Teaching the History of the English Language. Its position in the university curriculum and its relation to linguistic theory

Fischer, O.C.M.

Published in:

Citation for published version (APA):
Teaching the history of the English language: 
Its position in the university curriculum 
and its relation to linguistic theory

Olga Fischer 
University of Amsterdam

1. Historical linguistics in the university curriculum

I would like to address a number of questions I have struggled with both as a teacher of historical linguistics, and in my historical linguistic research, in the hope that others may profit from a discussion of them. There are two main problems or questions, which I have tried to capture in the title of this paper, and which are, it seems to me, very much linked to one another. The central point in both is the use of theory. What theoretical framework, if any, should we, historical linguists, use to do our research in, and how useful is theory or a theory for our teaching? The use of linguistic theory in historical language teaching is itself a topic that is linked to yet another question, and that is: what is the position of historical linguistics in the university curriculum? This was in fact the topic of a workshop I organized together with Nikolaus Ritt of the University of Vienna for a session at the ESSE conference in Hungary in 1997. Quite a number of points were raised there which are, I believe, relevant to the discussion here. I have put some of them under (1) and will briefly discuss each of them but not always in the order in which I have given them.

(1) (a) What is the aim of our teaching 
(b) The time available 
(c) Knowledge of and about language 
(d) The need for communicability between the subfields 
(e) ‘Alterity’ 
(f) The relation between linguistic theory and language change

In most university curricula nowadays we have but little time to teach our students an ever-increasing field of knowledge. We must teach them what they need with as little

1 After the workshop, a website was set up about this topic, where the contributions to the workshop can be found as well as the responses and the discussion that took place. The address is http://www.univie.ac.at/Anglistik/hoe.
redundancy as possible. It seems also clear that the aim of our teaching, especially at university level, must be not only the acquirement of knowledge, of facts about English and the history of English, but also and primarily an understanding of those facts, which boils down to an understanding of the nature of language. In other words, an educational perspective seeks to impart not just practical skills but also a conceptual frame of reference, some understanding of principles for the organization of facts, and following from that, an ability to relate what has been learned, to new unpredictable circumstances. After all a student should be able to operate outside the narrow confines of an English study once he or she has left the university, and entered the world at large.

So, the first point on the list in (1) leads us automatically to the second and third points, but also in a way to the fourth because it is lack of time and the need to avoid redundancy that forces us to recognize the need for links between the various subfields (links between literature and linguistics, between diachronic and synchronic investigations of language, between language theory and socio- and psycholinguistics, communication studies etc.). But this need for links between the subfields also loops back to the point about understanding the nature of language. It is clear that interdisciplinarity is highly conducive to a greater understanding of the phenomenon of language itself: we must approach it from various directions to achieve fuller knowledge. We must try to answer some of the essential questions about language from different perspectives.

We may raise a number of essential questions; a first one is: What are the external and internal forces that shape language and its development?

Here we see the link between the study of the theory of grammar, language history and socio- and psycho-linguistics. Language is the product of speakers, of the way their brain works, but speakers live in specific socio-cultural surroundings, which influence and shape their language. In addition, language itself is a historical product, and its own history may influence the shape of the language too, as we can see in the process of standardization, in conscious usage of archaisms or in language planning.

But there is yet another way in which history may influence the shape of the language, its grammar. To see this it might be useful to have another look at Andersen’s famous 1973 model of how grammar change takes place, which we find again and again in text-books on generative diachronic grammar.

(2)
It shows quite clearly that there can be no such thing as diachronic rules, i.e. of the Neogrammrian type OE [a:] > ME [e:] > ModE [ou:], because the grammars are discontinuous. These sort of Neogrammrian rules are in fact purely descriptive historical rules of what changes in the language at output level. No child would have such a rule in its grammar. Andersen’s model makes this difference very explicit.

In another respect, however, this picture of grammar change is rather too simple. Lightfoot (1979: 148) following Andersen writes that “there are no arrows relating the two grammars or the two outputs”, because there “is nothing for such arrows to correspond to” (Andersen indeed calls them “pseudo-connections”). Lightfoot continues, “since there can be no direct causal relation between discontinuous entities, the mere study of correspondences between consecutive stages of a language cannot yield any explanations of change” (p. 148). As I said, I agree with Andersen and Lightfoot as far as the Grammar is concerned, but the Output is a different matter. Grammar 2 (i.e. the grammar of the second generation of speakers so to speak) is not just based on the output of the speaker of Grammar 1, but on the output of many speakers, parents and peers, young and old, from different backgrounds and different dialect areas, and this output in itself embodies a diachronic dimension. So this output is not at all homogeneous; rather, it is highly diffuse, with variation built into it (this variation is itself the result of historical changes, of S-curves not quite completed). In a sense it is therefore impossible to make a neat or clear-cut distinction between Output 1 and Output 2 because this is, unlike Grammar 1 and Grammar 2, continuous. I believe that the ragged edges present in the Output (which, as I said, are the result of ongoing processes of historical change) may change the grammar of the next generation directly, since language learners will try to find a way out of these ‘frayed’ odds and ends by simplifying, analogizing and abducting so that a neater pattern emerges.

An example of how such a change may take place could be the following change taking place in Dutch, a change that has been happening during my own life-time (see also van der Horst and van der Horst 1999: 149–152) involving the use of the polite form $U$ versus informal $jij$ (sg ‘you’). The paradigm given in (3) is with the reflexive verb *zich vergissen*, literally ‘to mistake oneself’, i.e. ‘to be mistaken’, used in the perfect with the perfective auxiliary *hebben* ‘have’:

![Table](image-url)

---

1. It shows quite clearly that there can be no such thing as diachronic rules, i.e. of the Neogrammrian type OE [a:] > ME [e:] > ModE [ou:], because the grammars are discontinuous. These sort of Neogrammrian rules are in fact purely descriptive historical rules of what changes in the language at output level. No child would have such a rule in its grammar. Andersen’s model makes this difference very explicit.

2. In another respect, however, this picture of grammar change is rather too simple. Lightfoot (1979: 148) following Andersen writes that “there are no arrows relating the two grammars or the two outputs”, because there “is nothing for such arrows to correspond to” (Andersen indeed calls them “pseudo-connections”). Lightfoot continues, “since there can be no direct causal relation between discontinuous entities, the mere study of correspondences between consecutive stages of a language cannot yield any explanations of change” (p. 148). As I said, I agree with Andersen and Lightfoot as far as the Grammar is concerned, but the Output is a different matter. Grammar 2 (i.e. the grammar of the second generation of speakers so to speak) is not just based on the output of the speaker of Grammar 1, but on the output of many speakers, parents and peers, young and old, from different backgrounds and different dialect areas, and this output in itself embodies a diachronic dimension. So this output is not at all homogeneous; rather, it is highly diffuse, with variation built into it (this variation is itself the result of historical changes, of S-curves not quite completed). In a sense it is therefore impossible to make a neat or clear-cut distinction between Output 1 and Output 2 because this is, unlike Grammar 1 and Grammar 2, continuous. I believe that the ragged edges present in the Output (which, as I said, are the result of ongoing processes of historical change) may change the grammar of the next generation directly, since language learners will try to find a way out of these ‘frayed’ odds and ends by simplifying, analogizing and abducting so that a neater pattern emerges.

3. An example of how such a change may take place could be the following change taking place in Dutch, a change that has been happening during my own life-time (see also van der Horst and van der Horst 1999: 149–152) involving the use of the polite form $U$ versus informal $jij$ (sg ‘you’). The paradigm given in (3) is with the reflexive verb *zich vergissen*, literally ‘to mistake oneself’, i.e. ‘to be mistaken’, used in the perfect with the perfective auxiliary *hebben* ‘have’:

![Table](image-url)
My parents’ generation had a rule that anyone addressed who was not a peer, was to be addressed with the polite form, \( U \), written with a capital, and in fact a third person form (it is a shortened form of \( \text{Uwe edelheid} \) ‘Your Grace’); that it is third person can be seen from the verbal ending and also from the use of the third person \( \text{zich} \) as a reflexive (so it is similar to German \( \text{Sie} \) and Italian \( \text{Lei} \)). In the course of time, the \( u \)-form (now usually written without the capital) came to be interpreted as a second person, presumably because the addressee is the second person. Consequently, the verbal form changed from \( \text{heeft} \) to \( \text{hebt} \) (the usual form of the second person). This reduces the dissimilarity of the polite and the informal forms and may, in fact, have started a move towards the situation that is found in the third generation, where the \( u \)-form has almost disappeared, and has to be learned by children at a relatively late age. The situation now is that the \( u \)-form is no longer used by children to address their parents or aunts and uncles, and therefore they do not hear it nor is it instilled into them at home.

Next to this internal factor, there is also an external factor that is important here. In the seventies it became fashionable to use the informal \( jij \)-form towards everyone because authority was ‘out’, and equality was ‘in’. Whereas the second generation still possessed the (internal) grammatical rule which their parents’ grammar also had (i.e. use \( u \) whenever you are addressing someone older), they later developed an additional rule, which said, change \( u \) into \( jij \) if you do not want to appear old-fashioned. This too may have led to the confusion of \( \text{hebt} \) and \( \text{heeft} \), i.e. the mix between the third and second person. This second generation, then, was producing more \( jij \)-forms in spite of their internalized grammar rule. This of course meant that the frequency of \( u \) dropped tremendously, so that the third generation would internalize as a new rule: use \( jij \) whenever addressing someone. Only later would they acquire an additional rule, i.e.: use \( u \) in those circumstances where people expect you to defer to them or when you wish to be especially polite.

In other words, there is a mixture of factors at work here, and it depends on the frequency and the variation available which of the forms are actually used, and how the rule gets grammaticalized. One can also see that as long as there is variation, the process may actually be turned around again due to social circumstances, whereby the speaker may fall back on older \( u \)-forms, with even third person verbal forms because they still occur among the variants that he hears. My point then is that, where there is variation in the output (which is the result of a historical development), this may reflect directly on the (in)stability of the rule in the grammar that concerns this variation in output.

A second question we may ask is: Why is language so volatile, why do words change their meaning, why do people speak or write the same language so differently?

Again the above subfields are important, but here also the subfield of communication studies becomes relevant, and that of stylistics: the need to create a personality, to be expressive in language. It leads us to consider other aspects of the study of language change that play a role in language, such as the principles of economy and iconicity, discussed among others by John Haiman (1983); principles that also underlie such processes as grammaticalization and the evolution of a symbolic (i.e. arbitrary) system of grammar (cf. Plank 1979).
The principle of economy is related to our need to communicate fast and efficiently, so often used words get shortened and become grammaticalized. They lose referential meaning in the process. This has happened, for instance, to the infinitive marker to, which started life as a preposition expressing goal or purpose. The result of this grammaticalization may of course be that the original purpose function of to before the infinitive may need to be expressed differently. Use is then made of so-called ‘explorative expressions’ (the term is from Harris and Campbell 1995), which are always floating around. The choice of one such explorative expression, e.g. the phrase in order to or so as to, may then lead to a greater frequency of use and finally to a new cycle of grammaticalization, as for instance happened with Middle English for to and Dutch and German om te, um zu. These explorative expressions themselves may arise via metaphor and metonymy, or via pragmatic inferencing, which shows the importance of pragmatics too for the study of diachronic linguistics. This for instance happened in the grammaticalization of he is going to, which developed from a spatial expression into a temporal one (a well-known metaphorical path) because going to a place in order to do something implies (this is the pragmatic inferencing) that the thing you are going to do lies in the future.

A third question might be: How do we manage to learn our language so quickly?

Here we have links between the theory of grammar and theories of language acquisition and applied linguistics. No doubt other cognitive domains may come in here too, not directly relevant to language.

It seems evident then that in order to get an overall perspective on the phenomenon of language, we must study it from different points of view, from within different disciplines. It is also clear, however, that these disciplines must be connected. To my mind, the study of language change and the history of English provides one very good ‘connector’ between the various subdisciplines, because in order to understand change we must take most of the subfields mentioned (pragmatics, semantics, sociolinguistics, language acquisition etc.) into account. For that reason in itself, the study of the history of English must be theoretically orientated, because loose facts do not provide connections or a deeper understanding. However, I also feel that for this very reason (i.e. the link it provides between the subdisciplines) historical linguists must be eclectic in their choice of theory, because each theory only sheds light on one or two aspects of language, and in an historical study of language many aspects are important in order to understand its development. In language change one cannot, for instance, ignore external factors at the cost of internal ones, or vice versa, because the two are inextricably linked.

2. The history of historical linguistics

Since I have just ‘proclaimed’ that a historical viewpoint is helpful in gaining understanding, it seems to me a good idea also to look very briefly at the history of our own subject, at the way our subject has been studied and taught, to see what we can learn from that. We need to see in how far the different approaches here contribute to
a better understanding of language on the one hand, and to a more coherent university linguistic programme, cementing the various subfields together, on the other.

There are, and have been, various ways of looking at historical grammar or the history of English. There is first of all the philological approach, in which the study of the grammar of any particular period in the history of English is subservient to a proper reading and understanding of the texts of that period. A good and still fairly recent example of such a grammar is Bruce Mitchell’s *Old English Syntax* (1985). The great advantage of such grammars is that they contain a lot of material, are descriptively very accurate, and that they are indeed a great help when reading, as in this case, Old English texts. The approach also provides a necessary link with another part of our curriculum, namely the study of literature. The purely philological approach gives us not only a tool with which to read older texts, but it also makes us more sharply aware of the differences between Old and Modern English or Middle and Modern English. It makes us aware of language variation, and of the relativity of norms (this is linked to the fifth point in the list of [1], ‘alterity’). The disadvantage of the philological approach is that the material is ordered almost purely taxonomically; there is no attempt at a proper understanding of why Old English grammar looks the way it is, to understand it as a coherent system, and there is no real interest in any changes that may (have) take(n) place, and why. In other words, there is hardly a link here with the study of language proper (and therefore with any of the other subdisciplines mentioned).

When we look at the treatment of word order patterns in Mitchell, for instance, we are given a full description of the patterns that occur, but we are not given any understanding of them, e.g. why main clauses may show inversion and subordinate clauses do not, why an object pronoun often occurs in SOV structures, and why SOV structures are more current in e.g. relative clauses than in main clauses. These facts can be more fruitfully approached and understood with a proper grammatical theory of clitics or with discourse theory (cf. Koopman 1987). Similarly, Mitchell (1985: §§160–166) notes that adjectives may occur both before and after the noun in Old English, but he has no idea what determines this. I looked at this problem in some detail in the past two years (cf. Fischer 2000) and found that a lot of it is determined by iconic factors, as described for instance by Bolinger (1972) thus,

\[(4) \quad (a) \quad \begin{array}{c} A \\ \hline \end{array} \quad N \quad (b) \quad \begin{array}{c} N \\ \hline \end{array} \quad A\]

This picture is meant to indicate that an adjective placed before the noun changes the nature of the noun, it becomes a different class of noun, while an adjective placed after the noun only gives extra information about the noun but is not an inherent part of it. Thus, in the first Old English example in (5),
Teaching the history of the English language

(5) (a) ðone ilcan ceaddan iungne (L53 (Chad) 1.184)
the same Chad young

(b) *adolescens* geong mann (ÆGI 1 301.9)

we see that ‘young’ refers to the man named Chad when he was a young boy. When an Old English writer wished to talk about a young man or youngster, he would put the adjective before the noun, as we can see from the way *adolescens* is glossed by Ælfric in (5b). An interesting minimal pair is the following,

(6) (a) and þær sint swiðe micle meras fersce (Mitchell 1985: §126, Oros. 19.5)
and there are very big (many?) lakes fresh

(b) Swetum ferscum waterum (Læceboc2, 1.16.1.4)
with-sweet fresh water

In (6a) we have a postponed adjective because the writer refers to a lake that has fresh water in it; the lake itself is not qualified as fresh, in (6b), however, the water itself is fresh, it is sweet water as contrasted to a different kind of water, i.e. salt water, and the adjective now precedes (for a more detailed discussion concerning the position of adjectives in Old English see Fischer 2000).

Another traditional way of looking at grammar is the method found in works such as Jespersen’s *A Modern English Grammar: On Historical Principles*. Here it is the grammar of the contemporary period that stands central, but the way the language is structured is explained from a historical point of view. The ‘explanation’ that is lacking in the philological grammar mentioned above, is important here. But in this case we do not get a very full description of a grammar-phase in a particular period. The emphasis is also rather on language as a historical object, language as ‘output’ only. What is largely ignored is the role of the speakers, and of how they acquire their language. There is a tendency here to see language change in terms of ‘diachronic rules’ (see the discussion of the model in (2) above), which are merely descriptive and do not explain how or why the speaker changes his/her grammar, and how changes indeed may be connected to other changes taking place at the same time. It does, for instance not connect the fact that all the following structures that first appear in Middle English,

(7) (a) Moyses forbade swyn to etanne (AELS (Maccabees) 85)
Mozes forbade pigs to eat
‘Mozes forbade pigs to be eaten.’
(the passive infinitive here is a Middle English innovation)

(b) it is grotesque for *outdoor sports to be so* closely associated with habits that are notoriously dangerous to health (*Daily Mail* 9-3-77)

(c) ‘Come, come’, cried Emma, *feeling this to be* an unsafe subject

\[ ^2 \text{I have used the reference system of the DOE.} \]

\[ ^3 \text{Examples (7b, c, d) are taken from Duffley (1992).} \]
(d) they knew it to be a cloud. Yet they could see animals in it

(LOB corpus Fol 1371)

are all finally connected with the fact that the basic word order of English changed from SOV to SVO, so that an NP before an infinitive came to be looked upon as the subject of that infinitive rather than the object, which in turn led to a break in structure in that that subject was more part of the infinitival clause than it was as an object of the matrix clause. This break in structure also led to the influx of Latin-type accusativus cum infinitivo-constructions. There is no time to go into details here, but I refer the interested reader to Fischer et al. (2000: Ch. 7).

Since the work of de Saussure, it has become traditional to separate synchronic from diachronic grammar, and to see synchronic grammar as primary (I am here referring also to the sixth point on the list in [1]). This has led to a different way of studying grammar, which has had great influence on our understanding of how grammar works, and has led to a deep interest in the theory of language acquisition, and mental models of grammar. It has proven very fruitful and successful, and has laid bare many connections between aspects of grammar that had not been seen before. It has uncovered important principles and universals, such as the importance of basic word order, for instance, and the behaviour of clitics, both briefly mentioned above. It has also deepened our understanding of the internal mechanisms of language change.

The advantage of working with a theoretical linguistic model is that it creates the sense of a goal, we move beyond mere description, beyond texts, to a deeper understanding of how the human mind works. However, the Saussurean dichotomy has also led to a loss of interest in the theory of language change, or to put it differently, in an understanding of how language, as an object used by speakers against a historical socio-cultural background, changes. In fact, historical linguistics as a discipline now became subservient to the (synchronic) theory of grammar. This is particularly clear in the work of Lightfoot (1979, 1991). Lightfoot emphasizes that it is through knowledge of what is possible in change that we may discover what is possible in grammar, i.e. change may throw light on the contours of the theory of grammar. Although this approach did indeed at first mean a new lease of life for historical linguistics (which had become a rather dusty subject in the ‘structuralist’ period, and the word ‘philology’ itself had almost become a term of abuse), it in turn has led to distortions. Within generative theory, for instance, the emphasis is very much on the theory model, on speakers’ intuitions about their competence rather than on ‘hard’ data; thus, language-internal factors as an explanation for change are promoted at the cost of external ones, which play at least as important a role (see for instance the articles collected in Gerritsen and Stein 1992). So in fact what has happened is that only a narrow area of language change was put under scrutiny, which in turn has affected the ‘window’ (to use Kiparsky’s [1968: 174] words) it may provide on the form of linguistic competence or the theory of grammar. This window gives us only a partial view of what is possible in change, and it may therefore also distort the light it throws on linguistic competence. In addition, because of the emphasis on model building and on the elegance and simplicity of the model, the theory has a tendency to become more and more abstract. Changes are often described in terms of changes in rules, conditions or functional categories, that have no
surface manifestation. The more abstract the model, the more difficult it is to apply it with real explanatory power to its object of study, the changing language. A quite illustrative example of this tendency is the interpretation of the infinitive marker *to* in Old English by Kageyama (1992). In an attempt to explain the connectedness of a number of constructions that are used in Old English and which disappear or are changed in Middle English, he must give a highly abstract analysis of the particle *to*, which in his view is placed in a separate functional node, i.e. AGR. *To* is thus not seen as part of the infinitive or as a preposition governing the infinitive, but it is positioned in a higher node, above the infinitive. In this node, it functions as an agreement marker and also as an external argument, i.e. a subject, but one without a theta-role (Kageyama claims it absorbs its theta-role). So in order to prove the connectedness between various types of constructions, Kageyama must assume that *to* functions like an inflection on a verb and also as the subject of that verb, characteristics that intuitively we would not associate with *to* at all. There is no surface evidence for the change in *to* – this occurs later – there is only the postulated connectedness of the above-mentioned constructions. Circularity looms large when the explanation becomes almost purely theory-internal (for a discussion of the problems posed by Kageyama’s hypothesis, see Fischer 1996).

It seems to me that it is essential that historical linguistics remains an independent discipline, and that it serves the theory of grammar but does not become part of it (as Lightfoot advocates). In historical linguistics we must look at and take into account what is visible on the surface, not in the least because we do not have access to native speakers’ competence. In this light, it is important to use large data bases when describing and trying to explain language change and the historical development of the English language. Another reason to concentrate on what is visible is that it seems safe to assume that change usually starts out from the surface structure, not from deep, abstract structures. After all, change is brought about by speakers, who are not perfect linguists, as Joseph (1992) has emphasized. When speakers change their language, they are not necessarily steered by the elegance or simplicity of the overall grammatical model. It must be taken into account that factors such as analogy, abduction, metaphor and metonymy, pragmatic inferencing, discourse strategies, politeness strategies, folk etymology, surface adjacency and perceptual and iconic factors play an important role in language change, whereas these same factors hardly seem to come to the fore in formal theories of grammar.4

---

4 There is an interesting comment on this problem of how data should be classified or analysed in connection with the role of the observer or speaker, in McGinn (2000: 64),

[...] we classify things in very different ways. Sciences like chemistry classify according to the laws and internal structure of things; common sense tends to go by surface sensory appearance. Artifacts like chairs and bottles are classified by function; games are grouped together by Wittgensteinian ‘family resemblance’; numbers possess abstract principles of similarity. There is no one notion of likeness here, so our concepts are accordingly various in what they require of objects. Traditional empiricists neglected this point to their detriment.

It seems to me that this difference is the same for linguists and speakers; linguists, like chemists, will tend to analyse the data more deeply or abstractly according to their theoretical model (its ‘laws’), while speakers will analyse in a common-sensical, surface way, and thus use different notions of likeness depending on specific characteristics of the data in question.
3. Three illustrative cases of the use (non-use) of a theory

I would now like to look at a few examples of cases of change in English, where, on the one hand, the theory has played an important role in the explanation given by linguists: e.g. Lightfoot’s (1981) explanation of the loss of the so-called impersonal construction in Middle English in terms of a violation of a principle of the theory of grammar, or Brinton’s (1991) explanation of the rise of the semi-modal have to, which is explained in terms of grammaticalization theory. Or where, on the other hand, no explicit theory was used, as in the study of the English impersonals in Denison (1990). In all three cases I want to concentrate on the depth of explanation that each contribution offers, or, to put it differently, on the insight that each case ultimately provides with respect to the nature of language as a whole. I have chosen relatively simple and simplified examples because I have little time to go into them deeply, the examples are not recent but as an illustration they serve my purpose.

Lightfoot (1981) argues that the Old English impersonal constructions were characterized by the fact that the subject takes final position. His example is a made-up one taken from Jespersen:

(8) (a) tham cynge licodon peran

Basic word order in Old English is SOV, but in these impersonals the subject has been postposed according to Lightfoot into final position, as the arrow and the trace left in initial position indicate. In Middle English, due to morphological attrition, the same sentence has come to look as follows,

(8) (b) The king liked(en) pears.

Because the Middle English basic word order by now has changed into SVO (due to other causes), it becomes impossible to interpret this sentence as an OVS structure – i.e. as a structure with a final subject – because that would mean that the SVO speaker that produces this must move not only the subject into postverbal position, but also the object (which is postverbal in an SVO language) into preverbal position, and this would constitute a violation of the so-called Trace Erasure Principle (TEP), which stipulates that traces left by movement can only be erased by a designated morpheme (such as expletive it or there) and not by a random NP (this is made clear in 9b).

Thus the TEP, as it were, forces Middle English speakers, when confronted with the old form (9a), to abduct a surface SVO structure so that king becomes the subject, base-generated in subject position and pears becomes the object. The abduction also forces a semantic change on the verb, from causative ‘please’ to receptive ‘like’. In purely syntactic terms, this may provide a very neat explanation, but in terms of actual data, there are many problems. To mention but a few, the Old English impersonal verb

\footnote{For a more detailed discussion of the points at issue here, see Fischer and van der Leek 1981, 1983 and 1987.}
(9) (a) situation in OE (SOV)

(c) situation in ME (SVO)

constructions are much more varied than suggested by the made-up example in (8a); in fact, they may occur both with and without a subject, and when there is a subject it is not necessarily postposed:

(10) (a) [... ] bæt hi [acc] bæs metes [gen] ne recð (Met A6 13.44)
that them from-the food not care-is
‘that they do not care about the food’

(b) æghwylc man [nom] [...] þurh gode dæda Gode [dat] lician sceal
(HomS46(B1H1om 11)255)
each man through good deeds God please must
‘each man should please God through good deeds’

(10a) is an example where there is no nominative subject, and (10b) shows that the subject of the impersonal verb need not be postposed.

Secondly, in Middle English there are writers who use the verb *like* both with the meaning of ‘please’ as well as ‘like’ whereas only the latter should be possible after the change (cf. Fischer and van der Leek 1981). So there is a problem in that the data do not conform to the theory, and that the TEP explanation cannot be used for examples like (10b). An additional problem is that the TEP is a purely theory-internal projection. If the principle is later found to be unnecessary or wrong, the explanation falls away too, and we must start again. Of course, if the theory is highly developed, this need not be a problem. But as long as many principles of grammar are theory-internal and have not been given a surer footing in another domain (cf. Bybee 1988 and Fischer 1997 for a discussion of the importance of an explanation making use of a different domain), these problems will continue to occur.

Let us next look at the other side of the coin. Denison (1990 and 1993) considers all the possible types of impersonal constructions (against Lightfoot dealing only with one). There is quite a large variety of structures with what are called ‘impersonal verbs’ (as we see from 8 and 10). Semantically (except for ‘weather’ verbs like *to rain* etc) they are all of a type where the experiencer is unvolitionally or uncontrollably involved in the activity expressed by the verb. By using the semantic and syntactic criteria which Denison has established on the basis of true impersonal verbs, he comes to include further verbs and constructions which show partly similar behaviour, such as for instance *derian* ‘harm’ and *faegnian* ‘rejoice’:

Let us next look at the other side of the coin. Denison (1990 and 1993) considers all the possible types of impersonal constructions (against Lightfoot dealing only with one). There is quite a large variety of structures with what are called ‘impersonal verbs’ (as we see from 8 and 10). Semantically (except for ‘weather’ verbs like *to rain* etc) they are all of a type where the experiencer is unvolitionally or uncontrollably involved in the activity expressed by the verb. By using the semantic and syntactic criteria which Denison has established on the basis of true impersonal verbs, he comes to include further verbs and constructions which show partly similar behaviour, such as for instance *derian* ‘harm’ and *faegnian* ‘rejoice’:
These two verbs, *derian* and *faegnian*, share elements of meaning with *eglian* ‘ail’ and *sceamian* ‘shame’, which are truly impersonal (i.e. they have a dative experiencer, a third person verbal form and appear in a construction lacking a nominative subject). Syntactically, *derian* and *faegnian* have a non-accusative object-experiencer (11a) or a non-accusative object-cause, and when they are coordinated with impersonal verbs they may share the same object (cf. 11c). However, unlike true impersonals they do not appear in subjectless constructions. Denison (1990, 1993: 95–96) then invokes a notion of ‘gradience’ or ‘serial relationship’ to explain how by lexical diffusion the true impersonals slowly adopt the characteristics of the pseudo-ones like *derian* and *faegnian*. Although in this case we do get a very accurate description of the data, we miss a leading principle. We do not begin to understand why the structures are changing in this way; why, for instance, the change was not the other way around. There is no attempt here to uncover the underlying system or motive that may explain the direction of the change. Although Denison had widened the scope of his data considerably, he does not really take enough notice of the overall shape of the grammar at that point in time, which also may explain the direction of the developments.

My final example concerns the history of the semi-modal *have to*, which, according to linguists working within grammaticalization theory, developed gradually and slowly from a possessive verb, as in Old English,

(12) hæfst ðu æceras to erigenne? (ÆGram 135.2)

have you acres to plough?

into a semi-modal auxiliary. In (12) the infinitive *to erigenne* still functions as an adjunct to the noun (a kind of restrictive relative clause), and the noun *æceras* is the object of *have*. In (13), however, the infinitive has become a complement of the verb or rather part of the VP, and the noun now depends on the infinitive rather than on the finite verb,

(13) Have you (do you have) to plough acres?

Once this construction has developed, other constructions without any object become possible too, as in

(14) You have to go now.

Since there is a strong belief among linguists working within grammaticalization theory, that the process is semantically driven, and that it is gradual and unidirectional, most attention has been paid to the semantic change in this construction, to the slow semantic bleaching of *have* from a lexical verb meaning ‘possession’ into a modal verb expressing
obligation. It is not taken into account that in Old English, *habban* already occurred frequently in bleached form, and that constructions like the one in (12) already often possessed a modal colouring, but not necessarily one of obligation. It could also express possibility or desire as in the examples of (15):

(15) (a) *he Segeates selran [acc] næban to geceosenne cyning ænigne [acc]*
(Beo 1845)

the Sea Geats better not-have to choose king any
‘the Sea Geats do not have a better man to choose as king’ /
‘[...] whom they can choose as king’

(b) *Fela ic hæbbe eow to secganne (CP 35.237.11)*

Much I have to-you to say
‘Much I would like to say to you’

The usual story then is (as told by linguists like van der Gaaf 1931 and Brinton 1991) that the semantics of the construction slowly changes, and that the syntax changes in the wake of it, slowly too, step by step. I have looked at these constructions as they appear over a long period of time in the Helsinki corpus (cf. Fischer 1994) and have found that there are in fact no changes between Old English and the early Modern period, when the first unambiguous examples of semi-modal *have to* occur (unambiguous, because in these examples *have to* is followed by an intransitive verb, as in [14]). Thus, one cannot really maintain that the driving force behind the change is the process of grammaticalization and semantic bleaching, some other motivation must be found. I have suggested that it may be the basic word order change from SOV to SVO, which takes place in the Middle English period (and which is visible on the surface!) may be the real reason for the change. Through it the order in these clauses changed. In Old English the order was

(16) Old English: SOV

\[
\begin{array}{c}
S \\
\downarrow \\
N \\
\downarrow \\
I \\
\downarrow \\
a \text{story to tell} \\
\downarrow \\
\uparrow \\
\text{have} \\
\end{array}
\]

(heavy NP shift/VP raising in subordinate clauses)

(Verb-second rule in main clauses)

usual surface order in Old English in both main and subordinate clauses of *I have a story to tell*.

In main clauses due to V2 this regularly produced the surface order *I have a story to tell*, and in subordinate clauses this is a common surface order too (possibly due to heavy NP shift). The result of this surface order in Old English was that the NP object of *have,*
i.e. story, could also be interpreted as an object of tell. In infinitival constructions in Old English the infinitival object would precede the infinitive, and thematically there is a stronger link between story and tell than between story and have, because have could already be bleached of its possessive meaning, and therefore was not a verb that would strongly assign thematic roles. This ambiguous situation remained, but became problematic once the basic word order became SVO. If the object story was indeed felt to have a closer semantic relation with the infinitive than with the bleached verb have, then it is natural that this is shown syntactically too. In Old English, this created no difficulty, because the object in its intermediate position between the matrix verb and the infinitive could belong syntactically to both, but in Middle English, which had developed SVO word order, the syntactic object position was behind the infinitive, so we see a tendency emerging now to place the object there, as is shown in (17).

(17) I have to tell a story.

This put the to-infinitive adjacent to have. It is this adjacency which probably led to the reinterpretation of the verbal group have + to + infinitive as a unity. It is interesting in this respect to note that German and Dutch, which did not undergo the word order change to SVO, did not share the development of English have to, Dutch hebben and German haben followed by a te or zu-infinitive still have the same meaning as the construction had in Old English; they do not occur with an intransitive infinitive there, nor with double ‘have’ (I have to have ...), nor with an epistemic meaning (It has to be true that...).

To me this example shows that one may get carried away by the theory, in this case grammaticalization theory, and thus miss some of the facts. On the other hand, it is important to note that I myself really only came to think of the case in this way, because Brinton had strictly applied the theory to it. The good thing about a theory also is that it may be falsified more easily because it makes strict predictions, and hence through falsification, we may be getting further on the road towards explanation. It must be clear then from the above three cases that we need both sufficient data and a theoretical approach in order to advance our knowledge about language change and about language.

4. Concluding remarks

What is true for historical language research is also true for teaching in this area. We cannot properly teach the subject without a theoretical base. In section 1 of this paper, we had already concluded that historical linguistics can play a pivotal role in language teaching only if it is theoretically inclined. By combining forces with other theoretical approaches to language such as provided by pragmatic theory, sociolinguistics, discourse theory etc., we can begin to understand what factors play a role in language change, and thus elucidate the working of language itself. In teaching the history of English, therefore, I would plead for a historical linguistic approach that combines theory and description (following Fillmore 1990, who pleads for
a combination of the ‘armchair’ and the ‘corpus linguist’), so in fact for a combination of the good qualities of the approaches described above in section 2 (Mitchell, Jespersen and de Saussure). A historical linguist may be eclectic in his/her use of theory, because the aim of historical linguistics is not in the first place to (help) build a theory but to explain the ways in which language changes, in which both language and speaker play a role, and in which it is not possible to make a neat split between synchrony and diachrony. This approach may ultimately help us too to form a clearer idea of the shape of the theory of grammar (cf. e.g. Fischer 1997).

When one now considers the place of historical linguistics in the university curriculum at large, it should be clear, I hope, that that subject is worthwhile because through its emphasis on the usefulness of theory, its need for several theories, and also through its intimate connection with the nature of language itself, it may act like a kind of glue between the various subdisciplines. It provides a tool for understanding the literature of the period because it gives a description of its grammar, and it links up with the other linguistic topics because it can put into practice ideas acquired elsewhere, e.g. in discourse theory, generative theory, sociolinguistics etc.

References


