The Syntax of Early English

OLGA FISCHER
ANS VAN KEMENADE
WILLEM KOOPMAN
WIM VAN DER WURFF

Contents

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

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 impoverished evidence. The evidence available to the language learner consists of the speech output of her language environment, which contains many performance 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 language 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 understanding). 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*.

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
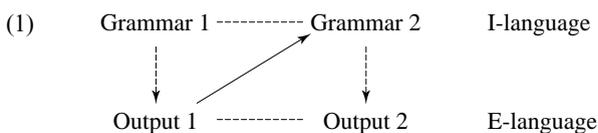
If we see output 1 as the speech of the parent grammar (their E-language, or external language), what this diagram shows is that the language learner constructs her grammar (grammar 2) on the basis of output 1. Crucially, this happens without reference to the grammar of the parent language, since the 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 language 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 important 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 parameters. 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 following 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)
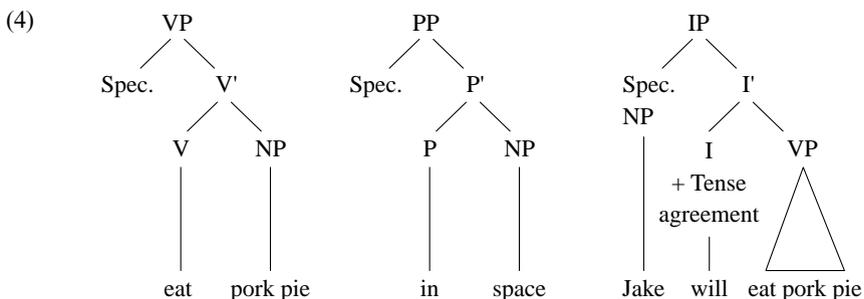
The basic principle of constituent structure is that each constituent has a head (V in VP; P in PP; I in IP etc.) with lexical properties determined in the lexicon. Heads are governors. Some heads are also case markers. In present-day

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 morphological 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 genitive 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 consider 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

flavour of the kinds of debates that currently dominate the scene, which really revolve around two interacting issues: the nature of the language evidence which is the potential input to the acquisition process; and the question of what it is that children actually do with this potential input: do they construct a grammar which matches this input as closely as possible? Do they create fragments of grammar that are later put together and may then turn out to be (partially) conflicting? Or do they keep an overall grammar in mind during the process? With respect to any case of grammar change, this leads to the question of what the change in the language environment is (to the extent that it is observable in our historical data set), how this is caused (if that question can be answered), how it triggers a grammar change, and how in turn the grammar change is reflected in the data set produced by the next generation. We will now review some of the literature with these issues in mind.
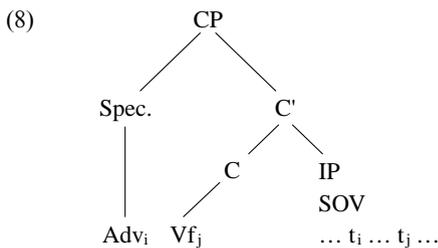
1.2.3.1 Degree 0 learnability and robustness

In the course of the 1980s, there was increasing emphasis on the question of how the language learner comes to set parameters in a different way. With this development, it has become important to appreciate the nature of the triggering evidence for the language learner. Lightfoot (1991) emphasizes very prominently that we should try to find acquisition triggers for the resetting of parameters, i.e. we should try to formulate clearly what the changes in the language environment are that induce the language learner to set a parameter differently. Such evidence must be robust, Lightfoot argues; in Lightfoot 1997, he quantifies this robustness as: exceeding a threshold level of thirty per cent of the potential of environments. An additional constraint that he formulates on the robustness of evidence is that it should come from simple, unembedded clauses. This is called *degree 0 learnability*. The rationale behind it is that the language environment of the learner consists largely of main clauses. They should therefore contain the evidence for the resetting of a parameter. Hence, if a language change formulated as a parametric change has to appeal primarily to embedded clauses for evidence, this does not qualify as an appropriate explanation for that change, according to Lightfoot. For example, modern Dutch is a 'Verb-Second' language, which means that the finite verb occupies the second constituent position in any main clause, and the first constituent position can be of any category or function: subject, object, adjunct, NP, PP, AdvP etc. Some examples are given here:

(7)   a. Jan *heeft* gisteren   het boek aan Marie *gegeven*.
           Jan has   yesterday the book to Marie given   (subject NP-first)
        b. Gisteren *heeft* Jan   het boek aan Marie *gegeven*.
           yesterday has   John the book to   Marie given   (adverb-first)

    c. Het boek *heeft* Jan   gisteren   aan Marie *gegeven*.
      the book has John yesterday to Marie given   (object NP-first)
    d. Aan Marie *heeft* Jan   gisteren   het boek *gegeven*.
      to   Mary has John yesterday the book given
      'John gave the book to Mary yesterday.'   (indirect object PP first)

It has become standard to assume that Dutch is an SOV language in the sense that objects precede the non-finite verb in surface word order. In main clauses, some constituent is moved to the specifier of CP, and the finite verb is moved to C, as in the following abbreviated structure:

(8)

```
                    CP
               ____/  \____
          Spec.           C'
            |           __/  \__
            |          C       IP
            |         / \      SOV
          Adv_i    Vf_j     … t_i … t_j …
```

First-language learners acquiring Dutch have to work out that in main clauses the first constituent position is a category-neutral one, a topic position rather than, say, a subject position, and that this position is followed by the finite verb, regardless of the structure of the rest of the clause. Since the input from the language environment contains many clauses which are subject-first, the crucial clue for the language learner should come from those clauses which are non-subject-initial, i.e. sentences like (7b-d). According to Lightfoot (1999: 153), statistical counts for Dutch, German, Norwegian and Swedish show that simple main clauses are subject-initial in about seventy per cent of the cases. The conclusion is, then, that the thirty per cent of non-subject-initial clauses suffice to trigger the Verb-Second grammar of Dutch, and that a threshold of robustness need not be more than thirty per cent. There is nothing magical about this number, but it should give an idea of what Lightfoot means by robustness.

1.2.3.2 Input matching

    Robin Clark and Ian Roberts (1993) offer an explicit perspective on the relation between acquisition and change in 'A computational model for language learnability and language change'. According to Lightfoot (1999), their approach to language learning is essentially what is called an 'input-matching' one: the learner, endowed with UG, constructs grammars and evaluates those grammars with respect to the input she hears in the language environment. In principle, all possible grammars are available to the learner.

Clark and Roberts propose, following Clark (1990; 1992) that *genetic algorithms* provide a computational model of learning for a principles and parameters theory. In their words:

> Genetic algorithms model the basic process of natural selection in the biological world: how certain patterns of genetic material are more adapted to their environment (i.e., fitter) than others, and hence tend to reproduce at the expense of others. Our account of language learning is analogous: the input text is the analogue of the environment, and so 'fitness' means consistency with this; parameter settings correspond to the genetic material of the biological world (and so a whole grammar would be a genome). Successful combinations of parameter settings 'reproduce' (i.e. contribute to the formation of new hypotheses about the target grammar) at the expense of others. In this way, the learning mechanism gradually eliminates 'unfit' hypotheses (those that are not consistent with the input text) and arrives at a single fittest grammar. Since nothing in the approach requires this grammar to be consistent with the one that underlies the input text, learners may arrive at final-state systems that differ from those of their parents; this, in essence, is our solution to the logical problem of language change.   (1993: 303)

Clark and Roberts propose that parametric change typically occurs when, because of various factors (phonological changes, morphological changes, etc.), the input data do not unambiguously force the setting of certain parameters to a definite value; several alternative grammars can adequately account for the input stream and so the choice of grammar cannot be uniquely determined on the basis of the language environment. This can happen when the evidence presented to the learner is formally compatible with a number of different and conflicting parameter settings. Clark and Roberts suggest that in such a situation the learner 'will turn in on itself, abandoning external pressure, and rely on its own internal structure to select from the alternatives at hand' (1993: 302).

This model of language learning assumes, then, that learners match all the potential values for each parameter against the language environment. Only those that fit the input data are then combined into larger grammar fragments, which are again matched, until one grammar is arrived at. This presupposes two important and contentious notions: the first is that this approach is highly UG-driven, in the sense that the learner actively explores all the options potentially offered by UG. This is in stark contrast to the approach in Lightfoot (1991), which gives pride of place to how the language environment triggers the acquisition of new parameter values. Clark and Roberts, on the other hand, postulate that the learner herself fuels diachronic change when confronted with a situation where the input is inconclusive. Thus, there is not necessarily a direct trigger in the language environment, merely a situation of indeterminacy that

the learner resolves with an appeal to UG. This is a situation not allowed for by Lightfoot, and it remains to be seen if that is correct for all language change.

A second contentious feature of this approach is that it assumes that the learner tries to match the input from the language environment in full. We will now briefly look at a case study by Alison Henry (1997), in which both these contentious issues play an important role. Henry studies syntactic change in progress in Belfast English imperative constructions and observes that, from one generation to the next, children ignore some of the positive evidence that they receive and adopt a grammar which systematically deviates from the parent grammar. Belfast English is undergoing a remarkable syntactic change in imperative constructions; the oldest of three generations has two types of imperative constructions, one with and one without inversion:

(9)   a. You sit down
        b. Sit you down
        c. Sit everybody down

Henry shows on the basis of a number of distributional criteria that the proper analysis of inverted imperatives in what she calls the 'Unrestricted Inversion' dialect has the finite verb in the C-position, much as in the Dutch Verb-Second constructions discussed above.

(10) $[_{CP} [_C$ Sit$_i$ $[_{AgrSP}$ you . . . t$_i$ . . . down]]]

This is the only construction left in Belfast English in which lexical finite verbs move to C; other contexts have *do*-support here. In the dialect of the oldest generation, when the verb is passive or unaccusative (i.e. where the subject is the underlying object), the subject may appear in AgrSP, as in (10), corresponding to (11a–b), but it may also remain in object position, as in (11c–d):

(11)   a. Be you picked for the team
        b. Be you going out of the door when they arrive
        c. Be picked you for the team
        d. Be going you out of the door when they arrive

(11c–d) are analysed as follows:

(12) $[_{CP} [_C$ be . . . $[_{VP}$ picked you for the team]]]

The middle generation also has inverted imperatives, but they are now restricted to constructions with passives and unaccusative verbs:

(13)   a. Go you away
        b. Arrive you on time
        c. Be picked you for the team
        d. *Eat you those vegetables
        e. *Run you around the room

Distributional evidence shows that the structure of inverted imperatives in what Henry calls the 'Restricted Inversion' dialect is quite different from those in the Unrestricted Inversion dialect. Where the structure of (9b) is as in (10), that of (13a) is as in (14):

(14)   [. . . [$_{VP}$ go you away]]

In other words, the only inverted imperatives now left are the ones we have in (13a–c), analysed as in (14): movement of lexical finite verbs as in (9b) analysed as (10) is no longer possible.

The transition from (9)–(12) to (13)–(14) is an interesting one, for a variety of reasons. With respect to the issues at hand, its most significant aspect is that the learners of the younger generation must have been exposed to a significant number of inverted imperatives during the acquisition process (perhaps imperatives are the most frequent type of sentence addressed to small children). They indeed arrived at a grammar which accommodates the superficially similar verb-first construction (13a–b), but the relevant sentences have a different structure, that in (14). Henry argues that this happened because some other aspects of the grammar militated against constructing a grammar in which lexical finite verbs move to C. This implies that learners made a UG-driven choice, ignoring part of the input data (the Unrestricted Inversion data), and accommodating another part of the input data (inverted unaccusatives) in a way compatible with the rest of the grammar.

Henry argues on the basis of this that the task of the language learner does not seem to be to set the parameters corresponding to all the data in the input. Rather, children select the grammar, from those made available by UG, that can accommodate the majority of data in the input, and may ignore other data, unless they are quite robust. Thus children can adopt grammars different from those of their parents, and be major contributors to language change. It is interesting to see that Henry's approach, like that of Clark and Roberts, is a strongly UG-driven one, where learners are assumed to run their UG-options by the language input in a continuous process of hypothesis creation and verification. Where the approaches differ is that, contra Clark and Roberts, Henry explicitly argues that learners leave part of the input unmatched.

1.2.3.3   Cue-based learning

The latest contribution by Lightfoot to the ongoing debate about language learning and change is Lightfoot (1999), in which he adopts ideas from some of the most recent work on language acquisition (the reader is referred to Lightfoot's book for references). He argues against the input-matching

approach: on the one hand, because it assigns too large a role to UG; on the other, because learners do not always match the input, a point that was discussed above in connection with Henry's work. Lightfoot pursues some recent work that argues in favour of what is called 'cue-based learning'. Where the cue-based child perhaps differs most sharply from the input-matching child is that the cue-based child does not explore all the parametric options that are made available by UG. Rather, the cue-based learner remains true to Lightfoot's earlier triggering experience, where the trigger now feeds into cues. We may think of cues as pieces of structure, little grammar fragments, deduced by the learner from robust pieces of degree 0 evidence in the language environment. As an illustration, we follow up the discussion on robustness above. The reader will recall that the trigger for acquiring the Verb-Second grammar of languages like Dutch presumably consists of the subset of main clauses introduced by a non-subject. What children have to learn is that the structure of those sentences is something like in (15), where XP can be a range of different phrases:

(15) $[_{CP}$ XP $[_C$ Vf $[_{IP}$ SOV $]]]$

Lightfoot (1999: 152–3) argues that the cue for the learner is that a robust number of main clauses begin with an arbitrary phrasal category. The account for why this is accompanied by movement of the finite verb to C must come from a UG condition saying that material in Spec,CP must be licensed by a lexically filled C. This is because the learner cannot know that movement of the finite verb is obligatory, since she does not have access to ungrammatical data: she does not hear sentences with [XP . . . Vf] that are marked with a star. This part of the account must therefore come from a UG condition. Hence, the trigger is, according to Lightfoot, the thirty per cent of main clauses beginning with XP, to which the learner assigns a piece of structure, the cue:

(16) $_{SpecCP}$ [XP]

This is an abstract representation of a partially analysed syntactic structure. A cue is, therefore, an element of I-language, which in turn feeds into parameter settings. Thus, the child gradually builds up a grammar, following a learning path. One point that is less than clear in this approach is that Lightfoot assumes that there will be a robust trigger in the language environment, hence a cue, for all and any of the parameter settings. One may reasonably wonder whether this is always true, and what the learner does if this is not the case. The latter situation is the very one that Clark and Roberts are concerned with: situations where the input is inconclusive; this is, according to Clark and Roberts, precisely when learners invoke some default value dictated by UG, or 'turn in on themselves', as they call it.

1.2.4   Synchronic variation and grammar competition

In the first few sections of this chapter, we outlined the generative approach to syntactic change in terms of the resetting of parameters. Of primary interest in this view are clusters of changes, like those in the history of the modals, which are analysed by Lightfoot as the result of a single underlying change (a categorial reanalysis from verb to auxiliary). In the scenario of changes sketched by Lightfoot, this kind of change is typically abrupt. The emphasis is on discontinuity, which follows from the fact that its focus is on the moment of grammar change rather than on the gradual change in the language environment. Generative work on syntactic change is often criticized for its emphasis on abruptness, which does not in general sit well with the surface graduality of language change. Let us evaluate as an example the changes involving case marking by adjectives discussed above. Adjectives in Old English could take nominal complements with dative or genitive case, and this option was lost in the course of the Middle English period. English generally lost the dative and (object) genitive cases, typically lexically selected cases. This would be a change in the ways in which case could be assigned. Verbs, adjectives and prepositions could no longer assign lexical cases. Supposing we look upon the availability of inherent case as a parameter, we might assume that this parameter was reset in the second half of the thirteenth century.[5] But after that date, examples of adjective + object will still be found, side by side with adjective followed by a PP. The examples in (17) illustrate this:

(17)   a. and tok   hem   out that were him lieve
             and took them out that were him dear
             'and took out those that were dear to him'
                                         (Gower *Confessio Amantis* 2.3395)
         b. This man to folkes alle was so leef
            this man to people all   was so dear
            'This man was so dear to all people'   (Hoccleve *Jonathas* 170)

If reanalyses are as radical as suggested, this situation is unexpected. This is also a problem from a theoretical point of view. One of the core ideas in the parameter-setting model is that when a parameter is set by the language learner, it is set once and for all. It then becomes a problem to deal with forms that suggest the opposite setting of the parameter. Forms that suggest the older parameter value, in this case the possibility of assigning lexical case, should presumably be handled by the learner in terms of some kind of adaptive rule. But it is questionable whether a restrictive theory of grammar should

---

[5] Arguments for this can be found in van Kemenade (1987), chapter 6.

allow adaptive rules that say the opposite of what the parameter says. We see then that there is considerable friction between the parameter-setting model of grammar change, which basically emphasizes discontinuity, and the fact of synchronic variation between older and newer forms in the language.

In ongoing work by Anthony Kroch and his associates (e.g. Kroch 1989; Pintzuk 1991), a view of synchronic variation and change has been developed which deals to some extent with this friction. The idea is that speakers are in a sense bilingual speakers of their native language: they acquire more than one grammar, and these grammars can differ with respect to one parameter setting. In the case at hand, they would have two grammars, which differ in that verbs, adjectives and prepositions can assign lexical cases in the one, and they cannot do so in the other. These grammars are in competition; there is a period of variation between the two, and eventually only the newer grammar remains as an option. Grammar change, according to Kroch, generally follows an S-shaped curve as in the diagram (18), in which the rate of occurrence is plotted on the y-axis, the time course on the x-axis: the change takes off slowly, then gathers momentum and there is a period of sharp rise, followed by a gradual petering out when the change nears completion:

(18)
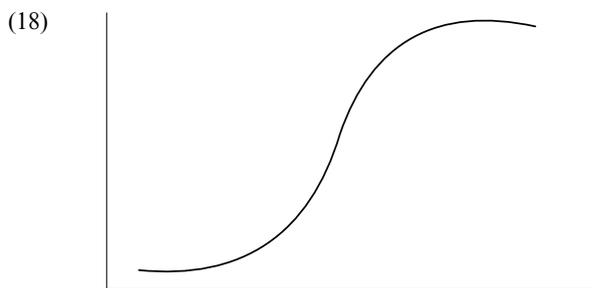
An important feature of modelling change in this way is that it allows the linguist to track clusters of changes. As noted above, grammar change is often seen on the surface as a cluster of changes spinning off from the grammatical reanalysis. It is then to be expected that changes in such a cluster occur at the same rate. Kroch (1989) calls this *the constant rate hypothesis*. Conversely, we can say that, if a cluster of changes follows the same quantitative curve, there is an argument for saying that the changes are grammatically related. Kroch (1989) illustrates this in an interesting way with the rise of *do*-support in the early Modern English period: having reworked the quantitative data gathered by Ellegård (1953) and established that the various contexts in which *do*-support is being established (negative sentences, various types of questions) follow the same curve of change, he presents his hypothesis that in the

grammar, *do*-support replaces movement of lexical finite verbs by showing that their curves of change are each other's converses, as in the following idealized graph:

(19)
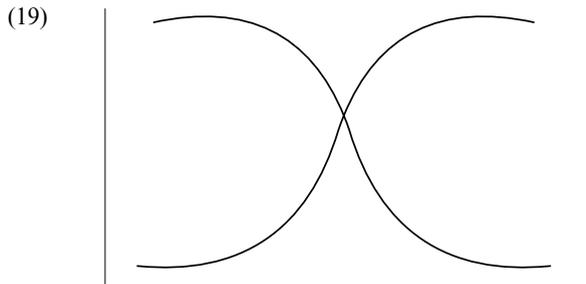
By showing that *do*-support and the contexts in which the lexical finite verb is moved have a complementary distribution, one gradually rising at the expense of the other, Kroch shows the analytical power of the constant rate hypothesis as a statistical tool in tracking down grammar change. Work from this perspective is currently expanding, due to the increasing availability of language corpora that allow for syntactic research.

It should be noted that this framework is primarily interested in synchronic variation between older and newer forms that is parametrically determined; it is designed to monitor the time course of change; in and of itself, it has very little to say about the motivations that push any particular change forward.

A good deal of synchronic variation is not conditioned by grammar-internal factors. Language contact of various kinds and social factors are other important sources of synchronic variation. Here again, we make a distinction between grammar and language, or I-language and E-language: language contact and social factors are facts of E-language; these may trigger grammar change because they change the language environment and hence may change the input for the language learner. But they are phenomena that can be distinguished from grammar change itself.

1.3   Grammar change and language change

So far we have discussed the notion of grammar change, as distinct from language change, concentrating on the acquisition process as the locus of (abrupt) grammar change. We have occasionally mentioned the friction between such an approach and the graduality of language change observable in the data. In this section, we want to broaden this perspective and discuss

some literature that is mainly concerned with language change. In doing so, we have two reasons. The first is to clarify further the distinction between grammar change and language change that we make in this book; discussing other approaches will put in relief what we want to achieve in the following chapters. The second is that important work has been done from a variety of perspectives on constructions discussed in this chapter and elsewhere in the book, with which we compare and contrast our approach here. The scholars whose work we will discuss would shrink from making a distinction between grammar and language, hence between grammar change and language change; indeed they probably feel this is wrong in principle, and shun the abstractness implicit in the notion of grammar as we see it here. In 1.3.1 we discuss work on the history of the modals in English, formulated from various lexicalist perspectives. In 1.3.2 we discuss work on impersonal constructions in Old English in which it is argued that lexical changes take place on the basis of surface similarities between uses. In 1.3.3 we discuss the kind of long-term change called grammaticalization, and suggest how grammaticalization as a phenomenon can be made compatible with our approach.

1.3.1   Prototypes and notions

In section 1.2.1.1 a summary was given of the basic changes affecting modal verbs in the history of English. Lightfoot (1979) believed that these set the scene for a radical change in word class from main verb to auxiliary. Anthony Warner and John Anderson have each tried to arrive at a better understanding of the nature of word classes and the changes in the modals, along the way refining the cataclysmic story told by Lightfoot to one that emphasizes graduality. The result is a more diffuse approach to syntactic categories, and hence to changes in word category, one that accommodates the diffuseness of the language facts in a manner that is sharply different from what is customary in the Principles and Parameters perspective.

Warner (1990; 1993) extends the theory of semantic prototypes formulated by Eleanor Rosch to grammatical categories.[6] At the heart of Rosch's approach is a general theory of human categorization, in which categories are defined in terms of opposing clusters of properties and in which certain groups of properties may be more central or prototypical, others less so. Warner's assumption with respect to word classes is that classification is done by grouping relevant similarities and differences. A word class such as 'verb' exists in opposition to other classes, so that a group of properties typical of one class (for verbs, this

---

[6] Surveys of this approach can be found in Rosch (1978; 1988).