Grammaticalisation: Unidirectional, non-reversable?
The case of to before the infinitive in English

Olga Fischer
University of Amsterdam

1. Introduction

In the literature on grammaticalisation it is quite generally assumed that this process is unidirectional and non-reversable, and also that it is essentially a process driven by semantic or pragmatic factors with grammatical and phonetic changes as it were automatically following (cf. Brinton 1988; Lehmann 1991; Traugott and Heine 1991; Hopper and Traugott 1993; Bybee, Perkins and Pagliuca 1994). I have tried to show on two other occasions (cf. Fischer 1994b, 1997b) that the syntactic and semantic changes do not necessarily go in tandem, and that the assumption of grammaticalisation as a diachronic process clashes with models of change that accept that change takes place within the grammar of each speaker, i.e. that each speaker builds up his grammar afresh and does not take account of processes that have started long before he was there. I agree with Harris and Campbell (1995) that it is probably methodologically more accurate to consider grammaticalisation not as a separate mechanism or cause of change, but as a process built up out of a series of reanalyses and extensions.1 A process that may proceed all the way, from lexical item to affix or even zero, but as often as not may stop halfway, or may even regress to some extent. Each station in the series or the chain is dependent on synchronic circumstances and determines the direction of the change afresh. In another paper (Fischer 1999), I also tried to show that many of the processes at work during grammaticalisation are motivated by iconicity, such as analogy, metaphor, isomorphism and persistence. I do not wish to go very deeply into the iconic aspects of grammaticalisation here (I refer the interested reader to Fischer 1999), but I would like to look in
detail at what happened in one particular case of grammaticalisation, and determine the nature of the process, and what may have caused it, in order to get a clearer picture of what grammaticalisation is. But first, before turning to this case study, I want to briefly show what I understand by iconicity, since this plays an important background role in the discussion that follows.

As I mentioned above, I believe that iconicity plays an important role in grammaticalisation processes, and may even upset some of its presumed unidirectionality. Since iconicity is used by linguists in different ways, a description of what I understand by iconicity is in order. Iconicity as a semiotic notion refers to a natural resemblance or analogy between the form of a sign (‘the signifier’, be it a letter or sound, a word, a structure of words, or even the absence of a sign) and the object or concept (‘the signified’) it refers to in the world or rather in our perception of the world. The similarity between sign and object may be due to common features inherent in both: by direct inspection of the iconic sign we may glean true information about its object. In this case we could speak of ‘imagic’ iconicity (as in a portrait or in onomatopoeia, e.g. ‘cuckoo’) and the sign is called an ‘iconic image’. When we have a plurality of signs, the analogy may be more abstract: we then have to do with ‘diagrammatic’ iconicity, which is based on a relationship between signs that mirrors a similar relation between objects/concepts or actions (e.g. a temporal sequence of actions is reflected in the sequence of the three verbs in Caesar’s famous dictum *veni, vidi, vici*: in this instance, the sign — here the syntactic structure of three verbs — is an ‘iconic diagram’). Obviously, it is primarily diagrammatic iconicity that is of great relevance to language. Both imagic and diagrammatic iconicity are not clean-cut categories but form a continuum, on which the iconic instances run from almost perfect mirroring (i.e. a semiotic relationship that is virtually independent of any individual language) to a relationship that becomes more and more suggestive and also more and more language-dependent (for a convenient overview of various types of iconicity see Fischer and Nanny 1999). Isomorphism, a term also used in the discussion that follows, is an example of diagrammatic iconicity of a rather abstract kind. It is used here in the sense used by Haiman (1980: 515–516), i.e. the principle of one form (signifier) corresponding to one meaning (signified). Iconicity also plays a fundamental role in the so-called naturalness approach to language theory (e.g. in the work of Wolfgang Dressler, Willi Mayerthaler, David Stampe, Wolfgang Wurzel and others), and in cognitive approaches to language (as in the work of Talmy Givón, Ronald Langacker, Paul Hopper, Sandra Thompson and others). Basic to both these approaches is (in the words of Dressler 1995: 22) that “it does not assume an autonomous module of grammar, but attempts to find cognitive and other extralinguistic bases ... for
grammatical principles and preferences”, that “questions of the relationship between language and the mind can be approached only by considering language in its natural functional context” (Hopper and Thompson 1984: 747–748), and that the cognitive strategies underlying language systems have a strong perceptual basis.

After this brief excursion, I now wish to consider what exactly the nature of grammaticalisation is. One of the problems I have with the way grammaticalisation has been dealt with in the literature is that the mechanistic side of it has been overemphasised, with the result, I think, that the mechanism has become too powerful as an explanatory tool or as a description of a diachronic process of linguistic change. Thus, the following quote from Bybee et al. (1994: 298) suggests that grammaticalisation is seen as an independent process with independent explanatory value: “Thus our view of grammaticization is much more mechanistic than functional; the relation between grammar and function is indirect and mediated by diachronic process. The processes that lead to grammaticization occur in language use for their own sake; it just happens that their cumulative effect is the development of grammar” (italics added). Similarly, Vincent (1995: 434) talks about “the power of grammaticalisation as an agent of change” (italics added), which suggests that grammaticalisation has explanatory value, that it has independent force. Finally, Heine et al. (1991b: 9) write that “Meillet followed Bopp rather than Humboldt in using grammaticalization as an explanatory parameter in historical linguistics” (italics added), and the authors themselves seem to follow this line too (see 1991b: 11).

Let us first look at the way the process has been described in the literature. Grammaticalisation is generally seen as a gradual diachronic process which is characterised as unidirectional, i.e. it always shows the “evolution of substance from the more specific to the more general and abstract” (Bybee et al. 1994: 13). The unidirectionality applies on all levels, the semantic, the syntactic and the phonological. Almost without exception, the process is seen as semantically driven, with bleaching of meaning playing a primary role. (This is not true for all linguists, notably Hopper and Traugott (1993) believe that bleaching only plays a role in the later stages of grammaticalisation.) Rubba (1994: 81), for instance, describes it as primarily a process of semantic change. Bybee et al. (1994: 17–18) even suggest that we can reconstruct the path of grammaticalisation with the help of the “hypothesis that semantic change is predictable”. The notion of graduality implies that grammaticalisation is seen as “an evolitional continuum. Any attempt at segmenting it into discrete units must remain arbitrary to some extent” (Heine and Reh 1984: 15, and see also Heine et al. (1991b: 68, 165 and passim). In this light it is not surprising to read that the mechanisms at work in and the causes of grammaticalisation are also seen as basically semantic/pragmatic in
nature. For most linguists writing on grammaticalisation, the main mechanisms involved are metaphoric and metonymic in nature. Metaphoric change can be related to analogy, it is a type of paradigmatic change whereby a word-sign used for a concrete object (i.e. the word *back* as part of the body) can be reinterpreted on a more abstract level as an indication of 'location' (because of some element that these concepts have in common), and then further interpreted along the metaphorical axis as an indication of 'time'. Metonymic change can be related to re-analysis, and functions on the syntagmatic plane. It takes place mainly via the "semanticization of conversational inferences" (Hopper and Traugott 1993: 84). A good example of metonymic change is the case of *to be going to* discussed in the introduction (this volume). Another example discussed by Hopper and Traugott (1993: 85) is the change in OE *hwile* 'while'; the meaning of temporal simultaneity that it has, may change into a causal meaning because in many cases "the conditions specified in the subordinate clause [i.e. the clause introduced by *while*] serve not only as the temporal frame of reference for those in the main clause, but also as the grounds for the situation".

As far as the cause of grammaticalisation is concerned, this is usually seen as being pragmatic in nature. Bybee *et al.* (1994: 300) write: "the push for grammaticization ... originates ... in the tendency to infer as much as possible from the input, and in the necessity of interpreting items in context". They show that grammaticalisation occurs in cycles or is self-propelling. The process of grammaticalisation (loss of concrete form) itself leads to a search for new expressive means to indicate the same function, and when the new expression has again grammaticalised, the search for a new concrete expression begins again. Likewise Hopper and Traugott (1993: 86) concur with Heine *et al.* (1991a: 150–151) that "grammaticalization can be interpreted as the result of a process which has problem-solving as its main goal". It is the result of a "search for ways to regulate communication and negotiate speaker-hearer interaction" (1993: 86).

Although I would agree with the views just now discussed, i.e. that re-analysis and analogy, or metonymic and metaphorical processes, are important in language change, and also that grammaticalisation may be caused by the need for expressivity and routinisation, I still cannot see that there is room for a separate or 'independent' process of grammaticalisation. Where most linguists see a unidirectional process from concrete to abstract, a process that cannot be cut up into segments, I can only see a more or less accidental concurrence. The processes underlying grammaticalisation may lead one way as well as another, i.e. there is no necessary link between one segment of the chain of grammaticalisation and another.
I also think that grammaticalisation processes themselves can only be discovered with hindsight, which means that if we have a preconceived notion of what grammaticalisation is, we will indeed discover mainly those processes that have run a full or ‘fullish’ course, and we will not realise that there may be many cases where the path of grammaticalisation proceeded differently. So it may only seem that grammaticalisation usually follows the same channel. Aborted and reversed processes are very difficult to find when one looks backwards in this way.4 So the similarities in known cases of grammaticalisation may have led to an overemphasis on a common core, and through that the idea may have arisen that grammaticalisation is an explanatory parameter in itself. To my mind it is the subprocesses that explain the change. I agree with linguists such as Lightfoot (1979, 1991, 1999) and Joseph (1992) that, logically, diachronic processes cannot exist because diachronic grammars do not exist. Each speaker makes up his own grammar afresh on the basis of data surrounding him, and on the basis of his general cognitive abilities or strategies (or, so one wishes, on the basis of some innate Language Acquisition Device). So why should a grammaticalisation process necessarily run from a to b, to c etc.? Why should there be unidirectionality?5 With Harris and Campbell (1995:20, 336ff.) (and see also Fischer 1997b) I would tend to accept that grammaticalisation has no independent status, no explanatory value in itself. It was when I looked at the so-called grammaticalisation of have to in English from a possessive verb to a modal auxiliary (see Fischer 1994b) that I began to realise that not all grammaticalisation change is driven semantically, and that unexpected, language specific factors may play a role. I think it pays to look at any hypothetical grammaticalisation process in detail, next to taking the wider, typological bird’s eye view. Here, therefore, I would like to consider the case of infinitival to in English.

2. The case of infinitival to

There is a widespread belief that the development of the original preposition to before the infinitive into a meaningless infinitival marker follows a well-known grammaticalisation channel. This is clear from Haspelmath’s (1989) study, the essence of which is expressed in his title, “From purposive to infinitive — a universal path of grammaticization”. Haspelmath shows that in many languages in the world the allative preposition — to in English — which expresses location, or rather the goal of motion, also comes to express goal or purpose more abstractly; and that in combination with the infinitive, the preposition begins to lose its original purposive function, ending up as a purely grammatical
element to indicate that the verbal form is an infinitive. This interpretation of the development is already seen in Jespersen (and cf. also Mustanoja 1960: 514),

In ... the to-infinitive, *to* had at first its ordinary prepositional meaning of direction, as still in "he goes to fetch it"; ... But gradually an enormous extension of the application of this to-infinitive has taken place: the meaning of the preposition has been weakened and in some cases totally extinguished, so that now the to-infinitive must be considered the normal English infinitive, the naked infinitive being reserved for comparatively few employments, which are the solitary survivals of the old use of the infinitive. This development is not confined to English: we find it more or less in all the Gothonic languages, though with this preposition only in the West Gothonic branch (G. *zu*, Dutch *te*), while Gothic has *du*, and Scandinavian *at* ( ... ) (Jespersen 1927: lO-U)

It seems to me that the expectations raised by the fact that this seems to be a frequent grammaticalisation pattern, has led us too much to see the English case as following the well-trod den path. I think it pays to look more closely at the linguistic details. I have compared the development of the infinitive marker in Dutch and English6 and come to the conclusion that *to* and cognate *te* have not grammaticalised in the same way. On the contrary, it looks as if *to* has been stopped early in its development and has even regressed in some respect. I think this could be characterised as a process of what Frans Plank (1979) has called *Ikonisierung*, a moving away from the symbolic pole back to the iconic one. I will first briefly explain what I mean by these two poles. It is well-known that in language there is competition between iconic and economic motivation (cf. Haiman 1983) or between the need for clarity and the need for processing speed. General erosion leads to the toss of expressivity and consequently to a constant need for new linguistic expressions. One could say therefore that language moves or is situated along an axis with two poles: an iconic, concrete pole at one end, and a symbolic (or perhaps 'arbitrary' or 'conventional' is a less confusing term here), abstract one at the other. In grammaticalisation, elements move along this axis, from concrete to abstract. One could also refer to the iconic pole as original and creative and to the symbolic as derivative and mechanistic.

I believe indeed that the forces behind the development of *to* have been to a large extent iconic (with 'persistence' and analogy or isomorphism playing an important role, see for more detail Fischer 1999), although there were some syntactic factors too, which I will come back to below. The main iconic factor indeed is isomorphism. I am using the term here, as used by Haiman (1980), meaning the existence of a one-to-one relation between signans and signatum, similar to von Humboldt's principle of 'one form-one meaning'. One can see that
through the grammaticalisation of *to*, the original isomorphic or one-to-one relation between the signans and the signatum (as given in (1a)),

(1) Structural stages in the grammaticalisation of *to*

a. \(\frac{\alpha}{x}\)  
b. \(\frac{\alpha}{xy}\)  
c. \(\frac{\alpha}{x} \frac{\beta}{y}\)

(\(\alpha = \) the signans *to*; \(\beta = \) the reduced signans of *to*; \(x = \) signatum ‘goal’; \(y = \) signatum ‘infinitival marker’)

is disturbed (as shown in (1b)). Through grammaticalisation, the sign *to* acquires two signata: the first is the original prepositional purposive ‘to,’ and the second the semantically empty, infinitive marking element ‘to’. The result is then an asymmetric, non-isomorphic situation as shown in (1b). This lack of isomorphism can be amended in two ways. The usual way according to the grammaticalisation hypothesis is for the new signatum to acquire its own distinctive linguistic form. This may be obtained through the phonetic reduction of *to* (this is another iconic principle, the ‘quantity principle’, see Givón 1995:49), which would then coexist with the full form *to* (stage (1c)). This development is most clear in Dutch, which has *infinitival te*, next to the earlier particle *toe*. But in Middle English, too, we find occasional *te* spellings or other spellings indicating the phonetic reduction of *to*. So with stage (1c) we have a new stable isomorphic relation. The other solution for the asymmetry of (1b) is to go back to the earlier symmetry (i.e. 1a). This also makes the relation isomorphic again, and it is more strongly iconic than (1c) because here the sign *to* is linked back up with its original meaning, i.e. it is re-iconicised, going back to the iconic pole (in Fischer 1999, I argue that this process, usually called persistence, is also iconic in nature). So my suggestion is that diacronically English *to* moved back to stage (1a), while Dutch *te* moved on to stage (1c). In what follows, I will have a look at the (comparative) facts, and also offer some suggestions as to why English *to* re-iconicised, both in terms of isomorphism and persistence.

2.1 The grammaticalisation of *to* in its early stages

It seems that at first, in the late Old English, early Middle English period, *to* developed very much like Dutch *te*. Evidence for this can be found in the following facts:

(2) a. strengthening of *to* by *for*  
b. phonetic reduction of *to*  
c. loss of semantic integrity  
d. occurrence of *to*-infinitive after prepositions other than *for*
2.1.1 To strengthened by for

First we find the need for an additional preposition (for) to emphasise the goal function of the to-infinitive. This use of for is attested from 1066 onwards (see Mustanoja 1960: 514) and steadily increases in the Middle English period until 1500 (see Table 1). A similar development can be seen in Middle Dutch, where om(me) te begins to occur quite frequently (see Stoett 1909: §283) and becomes more and more regular for the expression of purpose (see Gerritsen 1987: 143–147), becoming obligatory in many positions in Modern Dutch and remaining there whenever purpose or direction is intended.

2.1.2 Phonetic reduction of to

The phonetic reduction of to to te can be found in Middle English, as shown in Table 1. In Middle Dutch we already only find the reduced form, but this can be reduced even further to a single phoneme t attached to the infinitive (Stoett §283 gives the form tsine for te sine ‘to be’). I have found a few bound forms in the Helsinki corpus too, all from the later Middle English period (examples are given in note 7).

<table>
<thead>
<tr>
<th></th>
<th>1150–1250</th>
<th>1350</th>
<th>1420</th>
<th>1500</th>
<th>1570</th>
<th>1640</th>
<th>1710</th>
</tr>
</thead>
<tbody>
<tr>
<td>forto</td>
<td>1</td>
<td>29</td>
<td>46</td>
<td>46</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>forte</td>
<td>87</td>
<td>15</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>for to</td>
<td>14</td>
<td>91</td>
<td>323</td>
<td>251</td>
<td>41</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>for te</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(te, t’, to-</td>
<td>36</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

2.1.3 Loss of semantic integrity

We see the occasional use of the to-infinitive in Middle English in structures where it cannot possibly be goal-oriented, i.e. in positions where the plain infinitive and the present participle (which express simultaneity rather than purpose) had been the rule in Old English (cf. Fischer 1996: 119–121). The following is an example from a fourteenth century text, where to wepe clearly expresses a state not a purpose.

(3) And in my barm ther lith to wepe /Thi child and myn ...

and in my bosom there lies weeping thy child and mine

(Macauley 1900–1901, Gower, Conf.Am. III 302)
In Middle Dutch, too, the usual forms in these constructions were the plain infinitive and the present participle (but a coordinated construction is also quite often found, cf. Stoett 1909: §§10, 281). But here, too, the te-infinitive, which becomes the rule in later Dutch (cf. the examples in (4)), begins to make headway,

(4)  a. Hij lag te slapen
     he lay to sleep
     ‘He lay sleeping’

       b. Zij stond te wachten
       she stood to wait
       ‘She stood waiting’

2.1.4 Occurrence of to-infinitives after prepositions other than for

We see the occasional occurrence in Middle English of a to-infinitive preceded by another preposition which also governs the infinitive, making clear that to can no longer be prepositional. According to Visser (1969: §976), this structure does not occur in Old English, and is very rare again in later English. Most of his examples are from the period 1200 to 1500. Some illustrations are given in (5),

(5)  a. bliss of herte that comp of God to lovie
     the happiness of heart that comes from God to love
     ‘... from loving God’ (Morris 1965, Ayenbite 93)

       b. 7 himm birph zeornenn a33 bätt an, / Hiss Drihtinn wel to cwemenn / ... Wiph messess 7 wiph beness / 7 wiph to letenn swingenn himm
       and it behoves him to always desire that one (thing), to please his Lord well ... with masses and prayers and with to let scourge himself
       ... ‘and by having himself scourged ...’
       (Holt 1878, Orm. 6358–62)

In Middle Dutch, and more frequently in early Modern Dutch, the te-infinitive begins to occur too after other prepositions, such as van ‘of’, met ‘with’, na ‘after’, and especially sonder ‘without’, and these constructions can still be found in present-day Dutch, especially in colloquial speech (see Stoett 1909: §§282–283; Overdiep 1935: §§354–358). Some seventeenth century instances are given in (6),

(6)
2.2 New developments involving to

So the initial stages in Middle English look like a regular grammaticalisation process. However, towards the end of the Middle English period the trend seems to reverse. All the structures discussed in Section 2.1 above seem to disappear. Table 1 makes quite clear that the strengthening of the to-infinitive with for, disappears quite suddenly — at least from the Standard language — in the early Modern period. I believe that the reason for this is that to went back to its original meaning, again strongly expressing goal or direction (there is some difference with Old English usage, I will come back to that below). But apart from the disappearance of the grammaticalisation characteristics enumerated in (2), there are also new developments that indicate the renewed, semantic independence of to before the infinitive.8

(7) a. appearance of split infinitives
b. absence of ‘reduction of scope’
c. no loss of semantic integrity

2.2.1 The appearance of split infinitives

The first split infinitives are attested in the fourteenth century (see Mustanoja 1960: 515; Visser §977; Fischer 1992a: 329–330).9

(8) Blessed be thou lord off hevyn ... / That suche grace hath sent to his
Synful men for to bus lede / In paradice
(Cursor Mundi, Morris 1876, Laud Ms 18440–44)

This shows that the grammaticalisation of to is disturbed in that the usual process would have been for grammaticalised to to become more and more ‘bonded’ to the infinitive, in accordance with one of the grammaticalisation parameters distinguished by Lehmann (1985).
2.2.2 Absence of the ‘reduction of scope’
Another phenomenon showing ongoing grammaticalisation, also mentioned by Lehmann, is the reduction of scope, which is absent in English. In Dutch, on the other hand, there is reduction of scope: when two infinitives are coordinated in Modern Dutch, it is the rule for both infinitives to be marked by te, if the first one is so marked (for some idiomatic exceptions see Fischer 1996: 112–113). In other words the scope of te has been reduced to its immediate constituent. This is not the case in English where the first to can have both infinitives as its scope, as the literal English translation of the non-acceptable Dutch example in (9) shows,

(9) a. *Je kunt deze shampoo gebruiken om je haar mee te wassen en je kleren schoonmaken

b. You can use this shampoo to wash your hair and clean your clothes

2.2.3 No loss of semantic integrity
Another of Lehmann’s parameters, ‘the loss of integrity’, is relevant here too. It is clear that in Dutch, te has gradually lost its semantic integrity, i.e. it has become de-iconicised, and no longer expresses ‘goal’ or ‘direction’: this is now expressed obligatorily by om te. One result of this semantic loss in Dutch was already mentioned in Section 2.1.3 above. Another one is the appearance of the te-infinitive with a future auxiliary in Dutch. Overdiep (1935: §336) shows that they are quite regular already in early Modern Dutch. This again is a clear contrast with English where such a future infinitive simply never develops, neither in Middle English when shall and will could still be used in infinitival form, nor later with the new future auxiliary to be going to. Overdiep also mentions that zullen is especially common when the matrix verb itself is not inherently future directed, so after a verb like say. The reason for this difference may be clear by now. To itself expressed future and therefore had no need for a future auxiliary, whereas Dutch te no longer carried future meaning; it had become empty of referential meaning, and therefore the Dutch infinitive may need reinforcement.

The loss of the purposive meaning of te has also widened the possibility of using non-agentive subjects with a te-infinitive in Dutch (showing grammaticalisation along the ‘animacy’ hierarchy, cf. Hopper and Traugott 1993: 157). With a verb like dreigen ‘threaten’, the use of a non-agentive or an expletive it subject (i.e. with the verb being used epistemically) is quite common in Dutch, while it is more awkward in English, because of the stronger purpose meaning of to,
(10) a. *Het dreigde te gaan regenen, toen ik het huis verliet*

'It threatened to rain, when I left the house.'

Traugott (1993) notes that occurrences with expletive *it* are very rare in her corpora but that they do occur. It is interesting to note that all ten informants that I questioned, except one, do not like this sentence. They would much prefer, *It threatened rain* or *Rain threatened* or *It looked like rain*, and they find it also more acceptable with a progressive verbal form. Obviously it is not a construction they would comfortably use themselves. Traugott also notes (1993: 187) that although inanimate subjects with *threaten* occur, there is usually “something about the subject that leads to an expectation of the proposition coming into being”; in other words, there is a strong tendency still present to ascribe some agentive function to the subject. Similarly (10b) can be non-agentive in Dutch,

(10) b. *Hij dreigde van zijn fiets te vallen*

*‘He threatened to fall off his bike.’*

All my informants except one agree that (10b) is only possible in English if the subject wanted to injure himself and thus inflict pain on someone who cared for him. The epistemic meaning is the usual interpretation of this sentence in Dutch. When the subject has to be interpreted as agentive, the preferred construction would be with a finite clause: *Hij dreigde dat hij ....*

The same situation holds when *dreigen* is followed by a passive infinitive, making an agentive function of the matrix subject, which is also the subject of the infinitive, impossible (cf. Traugott 1995:34: ‘the passive demotes the inference that the subject ... is volitional or responsible with respect to the purposive clause’). (11) then, is a perfectly possible sentence in Dutch, but unacceptable for most English speakers,

(11) *Hij dreigde ontslagen te worden*

*‘He threatened to be fired.’*

And a construction like (12) is ambiguous in Dutch, but not usually in English,

(12) *Hij dreigde haar te doden*

‘There was a danger that he would kill her.’

‘He threatened to kill her’

English speakers strongly prefer the second, agentive interpretation. The reason for these differences is the fact, as I mentioned above, that *to* in English is still more strongly purposeful and therefore by default as it were one expects a controlling agent. It should be mentioned here that *threaten/dreigen* is not the
only verb that shows this difference in usage between Dutch and English. Traugott (1993) also discusses the behaviour of the verb *to promise*, which verb again in Dutch can be used non-agentively more easily than in English. In Fischer (1997a: 271–273), I also point out that in English *to*-infinitives regularly occur with the categories of verbs that Haspelmath (1989) has described as ‘irrealis directive’ and ‘irrealis potential’, but not with the categories ‘realis non-factive’ and ‘realis-factive’. The latter two categories contain clearly non-directional verbs, and it is interesting that in Dutch and German these last two categories do take *te/*zu infinitives much more easily than in English. Thus a verb like *affirm* does not take a *to*-infinitive in English, but its Dutch semantic equivalent *verzekeren* does (for more details see Fischer 1997a).

A final difference between Dutch and English is the formation of new modal auxiliaries in English consisting of a matrix verb that has semantically inherent future reference and the *to* element that belongs to the infinitive following the verb, as in *to be going to/gonna*, *to want to/wanna*, *to have (got) to/gotta* etc. Plank (1984: 338–339) notes that these verbs are unlike auxiliaries in that they occur with *to*, but notes at the same time that these same auxiliaries “allow the conjunction [i.e. *to*] to be reduced and contracted in informal speech”, even when this is not fast speech, and before pauses, indicating that this *to* has grammaticalised and become as it were affixed to the matrix verb. This amalgamation is possible because both *to* and the matrix verb express future modality. (So it seems *to* can become further grammaticalised in English only when it coincides with another future-bearing element.) In Dutch, however, this development has not taken place, because there was no meaningful ‘future’ or purposeful *te* for the matrix verb to attach to. In fact, whenever we do get a (semi-)auxiliary followed by a *te*-infinitive, it is clear that *te* goes with the infinitive. This is shown in the position of the adverb in examples such as the following,

(13) a.  *Ik zit nu te denken /*Ik zit te nu denken*  
I sit now to think /I sit to now think  
‘I am thinking now’

b.  *Het dreigt thans te mislukken /*Het dreigt te thans mislukken*  
It threatens now to fail /It threatens to now fail  
‘there is a possibility that it will fail’

In linguistically similar cases in English, the adverb can occur between *to* and the infinitive, showing that *to* and the infinitive do not form a cluster. That *to* in fact forms a cluster with the matrix verb is shown by cases in which matrix verb and *to* can be contracted as in the second example of (14),
3. The role played by the grammar

Now the question must be asked, what has caused the reversal in the grammaticalisation of *to*? I believe this is due to the grammatical circumstances under which *to* developed. In one respect English came to differ radically from Dutch, and this influenced the use and interpretation of *to*. In early Middle English the infinitive became much more strongly verbal than in Dutch (for instance, Dutch infinitives can be preceded by a possessive pronoun or an article, which is impossible in English (for more details see Fischer and van der Leek 1981:319). This verbal nature of the infinitive was strengthened by the fact that *to*-infinitives started to replace *that*-clauses on a grand scale in the Middle English period (cf. the rough statistics in Manabe 1989, and more specifically Los 1998); that is, they replaced clauses which have a tense-domain separate from the tense expressed in the matrix clause (cf. Fischer 1997c). This caused the element *to*, which originally expressed ‘goal’ or direction, to function as a kind of shift-of-tense element. What I mean is, *to* came to express a ‘break’ in time, a movement away from the time of the main clause; i.e. it again expressed ‘direction’. It is indeed only in English that we later (the first examples date from the late Middle English period) see the development of two different kinds of infinitival complements after perception verbs, where *to* becomes crucial in expressing a shift in tense.

(14)  I want to immediately go there
      I wanna go there immediately

(15) a.  *it thoghte hem gret pite / To se so worthi on as sche, / With such a child as ther was bore, / So sodeinly to be forlore*  
     ‘it seemed to them a great pity to see so worthy a woman as she was to be destroyed together with the child that was born to her’  
     (Macauley 1900–1901, Gower, Conf.Am. II, 1239–42)

b.  ‘*for certeynly, this wot I wel,’ he seyde, / ‘That forsight of divine purveyaunce / Hath seyn alwey me to forgon Criseyde,’*  
    ‘for certainly, this I know well, he said, that the foresight of divine providence has always seen that I would lose Criseyde’  
    (Benson 1988, Chaucer T&C IV, 960–62)

In both cases the *to*-infinitive refers to something happening in the future. The construction contrasts with the usual complement structure of physical perception
verbs, which until then had only allowed a bare infinitive, expressing the simultaneous occurrence of what had been seen, heard or felt, as in *I saw her cross(ing) the street*. In present-day English, this *to*-infinitive after perception verbs now no longer expresses future time, but it still expresses a shift in tense, making the experience indirect, as in (16).

(16) Alex saw Julia to have been in a hurry when she dressed (because she was wearing her T-shirt inside out) (the example is from van der Leek 1992: 13)

The type of construction shown under (15) was further strengthened by the influx of Latin-type accusative and infinitive constructions (aci) (as in (17)), which again appear in the late Middle English period (i.e. when *to* 'reverted') showing similar 'breaks' in tense between matrix verb and infinitive.

(17) I expect him to be home on time

These Latin-type aci constructions always have a *to*-infinitive in English. (For more details on this development, see Fischer 1992b, 1994a.). It seems that we can conclude that special syntactic circumstances as it were forced infinitival *to* to become more isomorphic again with the preposition *to*.

4. A brief conclusion

I have tried to show that grammaticalisation processes do not always run the same course, that there may be differences between similar languages, that the process may indeed be reverted, and that this relates to the specific grammatical circumstances that a language finds itself in. In other words, grammaticalisation need not be a process driven purely semantically, whereby the grammatical changes are the result solely of semantic and/or pragmatic change. The way the process developed in the case of *to* was (co-)determined by syntactic factors, specific to English, and by universal iconic constraints or patterns such as persistence and isomorphism.

Acknowledgments

I would like to thank audiences at Zürich and Vienna University, the editors of *Viewz* (in which an earlier version of this paper was prepublished [Viewz 7.1, 1998]), Elizabeth Traugott and Dieter Stein, and especially my colleagues Frederike van der Leek at Amsterdam University and Bettelou Los at the Free University of Amsterdam for their valuable criticisms and comments, and for their
willingness to discuss this paper with me. The paper has prospered from it, but of course those errors and weaknesses that have no doubt remained are all my own responsibility.

Notes

1. For a general discussion of the necessity to make a distinction between grammar change and aggregate language change, see also Lightfoot 1999: 220ff. Even though logically this distinction (i.e. between the process of language acquisition and the historical development of language) is clear-cut, the question remains as to what provides a proper explanation for language change. It must be clear that explanations are to be found both in language use (the triggering experience) and in the theory of grammar. The importance of each for explanation all depends on how ‘deep’ one’s model of grammar is, the shallower one’s grammar is, the more an explanation will have to be found in the changing circumstances of language use (see also Lass 1997 for a discussion of explanation and grammar-depth).

2. Bybee et al. (1994: 289ff.) recognise three other mechanisms of semantic change that play a role in grammaticalisation (it is quite clear that for them the mechanisms of semantic change are more or less equivalent to the mechanisms found in grammaticalisation, see p. 282), i.e. (3) generalisation, (4) harmony and (5) absorption of contextual meaning. It is clear from their description that all these mechanisms are essentially metonymic in nature, with metaphor playing a subsidiary role. Indeed they conclude (p. 297): “The most important point that can be made from the discussion of mechanisms of change is that context is all-important”.

3. Heine et al. (1991a) indeed refer to the process as a ‘chain’.

4. It is interesting to note that Bruyn (1995), who was looking at the developments taking place in a pidgin becoming a creole (where it is believed that grammaticalisation plays an important role), and so, as it were, looking for grammaticalisation evidence from another perspective (not on the basis of selected cases from many languages as is usually done, but on the basis of full data from one language where the process might be expected to apply according to the hypothesis), found very few cases where grammaticalisation ran its full course. She found that language contact (especially substratum influence) often caused divergence (p. 241ff.), or early abortion (p. 53ff.), or that sometimes a development was much more abrupt than is usual in grammaticalisation cases (pp. 237–239). Her investigation shows that it is important at each stage to take into account the synchronic circumstances, which will ultimately (and freshly) decide what will happen.

5. The principle of unidirectionality has been much in the limelight recently. The strongest adherent of the principle must be Martin Haspelmath (1999), who indeed believes it is without exceptions. This view came under strong attack at the recent (1999) conference on grammaticalisation held at the university of Potsdam, in papers delivered by among others, Johan van der Auwera (“Degrammaticalization”), Aidan Doyle (“Yesterday’s affixes as today’s clitics: functional heads and grammaticalisation in Irish”), Muriel Norde (“The final stages of grammaticalization: affixhood and beyond”), Günter Rohdenburg (“Degrammaticalization phenomena with ‘seeing’ and other adverbial conjunctions in English”), who all argued, to my mind convincingly, that true cases of deggrammaticalization exist, even though they are more rare. Roger Lass (this volume) goes into the philosophical, methodological and empirical problems that this notion entails.

6. For more details, also on the comparative development of zu in German, see Fischer (1997a). This article takes a different approach in that it considers the degrees of grammaticalisation of
to, zu and te only from the point of view of the parameters of grammaticalisation distinguished by Lehmann (1985).

7. See e.g. *King Horn* (Hall 1901: 25), te lyue, and signs of reduction in forms such as tobinde (*Havelok*, Smithers 1987: 56), tobe (Rolle, *Psalter*, Bramley 1884: 380), and tavenge (Caxton, *Reynard*, Blake 1970: 54). All these instances are from late Middle English, none have been found in earlier or later periods in the corpus I used (the Helsinki Corpus).

8. Some of the editors of *View* (in which this paper was prepublished, see *Vienna English Working Papers* 7.1, 5–24) raised the question here whether the developments mentioned under (7) are not “independent syntactic developments which only happen to affect the way in which the infinitive marker can be placed more or less accidentally”. This is of course a question of interpretation. I do not believe they are independent because they take place at more or less the same time, and because the same developments did not take place in either Dutch or German. These developments are indeed closely tied up with the process of grammaticalisation (as described in more detail in Fischer 1997a). The same editors mention as a possible counter-example the case of the ‘group genitive’ as in the teacher of music’s room, which according to them is not a case of de-grammaticalisation. I think it is interesting to observe in this connection that Janda (1980) has indeed suggested that it is a case of de-grammaticalisation (the genitive-s being reinterpreted as a possessive pronoun).

9. There are also cases of split infinitives found in early Middle English but these are of a different type, they involve infinitives preceded by the negative particle or by a personal pronoun. Presumably (cf. van Kemenade 1987) these particles/pronouns are still clitics, which explains their position next to the infinitive which was becoming more verbal around this time.

10. The importance of to for the development of these new modals is also emphasised by Hopper and Traugott (1993: 81), where they write, “[t]he contiguity with to in the purposive sense must have been a major factor in the development of the future meaning in be going to as an auxiliary”. A full discussion follows on pp. 81–83. And see also Fitzmaurice (this volume) on the link between the degrammaticalization of infinitival to and the grammaticalization of to in combination with semi-auxiliaries.

11. For more details on how this rather difficult example (of which there are three in Chaucer) should be interpreted, see Fischer 1995: 10–11.

**Texts**


References


Fischer, O.C.M. 1999. “On the role played by iconicity in grammaticalisation processes”. In M. Nanny and O.C.M. Fischer (eds), 345–373.


UNIDIRECTIONALITY: TO BEFORE THE INFINITIVE

