On the role played by iconicity in grammaticalisation processes

Fischer, O.C.M.

Published in:
Form Mimicing meaning: Iconicity in Language and Literature

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
On the Role Played by Iconicity in Grammaticalisation Processes*

Olga Fischer

University of Amsterdam

1. Introduction: The iconic and symbolic poles in language

It seems to be accepted by most linguists that the iconic drive or instinct is very strong in language users and that indeed language, both phylogenetically and ontogenetically, started/starts off iconically. Thus, Slobin (1985) in an article on “The child as a linguistic icon-maker” shows that many of the ‘mistakes’ that children make are iconically motivated, and Givón (1995a: 406), among others, characterises the rules of what he calls the ‘proto-grammar’ of pidgins — which are emergent, not fully evolved languages — as having as a “common denominator ... that they are extremely iconic”. Thus, what is concrete takes priority. This can also be seen in the fact that children acquire the lexicon much earlier than the grammar, and that the grammars of pidgins evolve from the lexicon. Givón notes too that animals can be taught lexical code labels but that teaching them “the natural use of anything remotely resembling human grammar — morphology and syntax — ... has been almost a uniform failure” (Givón 1995a: 401). It has also been widely noted that iconicity is not confined to the initial stages of language, or indeed to language itself. Plank (1979: 131) writes, “der frühkindliche Spracherwerb zeichnet sich durch eine ausgeprägte Präferenz zu ikonischer Zeichenbildung aus, die als ‘natürliches’ Substrat jeder Zeichenbildung zumindest latent wirksam bleiben dürfte, wenn auch nach Maßgabe von Symbolisierungsnotwendigkeiten” (‘language acquisition in young children is characterised by a very strong preference for the formation of iconic signs, which, as a natural substratum of all sign-formation probably remains at least latently active, even though
tempered by the need to symbolise’, italics added), and Givón writes in an article (1995b: 61), in which one of his concerns is to illustrate the long antecedence of isomorphic coding in biology, that “one must consider the pervasive iconicity of human language merely the latest manifestation of a pervasive preference for isomorphic coding in bio-organisms”.

On the other hand, it is also generally accepted that “human language is shaped by competing functional imperatives” (Givón 1995a: 393), i.e. what Haiman (1983) has referred to as iconic and economic motivation. In other words, language is subject to corrosive pressures, which are due to the need for processing speed and the general erosion (the loss of expressivity) caused by diachronic change. One could say then that language moves or is situated along an axis with two poles: an iconic, concrete pole at one end, and a symbolic (perhaps ‘arbitrary’ or ‘conventional’ is a less confusing term here), abstract one at the other (cf. also Plank 1979). One could also refer to the iconic pole as original and creative (less confined to language per se) and to the symbolic as derivative and mechanistic. But this division into iconic and symbolic, should not lead us away from the fact that even the symbolic is to some extent iconic (after all it is derived): it shows an iconicity of a more abstract order. It is here that the distinction between imagic iconicity and diagrammatic iconicity (first made by Peirce and Jakobson, and further elaborated upon by Haiman, Givón and others) has become so important. The mechanism in both is the same, but only in imagic iconicity, is there a straight iconic link between the verbal sign and the image or object (the ‘signans’ and the ‘signatum’), as for instance in onomatopoeia. Diagrammatic iconicity is more like a topographic map, where the relation between objects or concepts in the real world (as we see it) can be deduced from the relations indicated on the map. Thus, the idea of space in the real world is proportionally reflected in the map.

But even within diagrammatic iconicity, there are differences in terms of concreteness. It is interesting to observe, for instance, when Max Nänny applies the various types of diagrammatic iconicity distinguished by Haiman (1980) to poetry, that the use made of it there is more concrete than the examples that Haiman gives from the more conventional syntax of everyday speech. Let me illustrate this with one diagrammatic type, i.e. the ‘distance principle’, which Haiman (1983: 782) has characterised as follows: “the linguistic distance between expressions corresponds to the conceptual distance between them”. He gives as an example the difference between the following causative expressions,
(1)  
   a. I caused the chicken to die.  
   b. I killed the chicken.

The distance between the causative verb *cause* and the infinitive *to die* makes the activity — which is expressed by one verb, *kill*, in the second example (so with zero distance between agent and act) — less instantaneous, providing a possible slot for an intermediate agency (as Haiman adds, there might be an element of magic involved). Other often quoted examples where the distance principle is at work are,

(2)  
   a. He sprayed the door with green paint.  
   b. He sprayed green paint on the door.

and

(3)  
   a. Who has taught the children French?  
   b. Who has taught French to the children?

Here the distance between the verb and the two objects in each clause is an indication of the degree in which an object is affected by the verb: the closer it stands to the verb the more affected it is. Thus, the (a) example in each pair implies that the objects, *the door* and *the children*, have been totally affected, i.e. that *the door* is completely green, and that *the children* indeed know French, which is not implied by the (b) examples, because the same object is now distanced from the verb by the intervention of another. In Nanny’s (1986: 204) example of ‘distance’ from Wordsworth’s *Prelude* (Bk V, II. 79–80),

(4)  
   ... underneath one arm
   A stone, and in the opposite hand, a shell

the distance between the position of *stone* and *shell* is also expressed diagrammatically, or proportionally, by the distance between them on the page, but at the same time this distance is still concrete, it expresses *spatial* distance. This is similar to an example Nanny gives of another type of diagrammatic iconicity, i.e. the ‘principle of sequence’, where the order of the signs not only reflects the sequential order of realities referred to in the world, as in the well-known *veni, vidi, vici*, but again is more concrete in that a sign placed first, often contains the word *first*, and a sign placed at the end of the clause contains the word *last* or *end* or *adieu* etc. (Nanny 1986: 205–5).

It seems to me important to emphasise that the creative and the mechanistic,
or the iconic and the symbolic, go hand in hand, that they are as it were always present at the same time in everyday language use. Although at certain points we, as speakers, may rely more on the one or the other, i.e. we are more iconic when creating language afresh (as children, creators of a pidgin, poets), and are more mechanistic when conveying a message in everyday communicatory circumstances, we are still always at the crossroads of both possibilities. Fonagy has referred to the simultaneity of these two codes (he calls it ‘the dual encoding procedure’) in a large number of his articles (e.g. Fonagy 1982, 1995, and this volume). He refers to the arbitrary, symbolic structures as the primary message or code (i.e. the rules of grammar) and to the iconic mode as the secondary message. One could argue about the labels ‘primary’ and ‘secondary’. Which is seen as primary and which as secondary, depends on one’s point of view. The rules of Fonagy’s secondary code are primary in terms of language history; that is “they can be regarded as the vestiges of a pre-linguistic system of communication” (Fonagy 1982: 93); they are motivated and therefore easier to recognise. But they are also secondary because each time they are used they are “created through the meaningful modulation or distortion of the primary message generated by the grammar” and are “superimposed upon the primary message” (Fonagy 1982: 92). Another aspect that Fonagy stresses again and again is the playfulness of linguistic behaviour (see also Haiman 1994, and Lecercle 1990): we just love kicking over the traces, bending the rules. Indeed the ‘rules’ that Lecercle gives for the secondary or iconic code (which he calls the ‘remainder’) consist for a large part of overturning the rules of grammar “through either excess or lack” (Lecercle 1990: 122). It is clear that the iconic rules arise partly on the basis of the primary code.

Having touched upon our iconic needs, our iconic instincts, I would now like to turn to the process of grammaticalisation with this dual encoding system in the back of our minds. First some general points about grammaticalisation. 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 a description of a diachronic process of linguistic change. Let us first look at the way the process has been described. 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. 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 evolutorial 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’. Again the type of metaphor used here different from metaphors used in poetry, the shift is less concrete, it is less daring, more predictable. Metonymic change can be related to re-analysis and functions on the syntagmatic plain. It takes place mainly via the “semanticization [or grammaticalisation] of conversational implicatures” (Hopper and Traugott 1993: 84). As far as cause in 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 need to be more specific, in the tendency to infer as much as possible from the input, and in the necessity of interpreting items in context”. Likewise Hopper and Traugott (1993: 86) concur with Heine et al. (1991a: 150–51) 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”.

Although I would agree with the views discussed in the previous section 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 (but I also feel that attention ought to be paid to the much less controllable element of ‘play’ mentioned above), 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. Grammaticalisation processes 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. 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) 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. So why should a grammaticalisation process necessarily run from a to b to c etc.? Why should there be unidirectionality? 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.

What I would like to do in the remainder of this study is to show how strongly iconic the forces are that may lead to a process of grammaticalisation or its abortion, to further routinisation or to recreation.

2. Grammaticalisation and iconicity: General

In this section, I will discuss some general aspects of the relation between grammaticalisation and iconicity. Within diagrammatic iconicity, Haiman (1980) makes a distinction between isomorphism and iconicity of motivation (the latter comprising sequential ordering, markedness, centrality vs peripherality, distance vs proximity, repetition). Haiman states that isomorphism is the more general principle, and is universal, while the other types are not. I believe that isomorphism is also more basic in that it underlies some or perhaps all of Haiman’s
ON THE ROLE PLAYED BY ICONICITY IN GRAMMATICALISATION

motivation types, and it is also the foundation of metaphor and/or analogy (the mechanisms important, as we have seen, in grammaticalisation processes).

The principle of isomorphism (or ‘one form — one meaning’\(^{10}\)) predicts that synonymity on the one hand, and polysemy and homonymy on the other, are not allowed, or at least avoided as much as possible. In other words, if \(\alpha\) (in (5)) represents the signans and \(x\) the signatum, then there exists the basic principle that one \(\alpha\) (one language sign) should relate to only one \(x\) (one concept or object in the world as we view it). Thus the relationship between signans and signatum will preferably be as in (5a), and forms like (5b) should not be allowed:

\[
\begin{align*}
(5) \quad & \text{a. } \frac{\alpha}{x} \quad \text{(isomorphism)} \\
& \text{b. } \frac{\alpha\alpha}{x} \quad \frac{\alpha}{xx} \quad \text{(synonymy/polysemy/homonymy)}
\end{align*}
\]

This isomorphic or relational diagram (a term used by Hiraga 1994, who offers a lucid discussion of this) is also the basis for metaphor since metaphor arises through analogy of meaning; i.e. when another meaning, \(y\), comes to be related to \(x\), so that they share or become one \(x\), it is felt that likewise a single sign, \(\alpha\), is enough to express the signans/signatum relationship. Itkonen (1994: 45–46) shows how basic the “urge to analogize” is. He writes that analogy is a way of patterning and interpreting the world; a child must analogise in order to acquire knowledge of the external world and to survive in it. Itkonen shows how analogical inference is in fact a generalisation involving the properties of cooccurrence and succession. Thus, we observe by experience that a fire is hot, and infer by analogical generalisation that (every instance of) fire is hot. We can see then that Haiman’s iconicity of motivation types show a similar analogy (making use again of cooccurrence and succession); sequence, for instance, being represented by (6a) and repetition by (6b),

\[
\begin{align*}
(6) \quad & \text{a. } \frac{\alpha \rightarrow \beta \rightarrow \gamma}{x \rightarrow y \rightarrow z} \\
& \text{b. } \frac{\alpha \alpha}{x \times x}
\end{align*}
\]

This very basic, iconic analogy can be seen to be at work in grammaticalisation. It should not come as a surprise that I would claim (pace Heine et al 1991b: 25, and Hopper and Traugott 1993: 32ff.) that analogy is a \textit{cause} and not merely a mechanism in grammaticalisation processes (for a discussion of analogy as both
Analogy (or isomorphism, as a special type of analogy) then, can be seen to be responsible for the following factors in grammaticalisation, showing the influence of iconicity on grammaticalisation processes:

1. Metaphorical shift
2. 'renewal' and 'layering'
3. Erosion or phonetic reduction
4. 'persistence': the preservation of original meaning

ad (i): Metaphorical shift. This, I believe, needs no further elaboration in that I have already discussed the role played by metaphor in grammaticalisation processes in Section 1. The metaphorical shift from (more) concrete to (more) abstract also underlies a number of clines and/or hierarchies that have been distinguished in grammaticalisation studies, such as Heine et al. (1991a: 157) who make use of a 'categorial hierarchy',

(8) person > object > process > space > time > quality

which in turn is related to various other hierarchies in grammaticalisation, such as the 'case-hierarchy' (agent > benefactive > dative > accusative > locative > instrument); 'word-type hierarchy' (noun > verb > adverb/adposition > adjective) and the 'animacy hierarchy' (human > animate > inanimate > abstract) (see e.g. Heine et al. 1991a: 159–60; 1991b; Hopper and Traugott 1993: 157).

ad (ii): Renewal. Through grammaticalisation the isomorphic or transparent relation between signs and concepts often becomes obscured, which may then be 'repaired' by renewal. This may be cyclical in so far as the renewed form may again become grammaticalised. The renewal may even speed up the process of grammaticalisation, when the new and old forms are in competition (cf. Hopper and Traugott 1993: 123). Thus, for example, the opacity of strong past tense forms in English is being repaired by the emergence of new weak tenses, in which pastness (represented by y in (9)) is isomorphically related to a separate linguistic sign, -ed (which may go back to a cognate of the full verb do),

(9) \[ h_{old} > h_{elp} + ed \]

\[
\frac{\alpha}{x + y} > \frac{\alpha + \beta}{x + y}
\]

Renewal cannot be explained by the process of grammaticalisation itself, but is due to a constant pressure for a more iconic (more transparent) isomorphic
structure. Harris and Campbell (1995: 72–75) show in a discussion of what they call ‘exploratory expressions’, how new expressions are always floating around in language. They may become grammaticalised if the need arises but they may also disappear or remain as lexical idioms outside the grammar. Again I believe that many of these exploratory expressions can be explained iconically, i.e. they are motivated.

ad (iii): The process of phonetic reduction has often been shown to be explainable by the iconic principle of quantity, which according to Givón is also an instance of isomorphism. Thus, “a larger chunk of information will be given a larger chunk of code or on a more abstract level, less predictable or more important information will be given more coding material (Givón 1995b: 49).”

ad (iv): With reference to persistence, Bybee et al. (1994: 16) write that “certain more specific nuances of the source construction can be retained long after grammaticalisation has begun” (cf. also Hopper and Traugott 1993: 91, 120). I think it is important to note that this is especially the case (as the discussion in the literature shows) in circumstances of layering, i.e. when older and newer forms coexist expressing more or less the same function. For instance in English, a variety of more or less grammaticalised forms is used to express future time (the zero form, will, shall, to be going to/gonna), but the expressions are not synonymous. Their specific nuances can be explained diachronically, on the basis of their original meaning. In other words, the original relation between the signans and the signatum has remained in spite of grammaticalisation, preserving isomorphism.

3. Grammaticalisation and iconicity: special cases

I have on purpose emphasised the words in spite of in the last paragraph because I believe there are often forces within grammaticalisation that run counter to what is expected (cf. the description of grammaticalisation as given in Section 1). This, for instance, pertains to the idea that grammaticalisation is driven semantically, that it is unidirectional, and that generalisation or bleaching of meaning is a necessary element to start off the process (but cf. note 4). As I said before, looking at a case of grammaticalisation in more detail may show that not all developments proceed in the same direction and follow the general principles set out above. So I would like to conclude with an investigation of two cases in the
history of English which have generally been interpreted as more or less paradigm cases of grammaticalisation, comparable to similar processes of grammaticalisation taking place in other languages. The discussion of the first case — the grammaticalisation of *have to* from a full verb into a semi-modal auxiliary — will be brief since I have already described that at some length in Fischer (1994b) and (1997b). The other case concerns the grammaticalisation of *to* before the infinitive.

### 3.1 *The grammaticalisation of* have to

In line with similar developments involving a possessive verb like *have*, where *have* in combination with an infinitive grammaticalised from a full verb into an auxiliary (e.g. as happened in the Romance languages, cf. Fleischman 1982), it has usually been taken for granted that English *have to* represented a ‘regular’ case of grammaticalisation. Thus, van der Gaaf (1931), Visser (1969: §1396ff.), and Brinton (1991) all more or less accept the following developmental stages for the construction, *I have my work to do/I have to do my work*:

1. **(10)**
   - *have* at first is used as a full verb, meaning ‘to possess’
   - the NP functions as the direct object of *have*
   - the *to*-infinitive is not obligatory
   - the infinitive functions as an adjunct dependent on the NP
   - word order is not relevant, it does not influence meaning

In subsequent stages of the development, the meaning of *have* slowly generalises and it acquires modal colouring in combination with the *to*-infinitive, which now becomes obligatory. The infinitive no longer functions as an adjunct to the NP but as an object complement of the matrix verb *have*, and the original object of *have* becomes an argument of the infinitive. This object need not be concrete or ‘possessible’ any longer (i.e. it goes down the ‘category hierarchy’, referred to in Section 2). In the final stage we see the appearance of inanimate subjects (possessive *have* + infinitive always had animate subjects: so here we follow the ‘animacy hierarchy’ discussed above), and of intransitive infinitives, i.e. the original object can now be dropped altogether. Re-analysis or rebracketing from (11a) to (11b) now follows,

1. **(11)**
   - a. I have [my work to do]
   - b. I [have to do] my work
resulting in a fixed \textit{have} + \textit{to}-infinitive + NP word order (note the sudden shift in word order between (11) a and b, which is difficult to account for satisfactorily within this framework).

It is quite clear in this sketch of the putative development of \textit{have to} that the grammaticalisation proceeds along a path of semantic change, and that the syntactic changes are subordinate to it, following hard on the heels of the semantic change. Because the development is seen as gradual (as we have seen, one of the tenets of the grammaticalisation hypothesis), the various stages are extremely difficult to disentangle. This is noticeable also from the fact that van der Gaaf, Visser and Brinton do not agree as to when the different stages occur.

In my own investigation of this case (Fischer 1994b), I considered all the instances in which \textit{have} is followed by a \textit{to}-infinitive in the so-called Helsinki corpus (which covers the Old, Middle and early Modern periods). Looking at a total of 643 examples, I came to the conclusion that there is no evidence for a gradual semantic change as envisaged by the studies reported on above. The generalised meaning of \textit{have} (mentioned under the development in (10)) already existed in the earliest recorded (Old English) period, and modal colouring of \textit{have} was possible (but not necessarily with obligative meaning) in Old English too. Similarly, the object could be concrete or abstract from earliest times. All the more firm syntactic evidence for the change is very late (i.e. only from the early Modern Period). In fact, it can be shown that the syntactic changes follow upon a (general) word order change. The basic SOV word order of Old English, which persisted quite long in infinitival constructions in Middle English, ensured that the order of the three basic elements was in normal circumstances (so when no movement rules were involved) almost always \textit{have} + NP + \textit{to}-infinitive. The medially positioned NP could function equally well as an object of the main verb \textit{have} (due to the V2 rule in main clauses, which would move \textit{have} to a position before the NP object), and as an object of the infinitive. When the word order in late Middle, early Modern English became generalised to SVO everywhere, this kind of ambiguity was no longer possible: that is, the object NP had to shift to a postverbal position with respect to the verb which gave it its semantic role. Since \textit{have} usually had a generalised meaning in this construction, the usual position for the object became the one after the infinitive with which it had a stronger semantic bond. So it was the SOV > SVO word order change that fixed the order of the \textit{have to} construction, and which ultimately led to the re-analysis described in (11). Or, to put it differently, the word order change caused the adjacency of \textit{have} and the \textit{to}-infinitive, which in turn led to a semantic change,
in which *have* and the *to*-infinitive were considered one semantic unit. The data show quite clearly that all the clear grammaticalisation facts (intransitive infinitives, inanimate subjects, the use of ‘double’ *have to have*, the development of *have to* into an epistemic modal) date from after the word order change.

So what is notable about this case? What makes it different from the common or garden kind of grammaticalisation? We see,

(i) no gradual generalisation in meaning, no intertwining with gradual syntactic adaptations;

(ii) a quite sudden, unlinked syntactic change which sets the process towards auxiliary status rolling;\(^\text{13}\)

(iii) a semantic change in *have* to a modal auxiliary which is probably the result of the (syntactic) word order shift and of metonymic forces (conversational implicatures), caused by the inherent meaning of *to* (for which see below) and the occasional modal colouring that could be present all along in these *have* constructions depending on the other lexical items in the clause.

Further support for this sketch of the development may be found in the fact that in Dutch and German (closely related languages) the auxiliarisation of the cognates of *have to* did not take place, and neither did the SOV > SVO change, which, I believe, is behind the development. Why is it that the word order shift sets off this process? I believe it is iconic factors that are responsible here. In Fischer (1997b) I propose a reversal of Givón’s (1985: 202) ‘proximity principle’ to account for this, which in turn is derived from what he calls an iconic meta-principle.

(12) The closer together two concepts are semantically or functionally, the more likely they are to be put adjacent to each other lexically, morpho-tactically or syntactically.

One would expect the proximity principle also to be valid the other way around, i.e. the moment two elements are placed together syntactically or formally, it is likely that they will begin to function together semantically or functionally. Compare in this respect also Bolinger’s (1980: 297) remark: “the moment a verb is given an infinitive complement, that verb starts down the road of auxiliariness”.\(^\text{14}\)
3.2 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”. He 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 given in Jespersen (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: 10–11).

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-trodden 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 English and come to the conclusion that to and cognate Dutch te have not grammaticalised in the same way. On the contrary, it looks as if to was stopped early in its development and has even regressed in some respect. I think this could be characterised as a process of iconisation, a moving away from the symbolic pole back to the iconic one.

I believe indeed that the forces behind this process have been to a large extent iconic, although there were some syntactic factors too, which I will come back to later. The main factor is isomorphism. One can see that through the grammaticalisation of to, the original isomorphic relation between the signans and the signatum (as given in (13a) below) is disturbed (as shown in (13b)). The
sign to acquires two signata: the first is the prepositional purposive, allative ‘to’ and the second the semantically empty, infinitive marking element ‘to’. The asymmetric situation of (13b) 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, which would then coexist with the full form to. 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 (this would be stage (13c)). In fact with stage (13c) we have a new stable isomorphic relation. The other solution for the asymmetry of (13b) is to go back to the earlier symmetry (i.e. 13a). This also makes the relation isomorphic again, and it is more strongly iconic than (12c) because here the sign to is linked back up with its original meaning (cf. the general point I made about iconicity and persistence in (7iv)).

\[
\text{(13) stages of grammaticalisation of to} \\
\begin{array}{ccc}
\alpha & \beta \\
x & \text{goal} & y & \text{infinitival marker}
\end{array}
\]

So my suggestion is that diacronically English to moved back to stage (13a), while Dutch te moved on to stage (13c). In what follows, I will have a look at the (comparative) facts and also offer some suggestions why English to re-iconised.

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:

\[
\text{(14) the grammaticalisation of to in its early stages} \\
i. strengthening of to by for \\
ii. phonetic reduction of to \\
iii. loss of semantic integrity \\
iv. occurrence of to-infinitive after prepositions other than for
\]

ad (i): 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: 143–47), becoming obligatory in many positions in Modern Dutch and remaining there whenever purpose or direction is intended.

Table 1. The frequency of *for to* in the Middle English and early Modern English periods, based on the Helsinki corpus (taken from Fischer 1997a)

<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>for to</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>forté</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>

ad (ii): The phonetic reduction of *to* to *te* can be found in Middle English, see Table 1 above. In Middle Dutch we already find the reduced form only, 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 late Middle English period, as the last horizontal line in Table 1 shows.

ad (iii): 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) have been the rule in Old English (cf. Fischer 1996: 119–121). The following is an example from a fourteenth century text,

(15) *And in my harm ther lieth to wepe/ Thi child and myn ...*
    ‘And in my bosom there lies weeping thy child and mine’


In Middle Dutch, too, the usual forms in these constructions are 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 (16)), begins to make headway,
(16)  
_Hij lag te slapen_
he lay to sleep
'He lay sleeping'
_Zij stond te wachten_
she stood to wait
'She stood waiting'

ad (iv): 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 (17),

(17)  
a.  _bliss of herte that comp of God to lovie_
the bliss of heart that comes from God to love
'the happiness of heart that results from loving God'
   (Morris [Gradon] 1965, Ayenbite 93)
b.  _7 himm birhp 3eornenn a53 batt an/_ Hiss Drihhtinn wel to_
and him behoves yearn ay that one, his lord well to
   _cwemenn/_ ... _Wijp messess 7 wipp beness/7 wipp to letenn_
please, ... with masses and with prayers and with to let
   _swingenn himm_
scourge him
'and it behoves him to always desire that one thing, i.e. to please his Lord well ... with masses and prayers and and by letting himself be 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: §§282–83; Overdiep 1935: §§354–358),

(18)  
a.  _Hy starf, niet sonder seer beclaeght te wesen, den 8sten April_
he died, not without very lamented to be, the 8th April
'He died, not without being deeply lamented, the 8th of April'
   (van Mander, Overdiep 1935: 420)
b. ... sal ick eindigen naer mijn groetenisse aen alle
... shall I end after my greetings to all
\textit{de vrinden ghedaen te hebben}
the friends done to have

'I will end after having given my greetings to all my friends'
\cite{Reig.77/16, Overdiep 1935: 421}

Thus, the initial stages involving \textit{to} 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 (14i-iv) above seem to disappear. Table 1 makes quite clear that the strengthening of the \textit{to-infinitive} with \textit{for} disappears quite suddenly — at least from the Standard language — in the early Modern period. I believe that the reason for this is that \textit{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 (14), there are also new developments that indicate the renewed, semantic independence of \textit{to} before the infinitive:

\begin{enumerate}
\item new developments involving \textit{to}
\begin{enumerate}
\item appearance of split infinitives
\item absence of 'reduction of scope'
\item no loss of semantic integrity
\end{enumerate}
\end{enumerate}

ad (i): The first split infinitives are attested in the fourteenth century \cite{Mustanoja 1960: 515; Visser §977; Fischer 1992a: 329–30}.

\begin{enumerate}[20]
\item \textit{Blessid be pou lord off hevyn .../That suche grace hath sent to his} /\textit{Synfull men for to bus lede /In paradice} (\textit{Cursor Mundi}, Laud Ms 18440–44, Morris 1876)

'Blessed are you, Lord of heaven, who has sent his sinful people such grace as to lead (them) thus into paradise'
\end{enumerate}

This shows that the grammaticalisation of \textit{to} is disturbed in that the usual process would have been for grammaticalised \textit{to} to become more and more 'bonded' to the infinitive, in accordance with one of the grammaticalisation parameters distinguished by Lehmann's (1985).

ad (ii): Another phenomenon showing ongoing grammaticalisation, also mentioned by Lehmann, is the 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: 112–13). 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 (21) shows,

(21)  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

ad (iii): Lehmann’s fifth parameter, ‘the loss of integrity’, is also relevant here. It is clear that in Dutch, te has gradually lost its semantic integrity, i.e. it has become de-iconised, 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 under (14iii) 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, no 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 refers to the 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). 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,

(22) a. Het dreigde te gaan regenen, toen ik het huis verliet  
    "It threatened to rain, when I left the house"  
    b. Hij dreigde van zijn fiets te vallen  
    "He threatened to fall of his bike"
The same is true 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"). (23) is a perfectly possible sentence in Dutch, but unacceptable in English,

(23)  *Hij dreigde ontslagen te worden*
*He threatened to be fired*

And a construction like (24) is ambiguous in Dutch, but not in English,

(24)  *Hij dreigde haar te doden*
*There was a danger that he would kill her*

*He threatened to kill her*

English has only the second 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.20

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 in Dutch can be used non-agentively more easily than in English. In Fischer (1997a: 271-73) 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: 338–39) 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 that to could be further grammaticalised in English only when it coincided with another future-meaning 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,

(25) 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 (26).

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

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. Manabe 1989); that is, they replaced clauses which have a tense-domain separate from the tense expressed in the matrix clause. 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 (‘direction’) 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:

(27) 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 no longer expresses future time, but in examples like (28) (for those who accept this type of construction) it can still express a shift in tense, making the experience indirect,

(28) 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 (27) was further strengthened by the influx of Latin type accusative and infinitive constructions (as in (29)) appearing again in the late Middle English period, showing similar ‘breaks’ in tense between matrix verb and infinitive,

(29) I expect him to be home on time

These accusative and infinitive constructions always have a *to*-infinitive. (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 iconicity plays an important role in so-called grammaticalisation processes, both in a general sense, where it supports the hypothesised process of grammaticalisation, but also in specific cases, where it often runs counter to the expected development. The histories of have to and the to-infinitive show that the grammaticalisation path need not be one way, and need not be steered by the principles of semantic change alone. Iconicity plays an independent role, is in fact a more independent factor in syntactic change than grammaticalisation. It may be doubted even whether grammaticalisation should be looked upon as an independent causatory factor in the theory of linguistic change at all.

Notes

* I would like to thank my colleague at Amsterdam, Frederike van der Leek, for discussing this paper with me, for her valuable and inspirational comments, and for turning my attention to Holyoak and Thagard’s book, Mental Leaps. Analogy in Creative Thought. I also would like to express my gratitude to audiences at Zurich and Vienna for their questions and queries which have made me rethink and alter parts of this paper. Finally, and most importantly, I would like to thank Max Nanny, for stimulating me to make the ‘mental leap’ from linguistics to literature, a most fruitful analogy.

1. See also Song (1996: 5), who writes that typologically too there is an iconic relationship between the type of causative used and the (in)directness of the causation: “In order to express direct causation, languages tend to use the causative which exhibits a higher degree of fusion of the expression of cause and that of effect, whereas in order to express indirect causation, languages tend to use the causative in which there is a lower degree of fusion”.

2. Frederike van der Leek pointed out to me that what may be involved is not so much linear distance (because distance may change according to the construction used, cf. Which door did he spray with green paint?), but the fact that e.g. prepositions like on, with, to activate concepts of their own. Still, it could be said that these prepositions with their additional concepts create distance too between the verb and its argument.

3. For this mechanistic view of grammaticalisation see especially Bybee et al. (1994: 298), who write, “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 grammaticalization occur in language use for their own sake; it just happens that their cumulative effect is the development of grammar”. Although I would agree with the downtoning of the functional aspect, I do not believe that the process once set in motion is as automatic as they presume.
4. There is some difference of opinion as to the stage of grammaticalisation in which bleaching is most prominent. According to many investigators of the phenomenon, the bleaching of source concepts sets off the process (cf. Givón 1975, Lehmann 1982, Heine and Reh 1984, Bybee and Pagliuca 1985: 59–63, Heine et al. 1991a), but according to others (notably Traugott, see Traugott 1982, and Hopper and Traugott 1993: 87–93), bleaching is a process that occurs in the later stages of grammaticalisation, the semantic shift occurring at the beginning being one of pragmatic enrichment rather than loss.

5. 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 three 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”.

6. Heine et al. (1991b: 50, 60) indeed distinguish two types of metaphor. The type that occurs in grammaticalisation they call ‘experiential’ or ‘emerging’ metaphors, because these metaphors that arise in context (i.e. they are metonymic in nature), and they contrast them with ‘conceptual’ or ‘creative’ metaphors, which are much more likely to contain conceptual ‘jumps’ and cannot be predicted in any sense.

7. It is not at all clear from the literature I have studied what the status of grammaticalisation is in the theory of change. Vincent (1995: 434) talks about “the power of grammaticalisation as an agent of change”, which seems to suggest to me that he thinks it has explanatory value, that it has independent force. Most students of grammaticalisation describe it vaguely as a ‘phenomenon’, a ‘process’, an ‘evolution’. However, the fact that for most linguists one of its intrinsic properties is that it is unidirectional suggests to me that in their view the process must have some independence. 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 also p. 11 of their book).

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

9. 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), so as it where looking for grammaticalisation evidence from another perspective, 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. 237–39). 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. So far, evidence for cases of ‘regrammaticalisation’ or iconisation is not very abundant but examples are given in Plank (1979) and Ramat (1992).

10. I am here using ‘isomorphism’ in a somewhat narrower sense than is perhaps usual, following the definition given by Haiman (1980: 515–16) and other linguists quoted by Haiman (see also
Haiman, Kortmann and Ungerer, this volume). Isomorphism is usually seen as a mapping that satisfies “two fundamental structural properties”, i.e. the property of “being one-to-one” and the property of being “structurally consistent” (Holyoak and Thagard 1995: 29). Since here there are only two single elements involved, only the ‘one-to-one’ relationship applies. But as Haiman writes (1980, note 2), it is only “by virtue of this correspondence between individual signans and signatum [that it is] possible for the relationship of sets of signantia to mirror the relationship of sets of signata”.

11. See also Rubba (1994: 99), who notes that morphological boundness is related to semantic dependence, and the important work done by Bybee (e.g. 1985) on semantic ‘relevance’ as a factor in fusion and morpheme order in verbal forms.

12. For more information on this corpus, see Kytö 1991.

13. Concerning both (i) and (ii), it is quite generally accepted as part of the grammaticalisation hypothesis that the semantic and syntactic factors are seen as intertwined and the result of the same ‘motivation’. Heine et al (1991a: 168) write that “both grammaticalization and re-analysis are the result of one and the same strategy, namely the one which aims at expressing more ‘abstract’ concepts in terms of less ‘abstract’ ones”. We have seen that in the have to case the rebracketing or re-analysis is not linked to a semantic development.

14. Another example of the force of proximity can be found in Bolinger (1996: 16), his ‘oddment 14: proximity agreement’, where he discusses that the mere juxtaposition of a noun and a verb may lead to ungrammatical agreement, as in One of their sons were killed in Vietnam or The size of the fish are immaterial.

15. 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 from the point of view of the parameters of grammaticalisation distinguished by Lehmann (1985).

16. See e.g. King Horn (Hall 1901: 25), te lyue. In addition, there is also evidence of reduced to in forms where to is combined into one form with the infinitive, e.g. 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 Helsinki corpus.

17. 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.

18. Traugott (1993) notes that occurrences with expletive it are very rare in her corpora but that they do occur. It is interesting that all my ten informants, 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.
19. All informants except one agree that this 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 normal construction would be with a finite clause: Hij dreigde dat hij ....

20. There may be some problems here, which would need further investigation, since not every to before the infinitive may signal ‘purpose’. John Haiman, for instance, suggested that purpose is absent in To be or not to be .... However, I am of the opinion that to expresses some kind of shift here which is absent in the plain infinitive (or the ing-form for that matter). It is similar to what Duffley (1992: 106) writes on the difference between e.g. need followed by a bare or a to-in infinitive, “if the means of realizing the infinitive’s event are not felt to exist, then there is felt to be nothing real occupying the before-position [so a bare infinitive or ing-form would be appropriate]... If, on the other hand, the means are conceived as really or probably existing, then the speaker feels that there is something real occupying a before-position with respect to the infinitive’s event, and so uses to”. It is obvious that Hamlet is here considering possibilities (that is the question); should he be or not be? In other words, there is a ‘before’ and ‘after’.

21. The importance of to for the development of these new modals is also emphasised by Hopper and Traugott (1993: 81), where they write: “The 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.

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

References


Fischer, Olga. 1995. "The Distinction Between To and Bare Infinitival Complements in Late Middle English". *Diachronica* 12: 1–30.


