Onomasiological Profiles of Old English Texts

Analysing the Vocabulary of Beowulf, Andreas and the Old English Martyrology through Linguistic Linked Data

Thijs Porck | ORCID: 0000-0002-8537-6909

Leiden University Centre for the Arts in Society, Leiden University
m.h.porck@hum.leidenuniv.nl

Abstract

This article discusses proof-of-concept research into the structure of the vocabularies of three Old English texts, *Beowulf*, *Andreas* and the *Old English Martyrology*. With the help of the Web application Evoke, which makes *A Thesaurus of Old English* (TOE) available in Linguistic Linked Data form, the words that occur in these three texts have been tagged within the existing onomasiological structure of TOE. This tagging process has resulted in prototypes of ‘textual thesauri’ for each of the three texts; such thesauri allow researchers to analyse the ‘onomasiological profile’ of a text, using the statistical tools that are built into Evoke. Since the same overarching structure has been used for all three texts, these texts can now be compared on an onomasiological level. As the article demonstrates, this comparative approach gives rise to novel research questions, as new and distinctive patterns of vocabulary use come to the surface. The semantic fields discussed include “War” and “Animals”.

Keywords

onomasiology – Old English – Evoke – Digital Humanities – Linguistic Linked Data – *Beowulf* – *Andreas* – *Old English Martyrology*

1 Introduction

How many Old English words that express the concept of ‘warrior’ occur in *Beowulf*? Are the text’s *hapax legomena* confined to specific semantic fields or are they distributed evenly across the poem’s semantic range? How ambiguous (or: polysemous) are the words used in *Beowulf*? And how does this vocabulary use compare to that in other Old English texts, such as *Andreas* and the *Old English Martyrology*? These are relatively basic questions and it is not difficult to imagine how the answers to these questions might provide information about the particular style, nature and preoccupations of these Old English texts. Yet, there is currently no simple way of answering these questions other than manually leafing through published glossaries of these texts with specific questions in mind. This article introduces proof-of-concept research into the ‘onomasiological profiles’ of these three Old English texts that digitally facilitates answering these and other questions relating to the structure of the vocabularies of these texts.

After briefly covering some of the theoretical background and introducing the concept of an onomasiological profile, this article describes how the Web application Evoke (Stolk, 2018) was used to create prototypes of ‘textual thesauri’ for each of the three texts. Next, the vocabulary of these three Old English texts will be compared on an onomasiological level, answering such basic questions as the ones above as well as demonstrating how comparing the onomasiological profiles of different texts may give rise to novel research questions, as new and distinctive patterns of vocabulary use come to the surface.

2 Words and Worldviews: Cultural Conceptualisations and Onomasiological Profiles

Language can provide an insight into the culture of its speakers; this notion is well established in the fields of ethnolinguistics, which studies the relationship between language and culture, and cognitive linguistics, which studies how language is used to conceptualise ideas. Ethnolinguists, such as Wilhelm von Humboldt, Franz Boas and Edward Sapir, held that language is “the means by which men create their conception, understanding and values of objective reality” (for a summary, see Basilius, 1952). Cognitive linguists confirm the disposition that language can be studied not only as a medium of communication, but also as a means for speakers to conceptualise the world around them. The cognitive linguist Farzad Sharifian (2011), for instance, has shown that a

language reflects the thoughts and mental patterns of its speakers and, as such, each language can be analysed as a “collective memory bank for cultural conceptualisations” (39).

Although links have been proposed between culture and almost every aspect of language, including morphology, syntax and phonology (Mathiot, 1979; Palmer, 1996), vocabulary is generally regarded as the most direct link between a language and the worldview of its speakers (Wierzbicka, 1997; Sharifian, 2011):

Vocabulary is a very sensitive index of the culture of a people and changes of the meaning, loss of old words, the creation and borrowing of new ones are all dependent on the history of culture itself. Languages differ widely in the nature of their vocabularies. Distinctions which seem inevitable to us may be utterly ignored in languages which reflect an entirely different type of culture, while these in turn insist on distinctions which are all but unintelligible to us.

SAPIR, 1951: 27

The theoretical validity of this approach has been questioned in the past, since children can conceptualise ‘food’ before they acquire language, speakers can distinguish between pink and red, even if their language uses only one word for both, and we are often at a loss for words to express our heart-felt emotions (Hall, 2007: 12). Be that as it may, analyzing vocabulary, in particular how it is structured across semantic fields and through conceptual metaphors, is a well-established way of gathering insights into the culture of a language community (e.g., Goddard, 2015).

The study of Old English vocabulary has often been the starting point for discussions of cultural phenomena in Anglo-Saxon England (Strite, 1989; see also Díaz Vera, 2002). At a basic level, the absence or presence of words for particular concepts can have significance; the absence of Old English words for the temporal concepts ‘second’ and ‘minute’, for instance, may indicate that such precise levels of time measurement were not necessary in early medieval England (cf. Kopár, 2010).¹ Another important concept is ‘cultural elaboration’, the notion that the more culturally salient a concept is, the better

1 Hall (2007: 13) notes “People can of course conceive of things for which they lack words, and the absence of a word does not prove the absence of corresponding concepts. However, it is reasonable to suppose *a priori* that the distribution of words in a lexicon attests to the relative cultural salience of the concepts which they denote, with absences at least suggesting low salience.”

it will be represented in a language's lexicon, as demonstrated by the fact that the Hanunóo language of the Philippines has ninety different lexemes for 'rice' and that Arabic abounds in expressions for 'stone', 'camel', 'sword' and 'snake' (Wierzbicka, 1997: 10). Instances of cultural elaboration of vocabulary can also be found in the Old English lexicon, where the disproportionate size of the semantic field 'warrior' reflects the preoccupation with warfare of a sizeable portion of the Anglo-Saxon literary record (Bremmer, 2006: 75–76). Likewise, the relatively high number of Old English lexical items for 'bondage, slavery' suggests the importance and presence of a sizable group of unfree people in early medieval England (Momma, 2003: 80). Similarly, Rolf H. Bremmer Jr has noted that the multiple expressions for the instrument of Christ's execution reveal the importance of the Cross for the Anglo-Saxons: "A child that is loved has many names" (Bremmer, 2010: 231). In addition to these purely quantitative statements, studying the underlying structure within and overlap across Old English semantic fields has helped scholars gain an understanding of such diverse cultural concepts as emotions, the past, and the life course in early medieval England (Fabiszak, 2002; Taranu, 2013; Izdebska, 2022). The analysis by Caroline Gevaert (2002) of Old English expressions for anger is a case in point; she revealed that Anglo-Saxons associated anger with sadness, insanity and unkindness, as reflected in the polysemous Old English words *unrōt* 'sad, angry', *wōd* 'mad, raging, senseless' and *unmilts* 'severity, anger'.

A crucial development in the history of Old English semantics was the publication of *A Thesaurus of Old English* (*TOE*; see also the contribution by Roberts in this special issue). *TOE* provides an overview of Old English words organised by semantic principles and allows its users to find out which words were available for a particular notion or concept. It was first published in 1995, with a second imprint in 2000; since 2005 an updated version of *TOE* has been available online. The semantic classification system of *TOE* is based on the semantic categories developed for Glasgow University's *Historical Thesaurus of English* project (*HTE*). *TOE* uses eighteen major categories to classify Old English words, including "The Physical World", "Life and Death", "Work" and "Leisure". These categories are further divided into sub-categories, in a hierarchy moving from the most general to the most specific meaning. This semantic hierarchy is reflected in the numbering of the categories:

- 13 Peace, tranquillity
- 13.02 War
- 13.02.10 The military, soldiers
- 13.02.10.01 A man, warrior

By 2005, *TOE* covered 33,976 Old English lexemes and some of these lexemes are tied to more than one category (see Kay, 2005: 38). The noun *wulf*, for example, is included in two separate categories: “02.06.03.01.15. Wild animal: Wolf” and “08.01.03.09.11|03 Hardheartedness, cruelty, severity: A cruel man”. In total, *TOE* currently has 51,480 senses (that is: lexemes tied to categories).² *TOE* covers the vocabulary used in all Old English texts and does not distinguish between, for example, chronological phases or dialects of Old English. It does mark some usage features of its lexemes through its distribution flags “o” (rare), “p” (exclusively used in poetry), “g” (exclusively used in glosses) and “q” (questionable). Additional mark-up of the *TOE* data, e.g., whether a lexeme or sense is restricted to a particular text or author, has only recently become possible through the Web application Evoke (see the contribution by Stolk in this special issue and also below).

Using Evoke, the vocabulary of three Old English texts has been labelled within *TOE*’s onomasiological structure. This has produced, in effect, three ‘textual thesauri’ that enable scholars to look into the onomasiological profiles of these texts. Each text’s onomasiological profile provides insight into various aspects of vocabulary use, including (1) which (and how many) words were used in the text for a particular concept; (2) for which semantic categories these texts use the greatest variety of words; (3) in which semantic categories texts utilize vocabulary that is exclusively found in that text (i.e., *hapax legomena*); (4) the ambiguity of the vocabulary used in these texts (i.e., how much of its diction was polysemous); and (5) the specificity of the vocabulary used in these texts (i.e., how general or specific were the concepts potentially expressed by the words used?).

The first text that was selected for this proof-of-concept research was *Beowulf* (ed. Fulk, Bjork, and Niles, 2008), the vocabulary of which has drawn the attention of various scholars.³ Arthur Brodeur’s classic *The Art of Beowulf* (1959) remains one of the most influential studies of the poem’s vocabulary and his appraisal of the poet’s diction as “more vigorous, stately, and beautiful than that of any other Old English poem” (38) has met with wide approval. A conventional approach to the vocabulary of *Beowulf* is to focus on the headwords that belong to a given semantic field, e.g., ‘sea’, ‘hall’ or ‘warrior’, and discuss how the poet’s treatment of this semantic field differs from other texts in the Old English corpus (see, e.g., Brady, 1979; 1983). As will be demonstrated in

2 Tally based on analysis of the Linguistic Linked Data form of the *TOE* dataset made available in Evoke by September 2021.

3 For an overview of scholarship on the poem’s diction, see O’Brien O’Keeffe (1997).

this article, this approach can be facilitated with the establishment of textual thesauri through Evoke, which additionally allows one to analyse the distribution of a text's vocabulary across all semantic domains.

For the sake of a comparative approach, the vocabularies of two further Old English texts were also labelled within the onomasiological structure of *TOE: Andreas* (ed. North and Bintley, 2016) and the *Old English Martyrology* (ed. Rauer, 2013). The first is a poem that stands out for its close connection to *Beowulf*, as noted by Richard North and Michael D. J. Bintley (2016): “verbal parallels between *Andreas* and *Beowulf* are high in number, and there is a persistent likeness between these poems” (63). A comparison of the onomasiological profiles of both texts may shed new light on their possible relationship (e.g., how many words do both poems share and how is their shared vocabulary distributed across semantic fields?). Moreover, a comparative approach may test such impressionistic statements about the vocabulary of *Beowulf* as the following: “the style of *Beowulf* is consistent with what is seen in Old English poetry more generally, though the quality of expression here often strikes one as exceptional in its variety and inventiveness” (Fulk, Bjork, and Niles, 2008: cxvi). As demonstrated below, the statistical information derived from Evoke will allow scholars to quantify and further elaborate on such impressionistic qualifications. The third text to be included in this research is a prose text, the *Old English Martyrology*. In part, this text was selected because it has a vocabulary that is comparable to that of *Beowulf* and *Andreas* in terms of size, date and dialectal features. Like *Andreas* (North and Bintley, 2016: 46–48), the *Old English Martyrology* has Anglian dialect features (notably Mercian) and is dated to the ninth century (Rauer, 2013: 3–5), while *Beowulf* is probably also of Anglian origin though likely composed slightly before the onset of the ninth century (see various contributions to Neidorf, 2014). However, the vocabulary of *The Old English Martyrology* is also worth studying in its own right. As Christine Rauer (2013) has noted, the *Old English Martyrology* is “one of the longest and most important prose texts written in Anglo-Saxon England” (1) and “[o]ne of the most interesting linguistic aspects of the *Old English Martyrology* is its vocabulary” (7; see also Rauer, 2016).

3 Methodology: Tagging *Beowulf*, *Andreas* and the *Old English Martyrology* in *TOE* through Linguistic Linked Data in Evoke

Evoke, developed by Sander Stolk (2018), is a Web application that allows users to easily browse, filter, visualize and expand on thesauri that have been made available in the form of Linguistic Linked Data (LLD). The LLD form allows users to link their own information (e.g., custom tags) to existing datasets and,

in effect, create their own enriched and/or specialized onomasiological datasets. Using the OntoLex-Lemon vocabulary, Stolk was able to automatically port the dataset of *TOE* into LLD form (henceforth: *TOE*-LLD), which allows for this dataset to be enriched with a user's own tags (see Stolk, 2019). For the purpose of this article, the headwords listed in the published glossaries of *Beowulf*, *Andreas* and the *Old English Martyrology* were tagged in the *TOE*-LLD dataset, along with an additional tag “hapax” to indicate *hapax legomena*.⁴ Figure 1 shows *TOE* category “13.02.10.01 A man, warrior” as visualized in Evoke, along with the custom tags that belong to the textual thesauri of *Beowulf*, *Andreas* and the *Old English Martyrology*.

FIGURE 1 *TOE* category “13.02.10.01 A man, warrior” with tags from the *Beowulf Thesaurus v1.0*, *Andreas Thesaurus v1.0* and *Old English Martyrology Thesaurus v1.0*. The distribution flags “poetry”, “rare” and “questionable” are part of the original *TOE* dataset, the tag ‘hapax’ is an additional custom tag, indicating *hapax legomena*

4 The tagging process was performed by student-assistant Lucas Gahrman, under supervision of the author, with the help of the Alignment Tool, developed by Sander Stolk (see the contribution by Stolk to this special issue).

The technicalities of the automatic porting process that made *TOE* available in LLD form are outlined in Stolk (2019) and go beyond the scope of this article. However, one technical aspect deserves mention since it affects the way *TOE*-LLD has been used to create the textual thesauri of *Beowulf*, *Andreas* and the *Old English Martyrology*. Crucially, the OntoLex-Lemon vocabulary distinguishes between a *LexicalEntry* and a *LexicalSense* – a *LexicalEntry* corresponds to a headword in a dictionary (e.g., the Old English word *wulf*), whereas a *LexicalSense* corresponds to a lexeme within one of the *TOE* categories (e.g., *wulf* in “02.06.03.01.15. Wild animal: Wolf” or *wulf* in “08.01.03.09.11|03 Hardheartedness, cruelty, severity: A cruel man”). Naturally, one *LexicalEntry* (henceforth: entry) can have multiple *LexicalSenses* (henceforth: senses), in the case of polysemous words. Evoke allows tagging at entry or sense level: if a tag is added at entry level, all related senses will automatically show this tag as well; if a sense is tagged, this tag is not automatically carried over to the entry. The textual thesauri of *Beowulf*, *Andreas* and the *Old English Martyrology* were made by tagging at entry level. This decision was made because it is not always clear in what exact sense a word is used in a particular text. Moreover, an author may have used a particular word because of the variety of senses that word might ‘evoke’; in other words, the tags in these textual thesauri mark the senses that words used in these Old English texts could potentially have.⁵

Since the original *TOE* dataset consisted only of senses (lexemes tied to categories), the entries in *TOE*-LLD have been automatically generated during the porting process (see Stolk, 2019): if a sense matched another sense in spelling, part of speech and *TOE* distribution flags (i.e. “o”, “p”, “q” and/or “g”), they were assumed to belong to the same entry. While this automatic procedure has been mostly successful, the research conducted for this article has revealed a number of cases where two senses that are clearly related were not linked to the same entry.⁶ For instance, *sæne* in *TOE* category “05.12.03.02|04 Slowness: Slow” is not linked to the same entry as *sæne/sene* in *TOE* category “11.06|16 Disinclination to act, listlessness: Slow, dull, sluggish, inactive”, due to the

5 In the case of *wulf*, which appears in *Beowulf* and the *Old English Martyrology*, the tag at entry level means that both the sense *wulf* in “02.06.03.01.15. Wild animal: Wolf” as well as *wulf* in “08.01.03.09.11|03 Hardheartedness, cruelty, severity: A cruel man” show tags as if they appear in these two texts, even though the latter sense is unlikely to have been intended by the authors of these texts.

6 More rarely, lexemes that are unrelated are merged under one *LexicalEntry*, because they share the same spelling and distribution flags in the original *TOE* dataset, as is the case for *dung* ‘dung, manure’ in *TOE* category “04.02.04.02.01.01|02 To manure, dung: Dung” and *dung* ‘dungeon, prison’ in *TOE* category “14.05.08.01|01.01 Prison, confinement, durance: A prison, jail: A dungeon”. These two lexemes are merged because they both have the flags “o” and “g” in the original *TOE* dataset; incidentally, the flag “g” for *dung* ‘dungeon, prison’ is incorrect and should be “p”, since it only occurs in the poem *Andreas*, see *DOE*, s.v. **dung*.

differing spellings of the senses in the original *TOE* dataset, separate entries have been generated for each in the automatic porting process. Cases such as these indicate that the *TOE*-LLD dataset will need to be manually checked in order for analyses to be more fine-grained; moreover, they imply that the number of lexical entries in the textual thesauri discussed in this article may not always correspond to the number of headwords in the glossaries of the Old English texts.

In fact, there were various discrepancies between the headwords in the published glossaries of *Beowulf*, *Andreas* and the *Old English Martyrology* and the entries in *TOE*-LLD, due to differing editorial choices. As a result, five principles for the tagging of these texts have been established, which are outlined below.

3.1 *Concerning Optional Prefixes*

For many Old English verbs, the prefix *ge-* is optional and its presence rarely leads to semantic differentiation. *TOE* does not appear to have been consistent in its treatment of verbs with or without the prefix *ge-* and, therefore, the automatic porting process (Stolk, 2019) has occasionally resulted in separate entries for, e.g., *libban*, *(ge)libban* and *gelibban*. In dealing with these prefixed verbs, the following practice was devised:

- *libban* as a headword in the glossary – both *libban* and *(ge)libban* in *TOE*-LLD are tagged (but not *gelibban*)
- *gelibban* as a headword in the glossary – both *(ge)libban* and *gelibban* in *TOE*-LLD are tagged (but not *libban*)
- *(ge-)libban* as a headword in the glossary – *libban*, *(ge)libban* and *gelibban* in *TOE*-LLD are all tagged

This tagging principle naturally affects how the number of tagged entries in the textual thesauri correspond to the number of headwords in the glossary.

3.2 *Concerning Phrasal Verbs and Idiomatic Phrases in TOE*

While *TOE* includes phrasal verbs (e.g., *becuman on/becuman to*) and idiomatic phrases (*forþ gefremman*) in its dataset, most published glossaries do not list these as headwords. Wherever it was possible to identify the phrasal verbs and idiomatic phrases in the texts, these entries have been tagged in *TOE*-LLD.

3.3 *Multiple Entries in TOE-LLD Due to Spelling and Morphological Variants*

Due to the automatic porting process, *TOE*-LLD occasionally features separate entries for spelling variants (e.g., *eallgrene* and *ælegrene/eallgrene*) and morphological variants (e.g., *geardagas* and *geardagum*). If a headword in a glossary corresponds to such multiple entries, all of the corresponding entries have been tagged, regardless of the spelling or morphological form used in the text.

3.4 *Morphological Discrepancies between Glossaries and TOE*

A number of differences between the published glossaries and the *TOE* dataset concern the morphological form of single entries. For instance, the *Andreas* poet's use of the word "geaclod" [terrified] in line 805 triggered the inclusion of the headword *ge-aclian* 'to terrify' in the *Andreas* glossary (ed. North and Bintley, 2016), but this could only be matched with the past participle form *geaclod* 'terrified' in the *TOE* dataset. Another discrepancy concerns the preterite-present verbs. In most glossaries these occur as infinitive forms, but in *TOE* these verbs are found in their first/third person present tense form (e.g., *deag* for *dugan*; *dearr* for *duran*; *mot* for *motan*), with the exception of *sculan* which appears both as *sculan* and *sceal*. In all such cases, the entry in the form as it appears in the *TOE* dataset was tagged and no new entries were created next to those already present in *TOE-LDD*.

3.5 *Semantic and Lexical Information in the Glossaries That Is Not Present in TOE*

Occasionally, the published glossaries contained words or word senses that were not found in the *TOE* dataset. For instance, *TOE* has not consistently added comparative and superlative forms to its dataset (it has entries for, e.g., *ieldra*, *leofre*, (*ge*)*rumlicor*, *ieldest*, *gingest* and *oftost*, but not for *freondlicor*, *oftor*, *snotorlicor*, *deorest*, *leofest* or *strengest*), while most glossaries do consistently include comparative and superlative forms.⁷ Another issue concerns how editors have dealt with instances in a text where two words might form either a compound or a syntactic phrase. The headwords of the *Beowulf* glossary (Fulk, Bjork, and Niles, 2008), for instance, include *ealdmetod* 'god of old', *ealdsweord* 'old sword' and *fifdagas* 'space of five days' as compound nouns, while these words are not recognized by either the *DOE* or *TOE* and, instead, are assumed to constitute syntactic phrases.⁸ Wherever such discrepancies occur, no new entries were created next to those already present in *TOE-LDD*.

Each of the tagging practices outlined above is based on the overarching principle that the underlying dataset of *TOE-LDD* should form the basis for the textual thesaurus. Even though *Evoke* allows for *TOE-LDD* to be enriched with additional senses and entries (see, e.g., the contribution by Van de Poel and Stolk in this special issue), this step did not seem practical or worthwhile at

7 Similarly, and perhaps for more obvious reasons, *TOE* does not include all numerals, prepositions, personal pronouns and conjunctions.

8 On the issue of distinguishing between compounds and syntactic phrases in Old English, see Kastovsky (1992: 362).

this stage, especially since the comparative analysis relies on the same underlying dataset to be used for each textual thesaurus.

Following these tagging principles, three textual thesauri have been made, which are currently fully functional in combination with *TOE*-LLD within Evoke. The *Beowulf Thesaurus 1.0* consists of 3,549 entries, with 9,185 senses; the *Andreas Thesaurus 1.0* consists of 2,418 entries, with 7,251 senses; and the *Old English Martyrology Thesaurus 1.0* consists of 2,610 entries, with 7,843 senses.⁹ Each thesaurus has a version number (these prototypes are version 1.0), since it is anticipated that these thesauri will be made more fine-grained in the future, either by adjustment of *TOE*-LLD (e.g., resolving some of the multiple entries or adding headwords from the published glossaries) or by changing the tagging strategy (e.g., by tagging at sense level). It is important to note that, at present, these textual thesauri are examples of ‘top-down thesauri’ – the onomasiological structure of *TOE* is imposed upon the vocabulary of the text rather than deriving from the senses found in the text itself. While the latter practice may be more precise and fine-grained, the top-down method is a more effective and practical means of making a textual thesaurus and it has the added advantage of allowing a comparative analysis between different texts that are tagged within the same onomasiological structure, as is demonstrated below.

4 Results: A Comparative Analysis of the Vocabulary in *Beowulf*, *Andreas* and the *Old English Martyrology*

Evoke contains a number of built-in statistical tools for the analysis of onomasiological datasets in LLD form (see Stolk, 2021). These tools range from a simple ‘item count’ to visualizing the distribution of senses across the semantic hierarchy of the dataset; Evoke can also be used to generate statistical information concerning polysemy (the number of senses per entry), degree of synonymy (the number of synonyms available per sense) and degree of specificity (how far down the semantic hierarchy senses are found).¹⁰ The following subsections illustrate the potential of these statistical tools when applied to

9 The three datasets are made available in the digital repository DataverseNL (see Porck, 2021a; 2021b; 2021c) and the textual thesauri can be accessed on the Evoke platform via <http://evoke.ullet.net/content>.

10 Since the four datasets discussed in this paper did not show significant differences in terms of their degree of specificity, this feature is not discussed here. For possible implications of this aspect of an onomasiological profile, see the contributions by Stolk and Van Baalen in this special issue.

TOE-LLD in combination with the textual thesauri of *Beowulf*, *Andreas* and the *Old English Martyrology*.

4.1 *General Impressions: Analysing Full Datasets*

Using Evoke’s ‘item count’ functionality, a number of general features can be distilled from the full datasets. These item counts are summarized in Table 1.

These general item counts already allow for a number of observations. From the item counts of *TOE*-LLD, it can be established that about 15% of the Old English lexicon is restricted to the genres of glosses, while more than 11% is exclusively found in poetic contexts. The *TOE* distribution flag “p” in combination with the textual thesauri reveal how the lexicon of *Beowulf* consists of more entries exclusively found in poetic contexts (30.66%) than that of *Andreas* (20.60%), while the lexicon of the prose *Old English Martyrology* is, naturally, devoid of poetic vocabulary. The relative difference between the two poems can be partially attributed to the fact that the proportion of *hapax legomena* within the lexicon of *Beowulf* is higher than that within the lexicon of *Andreas* (18.71% vs. 6.2%; by default, these *hapax legomena* are restricted to poetry). It is important to bear in mind that these figures cannot directly translate into a

TABLE 1 Item counts of the full datasets of *TOE*-LLD, *Beowulf Thesaurus 1.0*, *Andreas Thesaurus 1.0* and *Old English Martyrology Thesaurus 1.0*, derived from Evoke

	<i>TOE</i> -LLD	<i>Beowulf</i>	<i>Andreas</i>	<i>OE Martyrology</i>
Entries	35,422	3,549	2,418	2,610
Senses	51,480	9,185	7,251	7,843
Entries with <i>TOE</i> flag “p”	3,962 (11.19%)	1088 (30.66%)	498 (20.60%)	[2] ^a
Entries with <i>TOE</i> flag “o”	9,574 (27.03%)	572 (16.12%)	149 (6.16%)	51 (1.95%)
Entries with <i>TOE</i> flag “g”	5,367 (15.15%)	0 (0 %)	[1] ^b	8 (0.3%)
Entries with <i>TOE</i> flag “q”	778 (2.20%)	6 (0.17%)	7 (0.29%)	3 (0.11%)
Entries with custom tag “hapax”		664 (18.71%)	150 (6.20%)	19 (0.73%)

a These entries, *wer* and *weard*, are incorrectly tagged “p” in *TOE*, given their occurrence in the *Old English Martyrology* and other Old English prose texts.

b This is the entry *dung* which is incorrectly tagged “g” in *TOE*, see footnote 6 above.

statement like ‘the lexis of *Beowulf* is more poetic than that of *Andreas*’, especially since matters of chronology, dialect and textual survival may also influence whether certain words are restricted to poetic contexts. Nevertheless, the figures may be indicative of a more versatile and inventive poet being responsible for *Beowulf* (cf. Brodeur, 1959), while the *Andreas* poet may have been more traditional and derivative in his lexical choices (cf. Orchard, 2016).

Some aspects of the textual thesauri do not seem to square with the underlying *TOE* dataset. Notably, *Beowulf* has more *hapax legomena* than entries with the *TOE* distribution flag “o”, which confirms that the *TOE* editors have not been consistent in applying the “o” flag for words that only occur in single texts.¹¹ The 8 entries in the *Old English Martyrology* that have the *TOE*’s distribution flag “g” showcase, on the one hand, that the flag “g” is not applied to words *exclusively* used in glosses and glossaries,¹² and, on the other hand, that the vocabulary of the *Old English Martyrology* has close ties to the Anglo-Saxon glossing tradition (for which, see Rauer, 2016).

Evoke also allows the ambiguity or polysemy of entries to be analysed. The degree of ambiguity of a text’s lexicon may be an insightful feature to distinguish between text types; for instance, it is generally assumed that Old English riddles employ ambiguous vocabulary in order to obscure their answers (see, e.g., Afros, 2015), whereas Old English legal texts may be assumed to use more specific and univocal lexis. The ambiguity analyses of the four datasets under discussion in this article is provided in Figure 2.

When compared to the ambiguity graph for *TOE*-LLD, the three textual thesauri show a clearly distinct pattern. While 78.40% of *TOE*-LLD entries has only 1 sense, the percentages of single-sense entries within the textual thesauri are notably lower: 52.04% for *Beowulf*, 43.40% for *Andreas* and 42.91% for the *Old English Martyrology*. The average degree of ambiguity shows similar differences: on average, a *TOE*-LLD entry has 1.453 senses, while the averages within the textual thesauri are 2.588 for *Beowulf*, 2.999 for *Andreas*, and 3.005 for the *Old English Martyrology*. The low average ambiguity for the *TOE*-LLD dataset is likely due to the fact that this dataset includes legal, medical and encyclopedic terms which are likely to be less ambiguous. While the differences between the

11 For this issue, see the discussion of the *TOE* distribution flags on the *TOE* website: “Many Old English words occur once only. Yet quite how to identify a word as truly a *hapax legomenon* is problematic if its appearance in multiple manuscripts of some one work be held to constitute evidence of its greater currency. The flag o should be viewed as a warning that a particular word form is very infrequent. Conversely, not all *hapax legomena* may have been identified.” <https://oldenglishtesaurus.arts.gla.ac.uk/flags/>.

12 On possible inconsistencies of applying the “g” flag, see the discussion on the *TOE* website, <https://oldenglishtesaurus.arts.gla.ac.uk/flags/>.

FIGURE 2 Ambiguity graph showing the percentage of entries on the Y-axis and the number of senses on the X-axis. Data derived from Evoke. Information about entries with more than 5 senses (<1%) is not included in this graph

two poems and the prose text are less pronounced, the vocabularies of *Beowulf* and *Andreas* are on the whole less polysemous than the words used in the *Old English Martyrology*. This difference is probably due to the higher proportion of poetic words and *hapax legomena* in the verse texts, the latter of which may have been coined for specific, unambiguous purposes. This suspicion is confirmed by the fact that the average ambiguity of all entries flagged “p” in the *TOE-LLD* dataset is indeed very low: 1.097, with 91.49% having only a single sense.¹³ In other words, the higher the proportion of “p”-flagged words in a text, the lower its average ambiguity.

While the onomasiological profiles of the two verse texts show a lower average degree of ambiguity than that of the prose text, the situation is reversed when the average synonymy is taken into consideration. In theory, the higher the synonymy of a text’s vocabulary, the more carefully selected its senses are assumed to be, given the fact that its inclusion in the text is the result of a more considered selection between alternatives (see Stolk in this special issue). On average, a sense in the *Beowulf Thesaurus 1.0* has 6.177 synonyms

13 This number is based on the 3,962 entries with the “p” flag in *TOE-LLD*. The 1,088 “p”-flagged entries in *Beowulf* have an average ambiguity of 1.115, while the 498 “p”-flagged entries in *Andreas* have an average ambiguity of 1.217. A distribution flag, e.g., the “p” flag, already limits the (distribution of) occurrences of a word, making it already more likely that fewer senses were recorded for it.

and in the *Andreas Thesaurus 1.0* each sense has an average of 6.056 synonyms, whereas the senses in the *Old English Martyrology Thesaurus 1.0* have an average of 4.949 synonyms (an average that is similar to that of *TOE-LLD*, in which every sense has an average of 4.917 synonyms). Future research will need to show whether this higher average degree of synonymy in verse texts is perhaps affected by the semantic distribution of senses in particular poems or whether this is indeed a distinctive feature of verse texts in general as opposed to prose texts.¹⁴

The distribution of the senses across the semantic hierarchy is perhaps the most salient difference between the four datasets. Figure 3 shows four graphs generated by Evoke, one for each dataset, which show the relative distribution of senses across the 18 main semantic categories identified by *TOE*.

While each graph shows a somewhat similar pattern, with the category “Aught, anything, something” being best represented (16.76% of senses in *TOE-LLD*; 20.77% in *Beowulf*; 20.58% in *Andreas*; and 21.45% in the *Old English Martyrology*) and the domain “Quiet, leisure, rest” showing the lowest percentage of senses (0.77% of senses in *TOE-LLD*; 0.61% in *Beowulf*; 0.57% in *Andreas*; 0.6% in the *Old English Martyrology*), the onomasiological profiles of the three texts show a number of distinct characteristics in terms of their sense distribution. Within the lexicon of *Beowulf*, for instance, the categories “Having, owning, possession” and “Peace, tranquillity” are relatively overrepresented (respectively, the two categories make up 3.14% and 5.19% of the senses in *Beowulf*, against 1.67% and 2.85% of the senses in *TOE-LLD*), which demonstrates that the poem’s interests in the exchange of items and violence (cf. Baker, 2013) are reflected in the vocabulary it uses (crucially, “War” is a subcategory of “Peace, tranquillity”). The onomasiological preference for warfare in *Beowulf* is particularly clear in comparison to the relative underrepresentation of the category of “Peace, tranquillity” in the *Old English Martyrology*, where it accounts for only 1.77% of the senses. By contrast, the category “Consumption of food/drink” is relatively underrepresented in all three texts (3.44% of the senses in *Beowulf*, 3.83% in *Andreas* and 5.28% in the *Old English Martyrology*, against 7.79% in *TOE-LLD*), which shows that single texts rarely feature a very wide variety of terms relating to food and drink. Be that as it may, *Beowulf* is a text famous for its feasting scenes while *Andreas* deals with cannibalistic monsters (see, e.g., Magennis, 1999; Battles, 2015), so the fact that the *Old English Martyrology* still has relatively more senses devoted to the category of

14 With regard to the latter, it is noteworthy that the 4,346 senses that belong to the 3,962 “p”-flagged entries in *TOE-LLD* have an average synonymy of 7.841.

FIGURE 3 Distribution of senses across the 18 main semantic categories of *TOE*

TABLE 2 Relative distribution of senses across the 18 main semantic categories

TOE category	TOE-LLD	<i>Beowulf</i>	<i>Andreas</i>	<i>OE Martyrology</i>
Action, operation	4.86%	6.9%	6.37%	5.88%
Aught, anything, something	16.76%	20.77%	20.58%	21.45%
Consumption of food/drink	7.79%	4.5%	3.83%	5.28%
Creation	14.94%	10.88%	10.18%	12.09%
Earth, world	3.5%	3.2%	4.01%	3.07%
Feelings	5.45%	6.61%	6.45%	4.63%
Having, owning, possession	1.67%	3.14%	2.81%	2.49%
Judgement, forming of opinion	3.38%	3.49%	3.61%	3.2%
Law, custom, covenant	2.86%	2.48%	2.8%	2.7%
Material, matter, substance	6.95%	6.33%	6.4%	7.61%
Peace, tranquillity	2.85%	5.19%	3.28%	1.77%
Power, might	8.06%	8.54%	8.36%	8.48%
Property, possessions, wealth	1.73%	2.06%	1.52%	1.47%
Quiet, leisure, rest	0.77%	0.61%	0.57%	0.6%
Speech, vocal utterance	3.67%	3.7%	4.51%	4.59%
Spirit, soul, heart	6.29%	6.85%	7.65%	5.78%
The extrasensorial world	6.68%	3.44%	5.76%	7.37%
Work, doings, actions, labour	1.79%	1.33%	1.3%	1.56%

“Consumption of food/drink” is remarkable and worthy of further research.¹⁵ A full overview of the data underlying Figure 3 is provided in Table 2.

While the application of Evoke on the entire datasets gives an immediate bird’s eye view of the onomasiological profiles of the thesauri, visualizations like Figure 3 and tabulations like Table 2, while insightful, only allow for very broad sketches concerning vocabulary use. The application in Evoke becomes even more revealing when it is applied to more specific parts of the onomasiological taxonomy, as the following two subsections demonstrate.

15 For a very brief discussion of food and consumption in the *Old English Martyrology*, see Frantzen (2014: 153–154).

Beowulf

Andreas

Old English Martyrology

FIGURE 4 Distribution of senses across the subcategories of “13.02 War”, generated by Evoke. The Y-axis shows the percentage of senses; the subcategories are listed on the X-axis. Note that Evoke currently does not allow for the Y-axis to be manually adjusted and, hence, it has automatically generated a different Y-axis for each thesaurus

4.2 *Zooming In: War and Warriors in Beowulf, Andreas and the Old English Martyrology*

The three Old English texts show marked onomasiological differences within the *TOE* category “13.02. War”. The category is best represented in *Beowulf* where 11.33% of all entries (= 402 entries) have at least one sense within this category; for *Andreas*, this percentage is 7.16% (= 184 entries), while in the *Old English Martyrology* the percentage is down to 4.3% (= 112 entries). Clearly, the poet of *Beowulf* felt a greater need to use and vary lexis relating to this subject, while the author of the *Old English Martyrology* wrote a text which did not require as wide a range of words to describe acts of war.

The manner in which the senses of this category are distributed across the subcategories of “13.02 War” show comparable patterns for each text: in each thesaurus, the categories “13.02.08 Military Equipment” and “13.02.10 The Military, soldiers” are best represented, but in *Beowulf* the former takes up a much larger proportion (42.26%) than it does for the other texts (23.26% in both *Andreas* and the *Old English Martyrology*) (see Figure 4). A larger onomasiological representation of “13.02.08 Military equipment” is expected, seeing as it is also the largest subcategory in *TOE*-LLD (where it accounts for 33.56% of the senses in subcategories of “13.02 War”), but the size of this subcategory in *Beowulf* is nevertheless remarkable. This distinctive pattern of vocabulary use may have to do with the *Beowulf* poet’s fondness of coining new terms relating to weapons (see Brady, 1979). Indeed, a sizeable proportion of the poem’s *hapax legomena* (92 out of 664) have at least one sense in this semantic category.

The second-best represented subcategory of “13.02 War”, “13.02.10 The military, soldiers” allows for gauging the onomasiological overlap between the three textual thesauri. *Beowulf* shows the greatest lexical variety in this semantic field with 87 entries having at least one sense within this category, out of which 24 are *hapax legomena*. In *Andreas*, this category is home to 49 entries, with 3 *hapax legomena*, and in the *Old English Martyrology* there are 24 entries with at least one sense within this category. In terms of onomasiological overlap, it is interesting to note that out of *Andreas*’s 49 entries, 37 also occur in *Beowulf*, which is testimony to the two poems’ close connection in terms of lexis (see North and Bintley, 2016: 63). The 24 entries in the *Old English Martyrology*, meanwhile, show different patterns of overlap: 17 are shared with *Beowulf* and 16 with *Andreas*, while 14 entries with at least one sense in “13.02.10 The military, soldiers” appear in all three texts and are, therefore, likely the most generic terms for warriors.

FIGURE 5 Distribution of senses across subcategories of *TOE* category “02.06 Animal”, excluding subcategories which had fewer than 1% of senses in *TOE*-LLD. The Y-axis shows the percentage of senses; the subcategories are listed on the X-axis

4.3 Textual Elaboration and Onomasiological Representation: Animals in *Beowulf*, *Andreas* and the *Old English Martyrology*

The item counts for entries with at least one sense in the *TOE* category “13.02 War” at the start of section 4.2 of this article already suggested that there may be a connection between the size of a particular semantic category within a text’s onomasiological profile and the salience of that particular topic within that text. *Beowulf*, a heroic poem about warriors, uses relatively more words to denote weapons and warfare than a prose martyrology whose protagonists are saints of all walks of life. This ‘textual elaboration’ is similar to the cultural elaboration described in section 2 of this article: the more salient a concept is within a text, the better it will be represented in the text’s vocabulary.

The premise of this notion of textual elaboration can be demonstrated through a brief analysis of the distribution of senses within the *TOE* category “02.06 Animal” (see Figure 5). In *TOE*-LLD, the three best represented subcategories are “Bird” (26.11%), “Domestic animals, livestock” (18%) and “Insect/small creature” (11.35%), but none of the three textual thesauri follow this hierarchy. Instead, each shows distinctive patterns of vocabulary use that can be related to the roles that animals play in the individual texts. The *Old English Martyrology* uses most senses in the subcategories “Domestic animals, livestock” (29.85%), “Bird” (14.93%) and “Wild animal” (13.43%), while the categories “Insect/small creature” and “Fish” are relatively underrepresented (5.97% and 4.48%, respectively, as opposed to 11.35% and 9.08% in *TOE*-LLD).

This onomasiological distribution matches scholarly observations about animals in medieval hagiography: typically, saints are served by domestic animals and birds, while wild animals either miraculously come to their aid or represent devils in disguise, while fish, let alone insects and small creatures, do not feature so prominently (see, e.g., Fiske, 1913; Alexander, 2008). In *Andreas*, as could be expected from a poetic saint's life, the animal-related vocabulary is also concentrated in the subcategories "Domestic animals, livestock" and "Bird" (both at 21.43%), but this text does not feature any senses in the subcategory "Wild animal". Instead, the categories "Animal parts/activities" and "Monster, strange creature" are relatively overrepresented (14.29% and 10.71%, respectively, against 9.88% and 6.24% in *TOE*-LLD). The *Andreas* poet likely employed the senses in these categories to describe the monstrous and cannibalistic inhabitants of Mermedonia – indeed, the description of their need "þæt hie tobrugdon blodigum ceafum/fira flæschoman him to foddorþege" [that they should draw apart with bloody jaws the flesh of men as fodder for their feasting] (ll. 159–160, ed. and trans. North and Bintley, 2016: 126) has notable animalistic undertones. Turning lastly to *Beowulf*, the subcategory of "Monster, strange creature" clearly stands out as it takes up 39.47% of the senses in the subcategories of "02.06 Animals". As with the other textual thesauri, it is not difficult to explain this particular onomasiological distribution; anyone familiar with the contents of the poem knows that monsters take central stage (Tolkien, 1936).

This excursus into the distribution of senses across subcategories of the *TOE* category "02.06 Animal" has demonstrated that onomasiological over- and underrepresentation can be related to the salience of concepts in texts. If it works for vocabulary related to animals, distinctive patterns in the use of words for emotions, colours, weather, stages of the life course and various other cultural concepts may equally be suggestive of the relevance of those concepts within these texts. In this way, the textual thesauri produced in *Evoke* can form the starting point of new inquiries into these age-old texts.

5 Conclusion

The aim of this article was to introduce a new and digital approach to studying the vocabulary of Old English texts. As sections 4.1–4.3 demonstrate, novel statistical information can be generated through the textual thesauri that have been produced within the Web application *Evoke*. This information highlights distinctive patterns of vocabulary use in these texts, which not only help to

substantiate and quantify existing statements about the lexical make-up of these texts, it also prompts new kind of inquiries.

The article has also highlighted a number of ways in which the results generated through Evoke might be fine-tuned in the future. In particular, it is important that the underlying *TOE* dataset in LLD form is manually checked for inadequacies at entry level that have resulted from the automatic porting process; moreover, since the *TOE* editors occasionally update elements of their dataset, these updates should also find their way into *TOE*-LLD.¹⁶ In addition, improvements within the textual thesauri presented here are also possible, by, for instance, adding headwords from the published glossaries that are currently not in the *TOE* dataset, or by tagging at sense level rather than at entry level. More generally, there are both advantages and disadvantages to the use of *TOE*-LLD as the onomasiological structure within which textual thesauri are produced. The main advantages are the relative speed and ease with which such textual thesauri can be created, as well as the fact that comparative analyses are enabled through the use of the same onomasiological hierarchy. A potential shortcoming is that the onomasiological structuring of *TOE* is based on an entire language and does not derive naturally from the vocabulary used in an individual text. As a result, the choices of the *TOE* editors are carried through in the textual thesauri. With respect to the animal vocabulary discussed in section 4.3, for instance, one might wonder whether words denoting “Monster, strange creature” really belong to the semantic field of animals and whether a ‘bottom-up’ textual thesaurus of *Beowulf* might not have chosen for a different semantic hierarchy, grouping monsters with the supernatural rather than with horses and birds.

These caveats notwithstanding, textual thesauri in combination with the Web application Evoke can become a valuable research tool for future analyses of Old English texts. As Michael D. C. Drout *et al.* (2016) have demonstrated, newly developed “computer-assisted statistical analyses” can “augment traditional philological and literary techniques” (2), especially in a field where the discovery of new insights relies increasingly on technological innovation, rather than newly discovered primary material. As a tool that uncovers hitherto unnoticed and distinctive patterns of Old English vocabulary use, within the entire corpus as well as in individual Old English texts, Evoke has the potential to influence the future of both digital and traditional philological research.

16 *TOE*-LLD available in Evoke at the time of writing this article is based on the version of *TOE* that was ported by Sander Stolk at 26 May 2017. A number of issues mentioned in this article have already been solved in the latest version of *TOE*.

Acknowledgement

This work was supported by the LUCAS Extra Resources Open Call-11 Grant 2020, awarded by the Leiden University Centre for the Arts in Society. The author would like to thank Sander Stolk and Lucas Gahrmann for their technical support, as well as the anonymous peer reviewers for their insightful comments.

References

- Afros, E. "Linguistic Ambiguities in Some Exeter Book Riddles." *Notes and Queries* 52 (2005), 431–437.
- Alexander, D. D. *Saints and Animals in the Middle Ages* (Woodbridge: Boydell Press, 2008).
- Baker, P. S. *Honour, Exchange and Violence in Beowulf*, Anglo-Saxon Studies 20 (Cambridge: D. S. Brewer, 2013).
- Basilius, H. "Neo-Humboldtian Ethnolinguistics." *Word* 8 (1952), 95–105.
- Battles, P. "Dying for a Drink: 'Sleeping after the Feast' Scenes in *Beowulf*, *Andreas*, and the Old English Poetic Tradition." *Modern Philology* 112 (2015), 435–457.
- Brady, C. "Warriors in *Beowulf*: An Analysis of the Nominal Compounds & an Evaluation of the Poet's Use of Them." *Anglo-Saxon England* 11 (1983), 199–246.
- Brady, C. "Weapons in *Beowulf*: An Analysis of the Nominal Compounds & an Evaluation of the Poet's Use of Them." *Anglo-Saxon England* 8 (1979), 79–141.
- Bremmer Jr, R. H. "Old English 'Cross' Words." In *Cross and Cruciform in the Anglo-Saxon World: Studies to Honor the Memory of Timothy Reuter*, eds. S. L. Keefer, K. L. Jolly and C. E. Karkov (Morgantown: West Virginia UP, 2010), 204–231.
- Bremmer Jr, R. H. "Old English Heroic Literature." In *Readings in Medieval Literature. Interpreting Old and Middle English Literature*, eds. D. F. Johnson and E. Treharne (Oxford: OUP, 2006), 75–90.
- Brodeur, A. *The Art of Beowulf* (Berkeley: University of California Press, 1959).
- Díaz Vera, J. E., ed. *A Changing World of Words: Studies in English Historical Lexicography, Lexicology and Semantics* (Amsterdam: Rodopi, 2002).
- DOE = Cameron, A., A. Crandell Amos, A. diPaolo Healey, et al., eds. *Dictionary of Old English: A to I Online* (Toronto: Dictionary of Old English Project, 2018).
- Drout, M. D. C., Y. Kisor, L. Smith, A. Dennett, and N. Piirainen. *Beowulf Unlocked. New Evidence from Lexomic Analysis* (s.l.: Springer, 2016).
- Evoke = Stolk, S. Evoke (Web application, 2018), <http://evoke.ullet.net/>.
- Fabiszak, M. "A Semantic analysis of FEAR, ANGER and GRIEF Words in Old English." In Díaz Vera (2002), 255–274.

- Fiske, C. F. "Animals in Early English Ecclesiastical Literature, 650–1500." *PMLA* 28 (1913), 368–387.
- Frantzen, A. J. *Food, Eating and Identity in Early Medieval England*, Anglo-Saxon Studies 22 (Woodbridge: Boydell Press, 2014).
- Fulk, R. D., R. E. Bjork, and J. D. Niles, eds. *Klaeber's Beowulf*, 4th edn (Toronto: University of Toronto Press, 2008).
- Gevaert, C. "The Evolution of the Lexical and Conceptual Field of ANGER in Old and Middle English." In Díaz Vera (2002), 275–300.
- Goddard, C. "Words as Carriers of Cultural Meaning." In *The Oxford Handbook of the Word*, ed. J. R. Taylor (Oxford: OUP, 2015), 380–398.
- Hall, A. *Elves in Anglo-Saxon England: Matters of Belief, Health, Gender and Identity*, Anglo-Saxon Studies 8 (Woodbridge: Boydell Press, 2007).
- HTE = Kay, C., M. Alexander, F. Dallachy, J. Roberts, M. Samuels, and I. Wotherspoon, eds. *The Historical Thesaurus of English*, 2nd edn, version 5.0 (Glasgow: University of Glasgow, 2021), <https://ht.ac.uk/>.
- Izdebska, D. "Weapon-Boys and Once-Maidens: A Study of Old English Vocabulary for Stages of Life." In *Early Medieval English Life Courses: Cultural-Historical Perspectives*, eds. T. Porck and H. Soper, Explorations of Medieval Culture 20 (Leiden: Brill, 2022), 47–89.
- Kastovsky, D. "Semantics and Vocabulary." In *The Cambridge History of the English Language. Vol. 1: The Beginnings to 1066*, ed. R. M. Hogg (Cambridge: Cambridge UP, 1992), 290–408.
- Kay, C. "The *Thesaurus of Old English* Online." *Old English Newsletter* 38:3 (2005), 36–40.
- Kopár, L. "Spatial Understanding of Time in Early Germanic Cultures: The Evidence of Old English Time Words and Norse Mythology." In *Interfaces between Language and Culture in Medieval England*, eds. A. Hall, O. Timofeeva, Á. Kiricsi and B. Fox (Leiden: Brill, 2010), 203–230.
- Magennis, H. *Anglo-Saxon Appetites Food and Drink and Their Consumption in Old English and Related Literature* (Dublin: Four Courts Press, 1999).
- Mathiot, M., ed. *Ethnolinguistics: Boas, Sapir and Whorf Revisited* (Mouton: The Hague, 1979).
- Momma, H. Review of *TOE, Notes & Queries* ns 50 (2003), 79–80.
- Neidorf, L., ed. *The Dating of Beowulf: A Reassessment*, Anglo-Saxon Studies 24 (Cambridge: D. S. Brewer, 2014).
- North, R., and M. D. J. Bintley, eds. *Andreas: An Edition*, Exeter Medieval Texts and Studies (Liverpool: Liverpool UP, 2016).
- O'Brien O'Keeffe, K. "Diction, Variation, the Formula." In *A Beowulf Handbook*, eds. R. E. Bjork and J. D. Niles (Lincoln: University of Nebraska Press, 1997), 85–104.

- Orchard, A. "The Originality of *Andreas*." In *Old English Philology: Studies in Honour of R. D. Fulk*, eds. L. Neidorf, R. J. Pascual and T. A. Shippey (Cambridge: D. S. Brewer, 2016), 331–370.
- Palmer, G. B. *Toward a Theory of Cultural Linguistics* (Austin: University of Texas Press, 1996).
- Porck, T. "Andreas Thesaurus." DataverseNL, V1 (2021a), <https://doi.org/10.34894/IHVHoZ>.
- Porck, T. "Beowulf Thesaurus." DataverseNL, V1 (2021b), <https://doi.org/10.34894/TOTFGZ>.
- Porck, T. "Old English Martyrology Thesaurus." DataverseNL, V1 (2021c), <https://doi.org/10.34894/QZCNW1>.
- Rauer, C. "The *Old English Martyrology* and Anglo-Saxon Glosses." In *Latinity and Identity in Anglo-Saxon England*, eds. R. Stephenson and E. V. Thornbury (Toronto: University of Toronto Press, 2016), 73–92.
- Rauer, C., ed. *The Old English Martyrology: Edition, Translation and Commentary*, Anglo-Saxon Texts 10 (Cambridge: D. S. Brewer, 2013).
- Sapir, E. "Language." In *Selected Writings of Edward Sapir in Language, Culture and Personality*, ed. D. G. Mandelbaum (Berkeley: University of California Press, 1951), 7–32.
- Sharifian, F. *Cultural Conceptualisations and Language: Theoretical Framework and Applications* (Amsterdam: John Benjamins, 2011).
- Stolk, S. "A Thesaurus of Old English as Linguistic Linked Data: Using OntoLex, SKOS and lemon-tree to Bring Topical Thesauri to the Semantic Web." *Proceedings of eLex 2019* (2019), https://elex.link/elex2019/wp-content/uploads/2019/09/eLex_2019_13.pdf.
- Stolk, S. "Evoke: A Web Application for Thesauri and Linguistic Linked Data." *Lexicala Review* 29 (2021), <https://lexicala.com/review/2021/evoke-a-web-application-for-thesauri-and-linguistic-linked-data/>.
- Strite, V. *Old English Semantic-Field Studies* (New York: Peter Lang, 1989).
- Taranu, C. "Senses of the Past: The Old English Vocabulary of History." *Florilegium* 29 (2012), 65–88.
- TOE = Roberts, J., and C. Kay with L. Grundy. *A Thesaurus of Old English* (Glasgow: University of Glasgow, 2017), <http://oldenglishthesaurus.arts.gla.ac.uk/>.
- Tolkien, J. R. R. "Beowulf: The Monsters and the Critics." *Proceedings of the British Academy* 22 (1936), 245–295.
- Wierzbicka, A. *Understanding Cultures through Their Key Words. English, Russian, Polish, German, and Japanese* (Oxford: OUP, 1997).