliffs came Grendel, bearing God's anger. The foul foe meant to trap one of the men in the high hall.' Leaving aside phonological and lexical differences, which are not our concern in this book, it is not difficult to spot differences in sentence construction between these four lines and the present-day language. For instance, the word order *Then came from the moor . . .* is at best a stylistically marked option in present-day English, and the word order with the finite verb in initial position in line 3 is ungrammatical: *meant the foul foe . . .* Other differences are the combination of the verbs 'come' and 'intend' with a bare infinitive, as in *com . . . Grendel gongan*, and *mynte . . . besyrwan*. A further difference is the word order of the nominal group *sele þam hean* 'hall the high'. Beside these, the language of *Beowulf* has a system of cases and of verb endings, and there are various other syntactic differences apparent from these four lines of text. Many of these differences will be discussed or touched upon in the chapters to come, though not all of them, for it is not the aim of this book to give an inventory of the syntactic changes that have taken place in the history of English. This is a task that is best left to handbooks, such as the various volumes of *The Cambridge History of the English Language*, which

1

contain excellent and extensive digests of the work that has been done. The approach in this book will be different, in that we pursue in detail the nature and causes of a number of cases of syntactic change in the history of English. The approach that we shall take in doing so is inspired by theoretical work in the vein of Chomsky's Principles and Parameters approach to syntactic theory. Looking at historical developments from this generative perspective has important consequences for our view of syntactic change, since it means that we will focus on change in grammar as conceived of in the Principles and Parameters approach, rather than on language change.[1]

In this introductory chapter, we explicate our approach and its consequences in the realms of syntactic theory and philology. We first sketch the basic ideas underlying the generative approach to syntactic change, and show how its emphasis on the grammar of the native speaker as the object of study both sharpens and complicates the study of historical change. We will also discuss some important recent contributions to the study of English historical syntax from perspectives other than our grammar-focussed one, to achieve a more comprehensive view of the syntactic changes in the history of English that we discuss in subsequent chapters. Section 1.2 will be on grammar change from the Principles and Parameters perspective; section 1.3 on grammar change and language change; and section 1.4 will concentrate on methodological issues and presents a discussion of problems that historical data pose for the linguist in general, and the generative linguist in particular.

1.2   Historical change, language acquisition and the Principles and Parameters model

1.2.1   *Language acquisition and grammar change*

The general framework for the study of syntax adopted here is Principles and Parameters theory. This is not one single set of ideas or theoretical notions, but rather an approach to the study of language. Its nature is perhaps best captured in the following quote from Chomsky:

> The study of generative grammar has been guided by several fundamental problems, each with a traditional flavor. The basic concern is to determine and characterize the linguistic capacities of particular individuals. We are concerned, then, with states of the language faculty, which we understand to be some array of cognitive traits and capacities, a particular component of

[1] For introductions to generative syntax, we refer the reader to Radford (1997) and Haegeman (1994).

> the human mind/brain. The language faculty has an initial state, genetically determined; in the normal course of development it passes through a series of states in early childhood, reaching a relatively stable steady state that undergoes little subsequent change, apart from the lexicon. To a good first approximation, the initial state appears to be uniform for the species. Adapting traditional terms to a special usage, we call the theory of the state attained its *grammar* and the theory of the initial state *Universal Grammar* (UG).   (Chomsky 1995: 14)

It follows from this characterization that in this perspective on the study of language, the object of study is the grammar of the native speaker, to be understood as one language learner's choices for her native language with respect to the abstract parameters that are part of Universal Grammar (UG).[2] One of the core aims of generative grammar, then, is to solve what has come to be called 'the logical problem of language acquisition', i.e. the question how it is that the language learner is capable of constructing a mature grammar of her native language in a surprisingly short time, and on the basis of impover- ished evidence. The evidence available to the language learner consists of the speech output of her language environment, which contains many per- formance errors, and little to no evidence about ungrammaticality. It seems that the role of correction by parents in the language acquisition process is very limited indeed, as illustrated in e.g. McNeill (1966). The starting point for the answer to the logical problem of language acquisition is that the human language capacity, the 'initial state' or 'UG' as Chomsky and Lasnik call it, is a highly structured system of abstract principles and parameters, the values of which are filled in by the language learner on the basis of exposure to the lan- guage environment. This system is called Universal Grammar and is assumed to be part of the genetic endowment of the human species.

If we consider historical change from this perspective, it follows that the focus of investigation is on grammar change rather than on language change. This distinction is crucial and has important ramifications for how we approach historical change. The distinction between grammar change and language change correlates with the distinction usually made in generative approaches between a speaker's *competence* (knowledge and understanding) and *performance* (what the speaker does with that knowledge and under- standing). The competence of the speaker, grammatical or otherwise, is reflected by what she knows about her native language. An important method for obtaining information about this grammatical knowledge is by eliciting a native speaker's wellformedness judgements. There may be a considerable

---

[2] Following frequent practice in the literature on language acquisition, we refer to the language-learning child as *she/her*.

4   The syntax of early English

discrepancy between competence and performance. Whereas competence is supposed to constitute the steady state referred to by Chomsky, performance very often reflects that steady state imperfectly, and is influenced by factors such as slips of the tongue, tiredness, boredom, external distractions and, as the case may be when working with historical data, factors that are beyond our reach, such as the possibility of a piece of written performance like a manuscript being a late copy of a copy of a translation from Latin, written in winter when the scribe's fingers were cramped by frost, with a quill that was badly in need of sharpening, while the candle was running low. What we aim at when we study historical change from this perspective is to isolate from the set of historical data, which comprises historical written performance material, those data that reflect changes in the competence of speakers, changes in grammars.

An implication of this view of grammar change is the notion that the process of acquisition of the grammar of the native language is the main locus of change. Data from language change are of particular interest to this approach because, as Paul Kiparsky first put it, they provide a window on the form of linguistic competence (Kiparsky 1982). Instances of change can show something about the grammars of languages, because we can get a clearer view of a partially hidden abstract system when it changes from one state to another. This in turn may throw light on the precise way the theory of grammar should be formulated.

The idea that we should look primarily to language acquisition for explanations of syntactic change has evolved with increasing emphasis since it was first formulated explicitly in this context in David Lightfoot's *Principles of Diachronic Syntax* (1979). In that work, Lightfoot reacts strongly against ideas about language change in terms of drift and teleology, and the notion of diachronic grammar, which were popular in the 1970s. Such notions presuppose that language change follows, even across many generations, a predestined direction. This, according to Lightfoot, cannot be right. Each speaker constructs her own grammar afresh. The language learner does not know anything about the history of her language, and hence cannot follow any predestined process. Lightfoot argues that the language learner is endowed genetically with the ability to construct a grammar of her native language on the basis only of the speech in her language environment. Example (1) (dating back to Andersen (1973)) illustrates this:
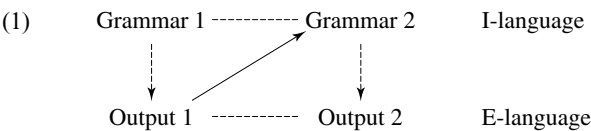
(1)   Grammar 1 ---------- Grammar 2   I-language

        Output 1 ---------- Output 2   E-language

If we see output 1 as the speech of the parent grammar (their E-language, or external language), what this diagram shows is that the language learner constructs her grammar (grammar 2) on the basis of output 1. Crucially, this happens without reference to the grammar of the parent language, since the learner has no access to that. The relationship is between output 1 and grammar 2; there is in principle no relation between grammar 1 and grammar 2. On such a view, there is no (direct) relation between the grammars of speakers, often called their I(internal)-languages, whether they belong to the same or to different generations. There is therefore no ontological basis for such notions as drift, teleology or diachronic grammar, since they presuppose that the language learner recognizes a change in progress as part of a master plan spanning many generations, to which she conforms. There is indeed no theory of change, since change is by definition synchronic, and takes place as each new language learner constructs her grammar.

There are, nevertheless, many long-term changes which often seem to follow a particular direction. This is the kind of change that inspires notions like drift and the emphasis on diachrony found in the work of grammaticalization theorists. For example, Hawkins (1990: 102–3) talks about 'diachronic universals' ('regular diachronic drifts'), and states that 'The causes of these drifts are various and constitute part of the theory of language change'. Because grammar change takes place in the acquisition process, it is a fallacy to analyse such phenomena as essentially diachronic. We discuss this more closely in section 1.3, and devote chapter 9 to a discussion of some case studies of long-term change.

Lightfoot (1979) gives an explicit methodology for work on syntactic change, which has the important quality of being falsifiable by virtue of its being explicit. Lightfoot argues that each language learner constructs her own grammar in an optimal fashion within the bounds set by the principles of UG. In his (1979) contribution, he assigns a major role to the *Transparency Principle*, a principle of grammar that requires derivations to be minimally complex, so that underlying structures are as close as possible to their surface structures. It is intended to minimize opacity in the derivation. In the course of historical development, a construction or category may acquire a number of marked characteristics through independent developments such as phonological changes, the loss of morphology and changes in word order. An example of this would be the precursors of the present-day English modals. The history of the English modals will be considered in greater detail in the next section, since it provides a good illustration of Lightfoot's view of syntactic change as well as that of others that will come up in the course of this and following chapters.

1.2.1.1   The history of English modals

Let us start with the standard assumption that in the present-day language, modals are auxiliaries, verbal function words. They occur as finite forms only, and in conjunction with an infinitive form without *to*, as in *I will do my homework*; *she might be going to the party*; *you can go to the party*. Syntactically, they function essentially as sentence modifiers: *I* in *I will do my homework* is the thematic subject of the predicate *do my homework*, not the subject of *will*. *Will* expresses future time reference, which is evidence that it is not a lexical verb. In the present-day language, modals lack inflections for person (first, second, third) or number (sg, pl), and although they have forms which reflect a present/past tense distinction historically, like *will/would*; *can/could*; *may/might*, these do not now necessarily mark a present/past distinction: for instance, the choice of *can/could* and *may/might* may reflect degrees of politeness, as in *can/could you pass me the salt?* or degrees of confidence of a positive reply as in *may/might I borrow your gold fountain pen?*

In the Old English period, modals had many more characteristics typical of lexical verbs. Evidence for this is that they could have objects and tensed clause complements, and, though they were part of the special class of so-called preterite-present verbs, they had a wider range of verbal inflections, including endings for the subjunctive mood. Lightfoot (1979) discusses the chain of events through which the Old English 'premodals', as he calls them, changed to the present-day modals as a paradigm case of a catastrophic change, a grammar change from one generation to the next. This account has been the subject of much criticism, not all of it justified: for instance, Plank (1984) argues that the history of the modals is a case of all graduality, but Warner (1990; 1993) shows that there is an abrupt shift in the behaviour of the modals in the early sixteenth century, although this is not a case of grammar change in the sense of a parameter of grammar being reset. Rather, to the extent that there is an abrupt change, it is a change in the lexical properties of modal verbs, the modal verbs being reanalysed from main verbs of sorts to auxiliaries, i.e. grammatical markers of mood.

The account in Lightfoot (1979) recognizes the following changes affecting the modals:[3]

(2)   a. Modals lost the ability to take a direct object. According to Lightfoot, this seems to have been complete in Middle English (fifteenth century) with the exception of *can*, which was a good deal more resistant (seventeenth century).

---

[3] Lightfoot adds a fifth change to this list, based on a highly theory-internal word order argument. We have omitted this for the sake of clarity.

b. Most premodals belonged to the inflectional class generally known as 'preterite presents'. The notable thing about this class is that the third person sg did not have the usual -*eþ* ending. Gradually, all the non-premodals of this class were lost. As a result, the premodals became a morphologically unique class.

c. Because of phonological similarities in the endings, the opposition between present and past as one of tense, and indicative and subjunctive as one of mood became increasingly opaque, so that the present and past forms and levelled subjunctive forms acquired separate modal meanings.

d. There were changes connected with the rise of the *to*-infinitive. In Old English, the premodals were never followed by *to*. The *to*-infinitive was firmly established in the course of the fourteenth century, except with premodals. Lightfoot concludes from this that at this stage the premodals were already beginning to be recognized as a separate class.

Following these changes, the premodals came to function as a separate class inflectionally, syntactically and semantically. Evidence for this is that the premodals (now modals) ceased to display a number of typically (main) verbal characteristics:[4]

(3)   a. They ceased to occur as infinitives.
         b. They could no longer occur with -*ing*-affixes.
         c. They could no longer occur in clusters.
         d. They could no longer occur with *have* and with -*en*-affixes.

According to Lightfoot, the modals have now acquired too many exception features to be learnable as lexical verbs. The Transparency Principle then predicts a reanalysis; the form of this reanalysis is constrained by other principles of grammar, and in this case the premodals changed into a different word category: that of auxiliaries, grammatical function words. In this view of the history of modals, the premodals were verbs and in one fell swoop underwent a radical categorial reanalysis, changing into modal auxiliaries.

While much of the ideology of Lightfoot's approach (1979) still stands, the Transparency Principle has proved to be an undesirable and superfluous addition to the theory of grammar. It is undesirable because it has no possible formal characterization like other principles of grammar, as it is not clear what opacity in a derivation really is. Also, it is implicit in the way Lightfoot illustrates the Transparency Principle that reanalyses are only forced as the result of accumulating exception features. This is not necessarily correct, as we will see below. Roberts (1985) argues that the Transparency Principle is superfluous

---

[4] The changes listed in (3) should be seen in perspective: the four changes reduce to one, i.e. the loss of nonfinite forms. But it is not the case that the modals before the reanalysis occurred in nonfinite forms on a large scale, and some of them (*e.g. may, must*) never had any nonfinite forms, as discussed in Warner (1983).

in that its results are incorporated in the parameter-setting approach to lan-
guage acquisition formulated in Chomsky (1981). This will be explained
further below.

1.2.2   The Principles and Parameters model

Work in the Principles and Parameters model has dominated the
generative scene since the development of Chomsky's 1981 theory of
Government and Binding (GB). In GB theory, UG is organized in terms of a
number of subsystems or modules, which interact with each other. One impor-
tant subsystem is the theory of Government, which started life as a structural
recasting of the notion of government in traditional grammar. Thus the head
of a constituent, say a verb or a preposition, governs its complement in a
constituent structure. A second subsystem is the theory of Binding, which
defines the grammatical conditions on the reference of nominal constituents:
anaphors like reflexive pronouns obligatorily refer back to (are bound by) a
subject antecedent in a local domain such as a tensed clause, so in *John likes
himself*, *himself* is bound by *John*, but in **John expects that Mary likes himself*,
it isn't. Pronouns may refer back in the discourse, but not to a noun phrase in
a local domain. Full noun phrases have their own reference.

Subsystems of grammar consist of quite general principles and of parame-
ters. Parameters define the dimensions along which languages may differ from
each other. As an example, we will consider in some more detail the theory of
Case, which is closely related to the theory of government. Consider the fol-
lowing bits of constituent structure (VP is a verb phrase, PP a preposition
phrase, IP an inflection phrase in which inflection for tense and agreement is
'coded'):

(4)
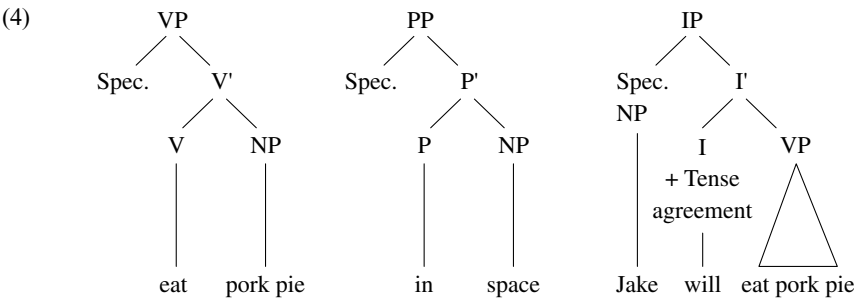
The basic principle of constituent structure is that each constituent has a head
(V in VP; P in PP, I in IP etc.) with lexical properties determined in the lexicon.
Heads are governors. Some heads are also case markers. In present-day

English, verbs and prepositions assign object case to their complements; a
tensed I-head assigns nominative case to the subject, which is in the specifier
of IP. Thus, for present-day English, Case is an abstract notion, since morpho-
logical case is visible only when the NP in question is a personal pronoun
(*he/him*, *she/her*). Case theory consists of the following general components:

(5)   a. *the Case Filter*, which stipulates that each NP must have one and only
            one case
         b. an inventory of heads which are possible case markers
         c. a definition of ways in which case can be assigned

(5a) is a good example of a principle of case theory, and is universal. In (5b)
and (5c), parameters come into play. Suppose that UG makes available a
possibility of case marking heads, but not all languages use all options. A
difference between Old English and the present-day language is that where the
present-day language has the case markers as in (4), Old English has adjectives
added to this inventory; adjectives may take complements that have dative or
genitive case, an option that was lost in the course of the Middle English
period. Contrast the Old English (6a) with Modern English (6b) and the
impossibility of (6c).

(6)   a. þeah hit þam cynge   ungewill   wære
            though it   the king (D) displeasing was
            'though it was displeasing to the king'   (*ChronE*(Plummer) 1097.22)
         b. though it was displeasing to the king
         c. *though it was displeasing the king

A complicating factor here is that adjectives combine only with dative or gen-
itive case, typically lexically selected cases. This brings us to (5c): there are at
least two ways in which Case can be assigned. It is assumed that in present-
day English, Case is assigned exclusively under structural conditions such as
those in (4) above. But lexically selected cases appear to be different: they are
probably specified in the lexical properties of the head, and therefore lexically
associated with that head, rather than purely structurally determined. Also,
the option of having lexical case probably presupposes that the language in
question is able to signal those cases by means of morphological case endings.
Languages differ, then, in the extent to which they have lexical cases. If we con-
sider the loss of case marking by adjectives in Middle English, there are at least
two ways of formulating the grammar change that must be associated with this
loss. The first could be that adjectives like *ungewill* in (6) dropped out of the
English-specific inventory of heads that were case markers, which would be a
change involving the inventory of case-marking heads; the second could be
that English lost the typically lexical cases dative and (objective) genitive,

probably because the Old English system of morphological case marking was lost. Since dative and genitive were the cases combining with adjectives, adjectives ceased to be case selectors. The change would then be a change in the ways in which case could be assigned. The latter view is the more interesting one. It is certainly the one with the most general validity. Dative and genitive cases did not only cease to combine with adjectives; the dative and (object) genitive cases were lost generally, with the loss of morphological cases, as we will see in chapter 3.

In a parameter-setting model of acquisition and change, the task of the learner is to decide, on the basis of the evidence in her language environment, how to fill in the values for the various parametric options allowed by UG. Choosing the values for the parameters for any particular language is the main task of language acquisition. With respect to the changes in Case marking by adjectives discussed above, we could say that in the Old English period, the language learner was able to incorporate lexical Cases in her grammar because the system of morphological case distinctions (in combination with some other properties) enabled her to learn a distinction between structural and lexical Cases. This is what, according to Roberts (1985), makes the Transparency Principle superfluous, since the nature of the acquisition process is such that the optimal grammar will be chosen. Roberts (1985, 1993) also suggests an explicit parametric account of the history of the English modals. While subscribing to Lightfoot's story of the history of modals as essentially a change in word-class resulting from the loss of specific main verb characteristics, Roberts shows that in addition, this categorial reanalysis interacts with and is furthered by other instances of grammar change, such as changes in verb placement, and changes in the system of subject–verb agreement. This makes it clear that the historical fate of the English modals was not necessarily shaped as a random accumulation of exception features, leading to a change in category forced by the Transparency Principle. The changes affecting the modal verbs interacted with other, independent changes.

1.2.3   *More on language acquisition and grammar change*

The general spirit of the Principles and Parameters approach to language acquisition and grammar change should be clear by now: language-learning children, on the basis of a richly structured innate UG, construct a grammar of their native language on the basis of the language they hear being spoken around them. There is no consensus in the literature on how children proceed to do this, and this lack of consensus makes itself felt in acquisition-oriented work on grammar change. In the following subsections, we give a