Markers of Futurity in Old High German and Old English: A Comparative Corpus-Based Study

Gabriele Diewald & Ilse Wischer

Abstract

This paper is a comparative corpus-based study of constructions that had the potential of marking future events in Old High German (OHG) and Old English (OE), i.e. modal constructions and those with be/become-verbs. Given the fact that both languages stem from a common source and probably had similar source lexemes for future grams, they nevertheless took diverging paths to develop a future tense, with werden in German and will/shall in English. The paper aims at comparing the earliest attestable stages of the two languages, i.e. Old High German and Old English to find out whether there are language internal differences with regard to the patterns of use of the possible source items. The database for our studies consists of OHG and OE text material dating from 790 to 1155, which we consider to be maximally comparable with respect to chronology, text type and content.

1. Introduction

Today, German and English, two West-Germanic languages, make use of different linguistic devices to refer to future events. In Present Day German (PDG), besides several less grammaticalized means, the werden & infinitive construction has been grammaticalized as a future marker,¹ as in example (1):

¹ The PDG werden & infinitive construction, beside its function as a future marker, has other uses in the domain of modality and evidentiality, which, however, will not be investigated here. Disregarding the question whether PDG should be attributed a full-fledged grammatical future category at all, it is taken for granted here that the werden & infinitive construction does behave like a future marker in some of its PDG uses, and, moreover, that it is the only serious candidate for this function in PDG.
Der Bund wird im laufenden Jahr rund 80 Milliarden Euro neue Schulden machen - so viel wie nie zuvor. (18.05.2010 Süddeutsche.de)

'The federal state will take on new debts of about € 80 billion in the running year - so many as never before.'

In Present Day English (PDE), on the other hand, this function is performed - among other linguistic devices, mainly by the modals shall and will, as in (2) and (3):

(2) However, we have grave concerns that this will have serious consequences for the security...financial assets. (27 May 2010 The Times)

(3) There we shall see one another as we really are, when all imperfection has been wiped away. (Roderick Strange, 13 November 2009 Times Online)

Another major difference between PDG and PDE is that while in German future marking is optional, in English it is to a large extent obligatory.

The question that derives from this observation is: Why did two closely related languages, which shared the same source items in Proto-West Germanic come to choose different source items for the grammaticalization of future markers and why did they follow different grammaticalization paths? As far as the diachronic lexical starting points are concerned, both languages – at least at first sight – seem to have had enough in common for developing similar future markers from cognate lexical sources. As is well known, in the earliest attested stages of the history of German and English both types of source lexemes, i.e. werdan on the one hand, and sculan/wellan on the other, were available as potential sources for future grams. The historical development of these forms in both languages, however, was divergent. Assuming that – beyond reasons of language contact, which are not the focus of this contribution – there are language internal reasons for this divergent development, in particular different patterns of usage and different frequencies of the source items (cf. Bybee 2010), we conducted a corpus-based study comparing the earliest attestable stages of the two languages, i.e. Old High German and Old English.

---

2 For reasons of simplicity we refer to the various formal occurrences of these source lexemes in OHG, MHG, OE and ME by citing the common etyma werdan, wellan and sculan.
The database we created for this study consists of OHG and OE text material dating from 790 to 1155, which is intended to be maximally comparable with respect to chronology, text type, content etc. The size of the corpus is about 80,000 words for each German and English. These texts were analysed according to the source lexemes that were available for future grams: *wellan, sculan, werdan*, and OE *beon*. This paper presents the results of this study and pinpoints several language internal factors for the diverging development of future markers in the history of German and English.

The paper is structured along the following lines: The next section describes our theoretical framework and gives a survey of earlier studies on the topic. Section 3 briefly comments on the empirical procedure. In section 4 our data will be presented. Section 5 discusses our findings, and finally, in section 6 we will summarize the results and draw some conclusions with respect to the question formulated in the beginning.

2. **Theoretical Background**

The rise of the German *werden*-construction and its use as a future marker has been the topic of a large number of studies since the nineteenth century. On the basis of an empirical analysis, Westvik (2000: 246ff.) suggests that its use as a future marker emerged in the first decades of the 14th century in the East Middle German and Upper High German dialect areas. In the course of the 14th and 15th centuries the construction expanded across the Upper High German and East Middle German dialects (cf. also Schmid 2000). This process obviously interrupted the grammaticalization of the modals *wollen* and *sollen*, which, at that time were well on their way to becoming future auxiliaries. In the 17th century, finally, the replacement of *wollen* and *sollen* by *werden* was completed (Bogner 1989: 82).

Only very recently were some studies on the subject published, which indicate a remarkable shift of interest insofar as they do no longer try to treat the rise of a grammatical marker in isolation, but to consider its language internal interaction with other grammatical markers on the one hand and comparative aspects with related

---

3 There are several highly divergent suggestions as to the exact date and origin of the *werden* & infinitive-construction; see Kleiner (1925), Saltveit (1962), Schieb (1981), Walther (1980), Leiss (1985), Schmid (2000) to quote only some of them. Some disagreement seems to be due to heavy differences in the quantity and quality of the diachronic language data used to propose a hypothesis. As Westvik (2000) provides an excellent survey on this, it is not necessary to report the research history anew. In addition to Westvik’s (2000) study, a further survey on that topic can be found in Harm (2001: 290ff.).
languages on the other. Among them is the study by Kotin (2003), who takes up the question of the rise of *werden* in connection with the development of the other German auxiliaries, and a very inspiring paper by Harm (2001), who – as far as we know – for the first time raises the issue of a comparative perspective with the aim of shedding light on the rise of German *werden* in opposition to the modals in other Germanic languages. Harm (2001: 294) also suggests taking into account the possibility of polygenesis of the *werden*-construction in several places in the linguistic area of German, which would render the notorious "where"-question redundant (further recent publications in the “environment” of our topic are Krämer 2005, Smirnova 2006, Hacke 2009, Rogler 2010).

In English, on the other hand, the cognate of German *werden* – Old and Middle English *weorþan* – was given up in the Late Middle English period. Instead, the *will/shall* & infinitive construction was grammaticalized as a future marker. Thus, from a seemingly comparable historical starting point, the languages made opposing choices. The history of *will* and *shall* as prime exponents of futurity in English is discussed in more or less detail in most handbooks on English historical syntax and in many treatments on the history of English modals (cf. e.g. Jespersen 1909, Mustanoja 1960, Kisbye 1971, Berndt 1982, Arnovick 1990, van Kemenade 1993). OE *weorþan* is mainly dealt with in connection with passive markers, although it is often claimed that passive constructions with *weorþan* in the present tense have a future connotation (Visser 1963-73: §1918; Mitchell 1985: §755; Kilpiö 1989: 61f.). Although the use of OE *beon* for future states of being or statements of eternal truth has attracted increased scholarly attention in recent years (cf. Kilpiö 1992, 1993, 1997; Lutz 2009; Wischer 2010), the development of *will* and *shall* into future grams in English is more or less taken for granted without considering other potential alternative sources in OE.

In our definition of a grammaticalized future marker we follow Bybee/Perkins/Pagliuca (1994: 244), who characterize a simple future as: “[a] prediction on the part of the speaker that the situation in the proposition, which refers to an event taking place after the moment of speech, will hold”. This is to say that a simple future is a grammatical marker by which the speaker asserts that the event expressed in the proposition will occur at a time yet to come. It indicates temporal distance from the moment of speech, whereby the direction on the time line is opposed to that of past tense markers. Or, to put it briefly, a simple future encodes a prediction.
Concerning the regularity of grammaticalization processes, the following facts have to be taken into account:

First, as we know from studies on grammaticalization paths (especially Bybee/Perkins/Pagliuca 1994: 253), the most frequent sources of future grams are movement verb constructions, followed by constructions with *be/become*-verbs, which correspond to the German *werden*-future and the OE use of *beon*. By contrast, the development of future auxiliaries from modals ("modal futures"), as in the case of English *will* and *shall*, is much less widely attested crosslinguistically. This means, while German seems not to have strayed from a well-trodden path concerning the development of its *werden*-future, the English case is rather exceptional.

Second, for the development of modal futures, the successive steps of semantic change leading from lexical sources with modal meaning to the notion of simple future are summarized as in Figure 1 (cf. Bybee/Pagliuca/Perkins 1991: 29 and Bybee/Perkins/Pagliuca 1994: 254-266):

Figure 1: Development of Modal Futures

- obligation (deontic modality) → desire (volitional modality) → root possibility (dispositional modality) → \{ \rightarrow INTENTION \rightarrow FUTURE \}

Third, the second grammaticalization path for future grams that needs to be taken into account here is the development from constructions with ‘be, become’, mentioned above. Bybee/Perkins/Pagliuca (1994: 263) treat them together with ‘have’/POSSESSION constructions as a subtype of obligation futures, although they suggest that there exist two alternative pathways for them, one path through obligation and a direct one with no intermediate obligation state, cf. Figure 2:

Figure 2: Development of ‘Be’, ‘Become’ Futures

- ‘be’, ‘become’, ‘have’/POSSESSION → OBLIGATION \{ \rightarrow INTENTION \rightarrow FUTURE \}
- ‘be’, ‘become’, ‘have’/POSSESSION → PREDESTINATION
3. Empirical Analysis (Methodology)

The principles concerning the design of the language corpora used and the basic considerations that have led to the building of our data-base are the comparability and the quantity of texts. Wherever possible we use maximally comparable texts in German and in English with respect to chronology, text type, content etc. The size of the English as well as the German corpus is about 80 000 words each. The length of the individual text sections varies according to availability.

The German corpus comprises the following text material dating from 790 to 1155 (see reference section for exact bibliographic data):

- **IS** = Isidor (ca. 790): about 5 100 word forms altogether, free translation from Latin;
- **TA** = Tatian (ca. 830): about 13 000 word forms, starting from the beginning of the text, gloss from Latin;
- **OT** = Otfrid von Weißenburg: "Das Evangelienbuch" (863/71): about 13 200 word forms, starting from the beginning of the text;
- **NO** = Notker von St. Gallen, translation of "Boethius: De consolatione philosophiae" (1025): about 15 000 OHG word forms from books 3 to 4 to match the section in the Helsinki corpus; the segment chosen comprises ca. 36 000 word forms altogether, roughly 21 000 of them Latin;
- **KA** = Kaiserchronik (1135/55): ca. 9 000 word forms, starting from the beginning of the text (pages 79-108 of the edition by Schröder);
- **SP** = Speculum ecclesiae deutsch (12th century): about 13 500 word forms starting from the beginning of the text, translation (free compilation) from Latin;
- **AL** = Alexanderlied (Vorauer Alexander) (ca. 1140/50): about 9 000 word forms, starting from the beginning of the Vorauer manuscript, translation from French (“Roman d’Alexandre”).

The German diachronic data can be accessed via [www.kali.uni-hannover.de](http://www.kali.uni-hannover.de).

The English corpus contains those sections from the following texts that are included in the Diachronic Part of the Helsinki Corpus of English Texts (for exact
bibliographic data see reference section). It comprises text segments dating from 880-1120:

- **AB** = Alfred’s Boethius (ca. 880): about 11 000 word forms, West Saxon dialect, translation from Latin;
- **AC** = Alfred’s Cura Pastoralis (ca. 885): about 18 000 word forms, West Saxon dialect, translation from Latin;
- **AO** = Alfred’s Orosius (ca. 885): about 9 000 word forms, West Saxon dialect, free translation from Latin;
- **WG** = West Saxon Gospels (ca. 990): about 10 000 word forms, West Saxon dialect, translation from Latin;
- **LG** = Lindisfarne Gospels (ca. 960): about 9 000 word forms, Northumbrian dialect, gloss from Latin;
- **C1** = Chronicle MS E (ca. 970-1050): about 9 000 word forms, West Saxon dialect;
- **C2** = Chronicle MS E (ca. 1070-1120): about 9 000 word forms, West Saxon dialect;
- **GG** = Gregory the Great (manuscript dating from ca. 1100; original from ca. 885): about 5 000 word forms, West Saxon dialect, translation from Latin.

These texts were analyzed completely. Additional language data as well as examples taken from earlier studies were used when it seemed illuminating for our argumentation (e.g. data from the 16th century).

4. **The Data: Source Lexemes in OHG and OE: Distribution and Frequency**

The source lexemes playing a role in the development of future marking in both languages are *wellan, sculan, werdan* and OE *beon*. These lexemes show remarkable differences with respect to frequency from the earliest stages. Tables 1a and 1b compare *werdan* in OHG and OE:

Table 1a: Frequency of *werdan* in the German corpus
The absolute numbers as well as the frequencies per 1,000 words show that German \textit{werdan} is much more frequent than its English cognate. However, other than in OHG, \textit{weorþan} has a rival form in OE, which partly overlaps with it functionally, namely \textit{beon}. As is known, in OE there are two competing verbs with the meaning 'to be, to exist': \textit{wesan} and \textit{beon}, which finally led to a suppletive paradigm for the verb ‘to be’ in Modern English. In Old English, \textit{wesan} usually refers to the real present situation, while \textit{beon} is used to denote general truths or future events.

Thus, while in OHG there is an opposition between the two lexemes \textit{werdan} and \textit{wesan}, in OE there are three lexemes, \textit{weorþan}, \textit{wesan} and \textit{beon}, to share the same functional space. In Table 2, a comparison between the OHG \textit{Tatian} and the OE \textit{Lindisfarne Gospels} rendering the same Latin text shows that where OHG uses \textit{werdan} or a form of \textit{wesan/sin} (‘to be’) to translate Latin futures, OE prefers \textit{beon}:

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|c|c|c|c|c|c|}
\hline
 & IS & TA & OT & NO & KA & SP & AL & total \\
\hline
\textit{werdan} \(\Sigma\) & 98 & 105 & 67 & 150 & 60 & 132 & 70 & 682 \\
\hline
frequency/1,000 words & 19.2 & 8.1 & 5.1 & 10.0 & 6.7 & 9.8 & 7.8 & 8.8 \\
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|c|c|c|c|c|c|}
\hline
 & AB & AC & AO & WG & LG & C1 & C2 & GG & total \\
\hline
\textit{weorþan} \(\Sigma\) & 33 & 54 & 44 & 15 & 50 & 20 & 49 & 9 & 274 \\
\hline
frequency/1,000 words & 3.0 & 3.0 & 4.9 & 1.5 & 5.6 & 2.2 & 5.4 & 1.8 & 3.4 \\
\hline
\end{tabular}
\end{table}

\footnote{The following detail from one of the OHG texts may support this observation: the OHG \textit{Tatian}, which altogether contains 765 verb types with 11 082 tokens, the verb \textit{werdan} (and \textit{furwerdan}) appears 362 times, and thus makes up 3.3\% of the total token frequency of verb forms (Sommer 1994:45, 84).}
Table 2: The rendering of Latin futures in the OHG Tatian and the OE Lindisfarne Gospels

<table>
<thead>
<tr>
<th>Latin</th>
<th>Tatian</th>
<th>Lindisfarne Gospels</th>
</tr>
</thead>
<tbody>
<tr>
<td>II.14 Et erit tibi gaudium...</td>
<td>Inti her ist thir gifehu... &amp; bið gefea þe...</td>
<td></td>
</tr>
<tr>
<td>et multi in ... gaudabant.</td>
<td>inti manage in ... mendent. &amp; monigo in ... biðon glæde.</td>
<td></td>
</tr>
<tr>
<td>II.15 Erit enim magnus coram...</td>
<td>Her ist uuârlieho mihhil fora... Bið forðon micel befora...</td>
<td></td>
</tr>
<tr>
<td>et spiritu sancto replebitur...</td>
<td>inti heilages geistes uuirdit gifullit... &amp; gaaste halge gefyllde bið...</td>
<td></td>
</tr>
<tr>
<td>II.20 Et ecce eris tacens...</td>
<td>Inti nu uuirdist suigenti... &amp; heono ðu bist suigendæ...</td>
<td></td>
</tr>
<tr>
<td>... quo haec fiant,</td>
<td>... in themo thisu uuerdent, ... of ðæm ðas geworðes...</td>
<td></td>
</tr>
<tr>
<td>... quae implebuntur in tempore...</td>
<td>... thiu thar gifultu uuerdent in ... ziti. ... ða ðe gefyllde biðon on tid...</td>
<td></td>
</tr>
<tr>
<td>III.35 ... et quod nascetur sanctum</td>
<td>thaz thar giboran uuirdit heilag, ... &amp; þætte acenned bið halig</td>
<td></td>
</tr>
<tr>
<td>vocabitur filius dei.</td>
<td>thaz uuirdit ginemnit gotes barn. bið geceid sunu godes.</td>
<td></td>
</tr>
<tr>
<td>III.45 ... quoniam perficientur ea quae...</td>
<td>... uuanta thiu uuerdent gifremitu thiu thar... forðon ðerh-geendad biðon ða ðaðe...</td>
<td></td>
</tr>
</tbody>
</table>

Table 3a shows that – according to the mere frequency of occurrences - in OE beon is much more frequent than weorpan, and this despite the fact that in Table 1b all occurrences of weorpan (including the past forms) were counted, while beon (in Table 3a) occurs only in the present tense and in the infinitive. If the past forms of weorpan are excluded from the frequency analysis as in Table 3b, an even higher discrepancy can be observed.

---

5 Luke I, 14ff. (Lindisfarne Gospels I: 14, 15, 20, 35, 45 [Skeat (ed.) 1874, pp. 17-23]; Tatian 2,6; 2,9; 3,7; 4,4 [Sievers (ed.) 1892/1966: 14-17].
Table 3a: Frequency of *beon* in the English corpus

<table>
<thead>
<tr>
<th></th>
<th>AB</th>
<th>AC</th>
<th>AO</th>
<th>WG</th>
<th>LG</th>
<th>C1</th>
<th>C2</th>
<th>GG</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>beon</em> Σ</td>
<td>191</td>
<td>248</td>
<td>9</td>
<td>35</td>
<td>57</td>
<td>12</td>
<td>3</td>
<td>31</td>
<td>586</td>
</tr>
<tr>
<td>frequency/1,000 words</td>
<td>17.4</td>
<td>13.8</td>
<td>1.0</td>
<td>3.5</td>
<td>6.3</td>
<td>1.3</td>
<td>0.3</td>
<td>6.2</td>
<td>7.3</td>
</tr>
</tbody>
</table>

Table 3b: Frequency of *beon* and *weorþan* excluding the past forms in the English corpus

<table>
<thead>
<tr>
<th></th>
<th>AB</th>
<th>AC</th>
<th>AO</th>
<th>WG</th>
<th>LG</th>
<th>C1</th>
<th>C2</th>
<th>GG</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>beon</em> Σ</td>
<td>191</td>
<td>248</td>
<td>9</td>
<td>35</td>
<td>57</td>
<td>12</td>
<td>3</td>
<td>31</td>
<td>586</td>
</tr>
<tr>
<td><em>weorþan</em> Σ</td>
<td>26</td>
<td>41</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>74</td>
</tr>
</tbody>
</table>

However, it must be noted that in both cases, for *beon* and *weorþan*, the dominant use is not that of marking futurity. Of the 191 *beon*-occurrences in Alfred’s *Boethius*, only 6 are used with a clear future reference⁶, cf. ex. (4), the others rather refer to statements of general truth, as in ex. (5). Similarly, of the 26 *weorþan*-occurrences in the same text, only 1 can be considered to have a clear future meaning, cf. ex. (6), the others rather express a current or general change of state, as in ex. (7). It is interesting to note that although the original meaning of *beon* was very similar to that of *weorþan*, namely ‘become’, in Old English it had already almost completely lost its inherent dynamic sense.

(4)  & for ðæm ege hi *beoð* simle swiðe earme. (AB, 117.28)
‘and because of that fear they will always be very miserable.’

(5)  *Fordy mon cwið be sumum goode þæt hit ne sie ful good, forðæm him bið hwæshwugu wana; …* (AB, 34.82.28)
‘Concerning any form of good we say that it is not perfect good, inasmuch as it lacks something; …’

(6)  *Gif þu þonne ænne stan tocliest, ne wyrð he næfre *gegadroð* swa he

⁶ Kilpiö (1989) also notices that *b*-auxiliaries in OE passives are rarely used for the future in the *Cura Pastoralis* and *Bede*. For the *Gospel of St Matthew* in the *Corpus Manuscript* and *Rushworth 1*, however, he notes that most of the instances of *beon* in passive constructions refer to the future. He concludes that the text type must have influenced the use of *beon/wesan*. For a similar conclusion cf. Bolze in this volume.
aer was; (AB, 34.92.28)

‘If therefore thou cleavest a stone it will never be united as it before was,’

(7) swa swa of þære sæ cymð þæt wæter innon ða eordan, & þær [{aferscað[};
cymð þonne up æt þæm æwelme, wyrd þonne to broce, þonne to ea, þonne andlang ea, ðod hit wyrd eft to sæ…. (AB, 34.86.22)

‘Even so from the sea the water makes its way into the earth, and there grows fresh; then it comes up at the spring, becomes a brook, then a river, then follows the course of the river until it comes again to the sea.’

Tables 4 and 5 are concerned with the modal source lexemes in OHG and OE. Tables 4a and 4b give the frequency counts for wellan, and Tables 5a and 5b those for sculan respectively.

Table 4a: Frequency of wellan in the German corpus

<table>
<thead>
<tr>
<th></th>
<th>IS</th>
<th>TA</th>
<th>OT</th>
<th>NO</th>
<th>KA</th>
<th>SP</th>
<th>AL</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>wellan Σ</td>
<td>2</td>
<td>15</td>
<td>30</td>
<td>62</td>
<td>36</td>
<td>30</td>
<td>54</td>
<td>229</td>
</tr>
<tr>
<td>frequency/1,000 words</td>
<td>0,4</td>
<td>1,2</td>
<td>2,3</td>
<td>4,1</td>
<td>4,0</td>
<td>2,2</td>
<td>6,0</td>
<td>2,9</td>
</tr>
</tbody>
</table>

Table 4b: Frequency of willan in the English corpus

<table>
<thead>
<tr>
<th></th>
<th>AB</th>
<th>CP</th>
<th>AO</th>
<th>WG</th>
<th>LG</th>
<th>C1</th>
<th>C2</th>
<th>GG</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>willan Σ</td>
<td>84</td>
<td>120</td>
<td>41</td>
<td>37</td>
<td>15</td>
<td>42</td>
<td>34</td>
<td>16</td>
<td>389</td>
</tr>
<tr>
<td>frequency/1,000 words</td>
<td>7,6</td>
<td>6,7</td>
<td>4,6</td>
<td>3,7</td>
<td>1,7</td>
<td>4,7</td>
<td>3,8</td>
<td>3,2</td>
<td>4,9</td>
</tr>
</tbody>
</table>

Table 5a: Frequency of sculan in the German corpus

<table>
<thead>
<tr>
<th></th>
<th>IS</th>
<th>TA</th>
<th>OT</th>
<th>NO</th>
<th>KA</th>
<th>SP</th>
<th>AL</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>sculan Σ</td>
<td>18</td>
<td>3</td>
<td>52</td>
<td>46</td>
<td>36</td>
<td>122</td>
<td>42</td>
<td>318</td>
</tr>
<tr>
<td>frequency/1,000 words</td>
<td>3,5</td>
<td>0,2</td>
<td>3,9</td>
<td>3,1</td>
<td>4,0</td>
<td>9,0</td>
<td>4,7</td>
<td>4,1</td>
</tr>
</tbody>
</table>

Table 5b: Frequency of sculan in the English corpus

<table>
<thead>
<tr>
<th></th>
<th>AB</th>
<th>CP</th>
<th>AO</th>
<th>WG</th>
<th>LG</th>
<th>C1</th>
<th>C2</th>
<th>GG</th>
<th>total</th>
</tr>
</thead>
</table>
These figures show that both modals were present with relative frequency in the corpora. While *wellan* is slightly more frequent in the English corpus compared to the German one, *sculan* predominates slightly in German, compared to the English corpus. This, however, may be due to the particularities of one text in OHG, the *Speculum Ecclesiae* (SP), with 122 instances of *sculan*. This text is a collection of sermons, which is highly instructive and therefore contains extremely many deontic uses of *sculan*.

According to their frequency of occurrence the potential source lexemes for future grams in both languages show the following distribution, cf. Table VI and Figure 3:

Table 6: Frequency of occurrence of potential source lexemes in OHG and OE

<table>
<thead>
<tr>
<th></th>
<th><em>werdan</em></th>
<th><em>wellan</em></th>
<th><em>sculan</em></th>
<th><em>beon</em></th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>OHG</td>
<td>682</td>
<td>229</td>
<td>318</td>
<td>---</td>
<td>1231</td>
</tr>
<tr>
<td>OE</td>
<td>274</td>
<td>389</td>
<td>205</td>
<td>586</td>
<td>1454</td>
</tr>
</tbody>
</table>

Already a look at the mere numerical relations points at different conditions for the grammaticalization of each individual source lexeme in the two languages. In the following section we will discuss our findings, including semantic and constructional factors for the development of future grams in English and German.
5. Discussion

Beyond frequency, there are further factors which are relevant for the divergent development of future markers in German and English, one is the degree of auxiliarisation, others are the inherent semantics and the constructional distribution of each item in contrast to the respective competitors. Since we cannot discuss all this at full length, we will concentrate on some of the most salient features.

5.1. The high degree of auxiliarisation of the OE modals

Earlier studies\(^7\) have proven that the German and English modals differ considerably as to the degree and the time of their auxiliarisation. While the English modals had already been highly auxiliarised in OE (cf. also Wischer 2006), the German modals never reached this degree (for details see Diewald 1999).

In our OE corpus, 84% of all *willan* tokens and 96% of *sculan* function unambiguously as auxiliary cf. examples (8) and (9).

(8) \[Hu ne meaht þu geson þæt ælc wyrt & ælc wudu wile weaxan on þæm lande selest þe him betst gerist \ldots\] (AB 91.13)
‘Canst thou not see that each plant and each tree will grow best in land that suits it best …’

(9) \[Ac þæt is swiðe dyslic & swiðe micel syn þæt mon þæs wenan scyle be Gode, oððe eft wenan þæt ænig þing ær him ware oððe betre donne he oððe him gelic.\] (AB 84.18)
‘But it is very great folly and sin to think thus of God, or to believe that anything existed before Him, or was better than or like unto Him.’

All of these examples are complemented by an infinitive, and in all of the cases it is the bare infinitive. Although the syntax of Old English does not allow a definite categorization of *willan* and *sculan* as auxiliaries since the word order is still rather

---

\(^7\) For English cf. Wischer (2006); for German cf. Diewald (1999).
flexible and the NICE properties\textsuperscript{8} are not applicable yet, the semantics, however, is often a clear indicator of their auxiliary status. As in examples (8) and (9) above, an interpretation in terms of their original lexical verb meaning does not make sense.

Another criterion that sets \textit{willan} and \textit{sculan} apart from other lexical verbs is the absence of non-finite forms. There is no single occurrence of a non-finite form of these two verbs in our text corpus, neither in the infinitive nor in the present or past participle. Furthermore, the negation of \textit{willan} differs from that of most lexical verbs. The proclitic negative particle \textit{ne} often merges with the verb stem giving forms like \textit{nylle} or \textit{nolde}. This is not possible with \textit{sculan} because of its initial [ʃ]. On the other hand, the initial [w] cannot be the only reason for the fusion of \textit{ne}, since it never occurs with \textit{weorþan} or \textit{wilnian}.

Thus, it should have become obvious that in their syntactic and semantic behaviour the majority of OE \textit{willan} and \textit{sculan} have diverged from the small rest of lexical \textit{willan} and \textit{sculan} and have adopted auxiliary status.

\textbf{5.2. Semantic and constructional aspects of the source items}

\textit{5.2.1. Modal futures}

As mentioned before, futures arising from modal sources have to pass a stage with intentional semantics, thus presupposing a conscious, intentional entity that exerts some influence on the event described by the proposition. In modals like \textit{will} this “modal source” is internal, which means it is co-referential with the subject of the sentence: the subject of the sentences is the source of the volition and – at the same time – the source of the intended action. In modals like \textit{shall}, on the other hand, the modal source is external, i.e. different from the subject of the sentence (see Diewald 1999: 93-111 for an extensive treatment).

Now, it follows from the definition of a simple future given in the beginning that to encode a purely temporal prediction, it must be devoid of any intentional meaning. For the development of future markers from modals this requires finally the abstraction of the semantic feature [+intentional].

The German modals \textit{wollen} and \textit{sollen} never reached the stage of encoding purely temporal notions. They never completely lost their intentional component, although they

\textsuperscript{8} Cf. Huddleston (1976: 333): N: they can be Negated by a following \textit{not/n’t}; I: in Interrogative clauses they undergo subject-verb inversion; C: they occur in post-verbal ellipses (Code) instead of \textit{do}; E: and
have reached a very high frequency since the OHG period (for details see Diewald 1999: 321-334).

It is interesting to note that as early as in the 16th century Veit Dietrich, who in an edition of his sermons and educational texts makes ample use of modal *wollen*- and *sollen*- constructions, always uses *werden* & infinitive for predictions, prophecies and so on, i.e. in cases which would naturally afford a simple future without modal connotations, as in example (10):

(10)  *wer an mich glaubt / d' wirt leben / ob er gleich stuerb / Vnd wer da lebt /vnd glaubet an mich / der wirt nimmermehr sterben.* (VD 79–21ff.)

He that believes in me, yet shall he live, though he would die. And whoever lives and believes in me shall never die.

Research on contemporary 16th century texts (especially on Luther) by Diewald/Habermann (2005) strongly supports the observation that in this period, rhetorically trained authors made a clear distinction between pure future (always encoded by *werdan*) and modal future uses.

Summarizing this brief excursion into the 16th century, we may say that the fact that the German modals semantically presuppose a modal source, i.e. an instigator of the state of obligation, volition etc., impeded their interpretation as a simple future tense marker. This means that, although these verbs, in the centuries we are talking about, could be used to express future time reference, they were not optimally suited to this function, and the interpretation of a modal as primarily referring to future time always remained a conversational implicature in the German modals.\(^9\)

Quite different from the situation in OHG, as early as in OE, *willan* and *sculan* in some uses had become quite close to future markers, expressing a mere prediction,\(^10\) cf. ex. (11) and (12):

---

\(^9\) As Harm (2001:297, 299) rightly points out, this, of course, does not imply that modals in general are bad candidates for future grams (in the light of the findings of many grammaticalization studies and the development of English this obviously cannot be true); it only shows that in the case of German they were not good enough, as there was a better candidate for this function.

‘Now, thou canst perceive how heavy and how difficult it is to explain all this; but nevertheless I will set to work to teach thee somewhat, for I have noted that it is a powerful remedy for thy sorrow if thou understandest, aught of this, though it be a long task for me to teach thee.’

‘he said that the doctrine of the Holy Spirit will flee falsehood.’

Although, even up to today, will and shall have not completely lost their modal colouring, will/shall + infinitive constructions “are the closest approximations to a colourless, neutral future” (Greenbaum/Quirk 1990: 57), and they could be used in such functions even in the earliest attested texts.

A check on the expressions of what comes closest to a mere future in the section of Alfred’s Boethius revealed the following frequency:

- willan (11)
- present indicative (5)
- beon (4)
- sculan (3)
- present subjunctive (1)
- weorpan (1)

Here willan clearly dominates with 44 % of all future expressions.

Although in Middle English, futurity is still much more frequently referred to by the simple present than in Modern English, will and shall are gaining more and more ground as future markers. Now shall becomes the usual means of indicating futurity, while will continues to carry a strong volitional meaning until the end of the Middle English period. In late Modern English, however, will had become the most frequently-used future marker, particularly in less formal registers.
5.2.2. ‘Be, become’ futures

German *werdan*, unfortunately, is not included in the language sample used by Bybee/Perkins/Pagliuca (1994). There is, however, one passage (262ff.), where German *werden* is mentioned as possibly belonging into a subclass of source lexemes, of "'be' and 'have' constructions" which follow "a direct path [towards future] with no intermediate obligation stage" (263), i.e. *werden* is classified together with the Latin -b- future, which derives from an Indo-European 'be'-verb. For such constructions the authors posit the following path *(ibid.):*

‘be’, ‘become’, ‘have’/POSSESSION > PREDESTINATION > INTENTION > FUTURE

Apart from the fact that we do not believe that an INTENTION-stage is necessary on the path to futures (cf. also Ziegeler 2006; Hilpert 2007: 38; Wischer 2008), or whether a sense of PREDESTINATION must precede the future meaning, we would argue that ‘become’-sources should be treated separately from ‘be’- and ‘have’/POSSESSION-sources because of their inherent aspectual ingressive sense, which is lacking in the other two.

In his very inspiring investigation, Fritz (2000: 43) describes the semantic structure of *werdan* as consisting of two parts, i.e. as the contrast between an original state and a final state whereby the focus is on the transformative moment, i.e. on the feature of a change of state; he also points to the fact that *werdan*, as opposed to the modals, does not imply an instigator (or source) of the change of state.

Consequently, *werdan* never had an intentional meaning that would have to be "bleached out" on its way to becoming a future. Thus, for ‘become’ sources we would suggest the following semantic path (cf. also Diewald & Habermann 2005: 237ff.):

‘become’/INGRESSIVE > FUTURE

Due to its lack of semantic restrictions, *werdan* has always been compatible with subjects and predicative elements of any kind, which make it a good candidate for auxiliarization processes.

Furthermore, it can be observed that throughout the history of German, *werdan* has displayed a high constructional variability. It has always been used simultaneously in a range of syntactic functions spanning from full verb via copula to auxiliary. The
construction types in which *werdan* occurred in our OHG corpus are summarized in Table 7a:  

Table 7a: Construction types of OHG *werdan*

<table>
<thead>
<tr>
<th></th>
<th>IS</th>
<th>TA</th>
<th>OT</th>
<th>NO</th>
<th>KA</th>
<th>SP</th>
<th>AL</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Past P</td>
<td>86</td>
<td>73</td>
<td>27</td>
<td>100</td>
<td>23</td>
<td>91</td>
<td>41</td>
<td>441</td>
</tr>
<tr>
<td>Copula</td>
<td>6</td>
<td>9</td>
<td>19</td>
<td>14</td>
<td>23</td>
<td>30</td>
<td>18</td>
<td>119</td>
</tr>
<tr>
<td>Intrans.</td>
<td>6</td>
<td>9</td>
<td>20</td>
<td>33</td>
<td>14</td>
<td>6</td>
<td>11</td>
<td>111</td>
</tr>
<tr>
<td>+ Pres P</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>+ Inf</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Σ</td>
<td>98</td>
<td>105</td>
<td>67</td>
<td>150</td>
<td>60</td>
<td>132</td>
<td>70</td>
<td>682</td>
</tr>
</tbody>
</table>

The construction types for OE *weorþan* and *beon* are listed in Tables 7b and c:  

Table 7b: Construction types of OE *weorþan*

<table>
<thead>
<tr>
<th></th>
<th>AB</th>
<th>CP</th>
<th>AO</th>
<th>WG</th>
<th>LG</th>
<th>C1</th>
<th>C2</th>
<th>GG</th>
<th>Σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Past P</td>
<td>11</td>
<td>32</td>
<td>27</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>31</td>
<td>5</td>
<td>118</td>
</tr>
<tr>
<td>Copula</td>
<td>18</td>
<td>17</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>7</td>
<td>3</td>
<td>63</td>
</tr>
<tr>
<td>Intrans.</td>
<td>4</td>
<td>3</td>
<td>11</td>
<td>10</td>
<td>47</td>
<td>4</td>
<td>11</td>
<td>1</td>
<td>91</td>
</tr>
<tr>
<td>+ Pres P</td>
<td>-</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>Σ</td>
<td>33</td>
<td>54</td>
<td>44</td>
<td>15</td>
<td>50</td>
<td>20</td>
<td>49</td>
<td>9</td>
<td>274</td>
</tr>
</tbody>
</table>

Table 7c: Construction types of OE *beon*

<table>
<thead>
<tr>
<th></th>
<th>AB</th>
<th>CP</th>
<th>AO</th>
<th>WG</th>
<th>LG</th>
<th>C1</th>
<th>C2</th>
<th>GG</th>
<th>Σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Past P</td>
<td>38</td>
<td>114</td>
<td>2</td>
<td>13</td>
<td>34</td>
<td>2</td>
<td>-</td>
<td>19</td>
<td>222</td>
</tr>
<tr>
<td>Copula</td>
<td>126</td>
<td>115</td>
<td>5</td>
<td>22</td>
<td>13</td>
<td>10</td>
<td>3</td>
<td>10</td>
<td>304</td>
</tr>
<tr>
<td>Intrans.</td>
<td>24</td>
<td>15</td>
<td>-</td>
<td>-</td>
<td>6</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>46</td>
</tr>
<tr>
<td>+ Pres P</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>-</td>
<td>4</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>9</td>
</tr>
</tbody>
</table>

---

11 Harm (2001: 298) gives a fine-grained account of the various uses of *werdan*, which, among other things, includes *werdan* in so-called "Funktionsverbgefügen" (phrasal verbs), that is, in combination with prepositional phrases (e.g. MHG. *ze leide werden* 'to turn into grief (for somebody)'), see Harm 2001: 298). Furthermore, OHG *werdan* is also used in possessive constructions, as in Tatian (2, 2): *Inti ni uuard in sun lat.: et non erat illis filius* 'and they did not have a son'. For our purpose, however, it is sufficient to distinguish the four classes in table 7a.
For a comparison of the lexical and periphrastic construction types of *werdan/beon* in OHG and OE see Table 8:

Table 8: Comparison of the lexical and periphrastic construction types in OHG and OE

<table>
<thead>
<tr>
<th>Construction Type</th>
<th>OHG <em>werdan</em></th>
<th>OE <em>weorþan</em></th>
<th>OE <em>beon</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intransitive</td>
<td>16 %</td>
<td>34 %</td>
<td>7 %</td>
</tr>
<tr>
<td>Copula</td>
<td>17 %</td>
<td>22 %</td>
<td>52 %</td>
</tr>
<tr>
<td>Σ Lexical Verb</td>
<td>33 %</td>
<td>56 %</td>
<td>59 %</td>
</tr>
<tr>
<td>+ PastP</td>
<td>65 %</td>
<td>43 %</td>
<td>38 %</td>
</tr>
<tr>
<td>+ PresP</td>
<td>1 %</td>
<td>1 %</td>
<td>2 %</td>
</tr>
<tr>
<td>+ to-Inf</td>
<td>-</td>
<td>-</td>
<td>1 %</td>
</tr>
<tr>
<td>+ bare Inf</td>
<td>0 %</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Σ Potential Aux</td>
<td>66 %</td>
<td>44 %</td>
<td>41 %</td>
</tr>
</tbody>
</table>

As can be seen, the potential auxiliary uses of *werdan* were much higher in OHG than in OE. It even occurred - though quite rarely - with a bare infinitive in OHG, cf. ex. (13).

(13) *Súnter thaz giscrib min wirdit bêzra sin, búažent síno gúati thio mino míssodati.* (Otfrid V 25,45)
Sondern das Geschreib mein wird besser sein, büßen seine (des Lesers) Qualitäten meine Versäumnisse.
But my writing should be better, his (the reader's) qualities atoning for my failures.

6. Summary
Our comparative analysis of potential source lexemes for future grams in an OHG and an OE text corpus provided evidence that the two languages display remarkable differences with respect to the relevant items even in their earliest attested stages.

These differences concern the frequency of usage, the value (in the sense of de Saussure) of each item, i.e. the relations to its closest competitors / neighbours in the word field, semantic features, and structural and constructional factors.

While from the perspective of cross-linguistic relevance (and frequency) of grammaticalization paths for futures, the be/become futures surpass the modal futures and thus should win out in a situation where both sources are present in a language – as werden did in German, Modern English opted for the modal future due to disadvantages of OE werdan in terms of frequency and competition by beon.

This is a clear indication that general grammaticalization paths, invaluable as they are for an overall evaluation of probabilities, have to be checked carefully against the respective internal linguistic situation in a given synchronic stage of a language under investigation.

We have shown that these internal factors provide a powerful motive for the development “against” the more frequent universal pathways, and may promote a divergent development of closely related languages with almost identical starting conditions.

In this context, however, it is necessary to look at sociolinguistic factors as well, in particular at language contact, which we did not do in this paper. So, this remains a task for future work.

As to possible grammaticalization paths for futures, we suggest, on the basis of our findings, to refine previous suggestions and to distinguish an additional grammaticalization path for become-sources of futures, as set apart from be/have-sources.

**BIBLIOGRAPHY**

**Primary Sources:**

OE


OHG


References


1. Introduction (by Diewald, Gabriele)
2. "haitan in Gothic and Old English (by Cloutier, Robert)
3. Incipient Grammaticalisation: Sources of passive constructions in Old High German and Old English (by Mailhammer, Robert)
4. Passive auxiliaries in English and German: Decline versus grammaticalisation of bounded language use (by Petre, Peter)
5. Causative habban in Old English: Tracing the Development of a Budding Construction (by. Language Contact and the Origins of the Germanic Languages (Routledge Studies in Linguistics Book 13). Peter Schrijver. 4.0 out of 5 stars 1. Kindle Edition. $48.40. New Comparative Grammar of Greek and Latin. Andrew L Sihler. 3.8 out of 5 stars 13. This is definitely not a text for beginners -- readers should have a good background in Comparative Indo-European Linguistics before tackling this book. Those who have the patience to work their way through the book will be immeasurably rewarded (do not skip the endnotes!). Read more.