



Universiteit  
Leiden  
The Netherlands

## **Digital thesauri as semantic treasure troves: a Linguistic Linked Data approach to "A Thesaurus of Old English"**

Stolk, S.S.

### **Citation**

Stolk, S. S. (2023, May 31). *Digital thesauri as semantic treasure troves: a Linguistic Linked Data approach to "A Thesaurus of Old English"*. Retrieved from <https://hdl.handle.net/1887/3619351>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3619351>

**Note:** To cite this publication please use the final published version (if applicable).

## Chapter 8

## 8. Research case studies with *A Thesaurus of Old English* and Evoke

### 8.1. Introduction

This chapter assesses the usefulness of *A Thesaurus of Old English* (*TOE*) as Linguistic Linked Data, both for research and educational purposes, through a discussion of a number of case studies. These case studies were part of a collaborative project titled ‘Exploring Early Medieval English Eloquence’ (EEMEE), established specifically for this purpose.<sup>1</sup> The project brought together scholars from universities and lexicographic institutions from across Europe to explore – and elaborate on – the contents of *TOE* using the web application Evoke. Three workshops, organized between 2019 and 2020, facilitated the development and presentation of case studies within the project (see Appendix 8.B). These events informed further refinements to the requirements for Evoke. As described in Chapter 2, functionalities identified and requested for research purposes have been implemented as part of newer iterations of this web application.

Case studies in the EEMEE project approached the thesaurus information from the perspectives of various disciplines: linguistics, literary criticism, history, lexicography, and philology. Participating researchers (and in educational settings, their students) used Evoke to explore the contents of *TOE* alongside additional material. These extended sets of information were fashioned either by linking an existing source to *TOE* data or by using the annotation system within Evoke. Those case studies which were presented at the EEMEE workshop at the 21st International Conference on English Historical Linguistics have been published in a special issue of the international, peer-reviewed journal *Amsterdamer Beiträge zur älteren Germanistik* titled ‘Exploring Early Medieval English Eloquence: A Digital Humanities Approach with *A Thesaurus of Old English* and Evoke’.<sup>2</sup> Many of the datasets created in this project have been made available publicly, under an open license (CC-BY-SA), in the DataverseNL data repository and in Evoke (see Appendix 8.A). Their open access availability is intended to encourage future study beyond the scope of the EEMEE project and not restricted to any single software application.

The remainder of this chapter is structured as follows. Section 8.2 provides background information on *TOE* and its relation to Old English language and

<sup>1</sup>EEMEE was supported by the LUCAS Extra Resources Open Call-II Grant 2020, awarded by the Leiden University Centre for the Arts in Society, and the LUCDH Small Grant 2018, awarded by the Leiden University Centre for Digital Humanities.

<sup>2</sup>*Amsterdamer Beiträge zur älteren Germanistik* 81.3-4. Figures included are for illustrative purposes only, stemming from articles or presentations, and are used with permission.

culture. Subsequently, sections 8.3-8.7 present, and reflect on, notable EEMEE case studies on the topics of the history of Old English lexicography, stylistics of Old English writing, diachronic developments of Old English, comparative analyses of Old Germanic languages, and teaching Old English language and culture. The case studies are followed by an evaluation of the usefulness of the Linguistic Linked Data form of *TOE* and the web application Evoke for research purposes in section 8.8.

## 8.2. A *Thesaurus of Old English* and its relation to Old English language and culture

Originating as a pilot study for University of Glasgow's *Historical Thesaurus* project, *TOE* was quickly recognized as an essential resource in its own right.<sup>3</sup> Since its first publication in 1995, this thesaurus has been the point of departure for many scholars studying aspects of the language and culture of early medieval England. Scholars performing semantic field studies of the language or literary-critical analyses of Old English texts have profited from the ability to look up which Old English words were available to denote particular concepts.<sup>4</sup> The current section explores the relation of this thesaurus to Old English language and culture, examining whose language has been recorded and its relation to the culture of England's early medieval inhabitants. As will become evident, thesauri and related bodies of knowledge can act as "a series of step-by-step doorways into the heart of a national culture".<sup>5</sup>

For their lexis, *TOE* and other dictionaries of Old English have relied on surviving texts of the early medieval language. Spoken between roughly 500 and 1100, the language is found, in runes or a version of the Latin alphabet, in "a rich diversity of records written on parchment, carved in stone and inscribed in jewelry".<sup>6</sup> In total, the corpus of surviving texts, as found in *DOEC*, contains 3,037 texts with over 3 million Old English tokens, i.e., orthographic words rather than lemmas found in dictionaries. The size of this corpus, which is said to equal "almost five times the collected works of Shakespeare",<sup>7</sup> is only a small fraction of the surmised body of texts that must have existed.<sup>8</sup> The variety of the surviving texts notwithstanding, scholars should have an understanding of the writing culture in which these texts were produced in order to interpret their findings based on *TOE* correctly.

A select group in early medieval society could read; an even smaller group

---

<sup>3</sup>Bremmer, 'Treasure Digging', p. 109; Dance, Review of *TOE1*, p. 312; Görlach, Review of *TOE1*, pp. 398-9.

<sup>4</sup>See, for instance, Bouwer, *Studien zum Wortfeld um eald und niwe im Altenglischen*; Moriyama, 'Synonyms and Synonymous Expressions in the Old English Semantic Field "Hospitality, Harbours, and Entertaining"'; Roberts, 'The Old English Vocabulary of Nobility'.

<sup>5</sup>*ScT*, p. ix.

<sup>6</sup>About the Dictionary of Old English'.

<sup>7</sup>*Ibid.*

<sup>8</sup>To illustrate, recent research utilizing unseen species models from ecology estimates 38.6% of narratives from Old and Middle English has survived (Kestemont et al., 'Forgotten books').

could write. Scribes of the surviving Old English texts were often learned monks and their writing reflects their interests: religious topics dominate the corpus.<sup>9</sup> Vivid examples include the many Saints' lives, such as Ælfric's *Lives of Saints*, and adaptations of Biblical books, such as the Old English *Genesis* and *Exodus* in the Junius Manuscript. Any lexicographic work concerned with the Old English lexicon is therefore bound to reflect this writing culture rather than representing the vocabulary of the majority of the populace of the early medieval kingdoms in England. Even so, much can be said about the lexis available to the population. Old English authors shared the vernacular language with the population at large and their writing influenced their society beyond the monasteries, too. Sermons, such as the Old English *Sermo Lupi ad Anglos*, will have been preached in the vernacular to the laity. The content of charters, which recorded transfer of land property, and legal texts will have affected inhabitants in their daily lives. Moreover, thematic domination in the corpus is mitigated in *TOE* through its inclusion of all words and word senses discovered in the corpus, regardless of the frequency with which they occur. Words that are attested infrequently, perhaps only once throughout the corpus, are incorporated alongside those that occur often and across numerous texts.

Analysing the Old English lexicon can offer insights into the culture of the medieval population who spoke it as well as the composition of the historical language. Alaric Hall argues this case convincingly:<sup>10</sup>

There is [...] a well-established and theoretically justified supposition that language reflects culture [...]. This, as a generalisation, can hardly be denied – if language did not reflect culture then it would be an absurdly ineffectual tool for communication.

That language can convey aspects of the culture using it is by no means a novel notion. Decades ago, Edward Sapir already argued that vocabulary is a “very sensitive index of the culture of a people”.<sup>11</sup> Indeed, this notion forms the cornerstone of cultural linguistics, in which the relationship between language and culture is explored.<sup>12</sup> *TOE*, through explorations of the Old English lexis, offers such glimpses. Degrees of lexicalization presented by the thesaurus, for instance, are indicative of the salience of concepts and semantic domains.<sup>13</sup> A high degree of lexicalization suggests the need for many nuances in communication by the language community concerned; an altogether lack of words for a concept suggests unimportance or absence. Low degrees of lexicalization for concepts may convey unimportance, too, although they may also be the result of ritualization or other linguistic processes.<sup>14</sup> These and other indicators of cultural aspects

<sup>9</sup>David and Simpson, ‘The Middle Ages to ca. 1485: Introduction’, p. 5.

<sup>10</sup>Hall, *Elves in Anglo-Saxon England*, p. 13.

<sup>11</sup>*Selected Writings of Edward Sapir in Language, Culture and Personality*, p. 27.

<sup>12</sup>Sharifian, ‘Cultural Linguistics and World Englishes’, p. 515.

<sup>13</sup>Wierzbicka, *Understanding Cultures through Their Key Words*, pp. 10–11.

<sup>14</sup>In his book chapter ‘Old English “Cross” Words’, Rolf H. Bremmer Jr identifies six native words in Old English for the cross that bore Jesus Christ, whereas “present-day English today possesses only one current word, ‘cross’,” a development which could be attributed to ritualisation of the religious vocabulary (p. 231).

encapsulated in the Old English lexicon have been investigated through *TOE* in various case studies of the EEMEE project.

In addition to surviving texts, *TOE* has relied on prior lexicographic work on the language. The process of creating the thesaurus has been described in the prefatory matter of its various editions and, more fully, in a recent article by Jane Roberts, one of its editors.<sup>15</sup> Existing dictionaries of Old English supplied the thesaurus editors with words and word senses, which were copied onto slips of paper and rearranged to form groups of synonyms to be positioned in an overarching, topical structure. Revisions on *TOE* are ongoing, as its current editor updates its content based on the findings of the team behind University of Toronto's *Dictionary of Old English (DOE)*. The knowledge of the Old English lexis recorded in *TOE* is therefore affected by lexicographic practices and choices during the compilation of these source dictionaries and the resulting thesaurus. The following section will delve more deeply in the history of Old English lexicography and present an EEMEE case study on the subject.

### 8.3. History of Old English lexicography

Several dictionaries of Old English have been published since the early modern period. The first was William Somner's *Dictionarium Saxonico-Latino-Anglicum*, which appeared in 1659. Almost two centuries later, Joseph Bosworth published his *A Dictionary of the Anglo-Saxon Language* in 1838. Fifty-six years later, John R. Clark Hall issued his *A Concise Anglo-Saxon Dictionary for the Use of Students* in 1894. Both Bosworth's and Clark Hall's dictionary were further developed and expanded, which resulted in publications of supplements and revisions in the late nineteenth and the twentieth century.<sup>16</sup> *TOE*, alongside University of Toronto's *Dictionary of Old English*, is one of the more recent, major lexicographical works of Old English that is still being updated. Developments in Old English lexicography can be scrutinised through the semantic domains available in *TOE*, as demonstrated by the case study below.

#### 8.3.1. Charting lexicographic developments

In her case study, Rachel Fletcher posited that Evoke and the Linguistic Linked Data version of *TOE* enable investigations into the representation of the Old English lexicon throughout the history of Old English lexicography.<sup>17</sup> As mentioned in the previous section, *TOE* presents a filtered image of the early medieval English lexicon: it is based on the surviving texts, interpreted by scholars, and on prior dictionaries. Each lexicographic work reflects its aim,

<sup>15</sup>Roberts, 'A Thesaurus of Old English: The Pilot Study for the Glasgow *Historical Thesaurus*'.

<sup>16</sup>In 1898, Thomas N. Toller published a revised edition of Bosworth's dictionary under the title 'An Anglo-Saxon Dictionary'. A supplement became available in 1921 and an edition with "enlarged addenda and corrigenda", by the hand of Alistair Campbell, appeared in 1972. Clark Hall's dictionary exists in four editions, of which the last published in 1960.

<sup>17</sup>Rachel Fletcher presented the case study in a paper titled 'Evoke and the History of Old English Lexicography: Preliminary Explorations' at the second EEMEE workshop (see Appendix 8.B).

target audience, the knowledge of the lexis at that particular point in time, and the editorial choices inherent in lexicography. Utilizing Evoke, Fletcher's case study explored how the representation of the Old English lexicon in *TOE* compares to its representation in earlier dictionaries, aiming to “shed light not only on the subjectivity of the *Thesaurus of Old English* (of which users should be aware) but on changing attitudes towards Old English and early mediaeval English culture over the history of scholarship”.<sup>18</sup>

The case study investigated five dictionaries alongside *TOE*, including the two principal dictionaries on which *TOE* is based. The five dictionaries are listed below in order of publication date.

- W. Somner, *Dictionarium Saxonico-Latino-Anglicum* (Oxford, 1659).
- E. Lye, *Dictionarium Saxonico et Gothico-Latinum* (London, 1772).
- J. Bosworth, *A Dictionary of the Anglo-Saxon Language* (London, 1838).
- J. Bosworth and T. N. Toller, *An Anglo-Saxon Dictionary Based on the Manuscript Collections of the Late Joseph Bosworth* (London, 1898), *Supplement* by T. N. Toller (Oxford, 1921), with *Enlarged Addenda and Corrigenda* by A. Campbell (Oxford, 1972).
- J. R. Clark Hall, *A Concise Anglo-Saxon Dictionary*, 4th edn, with a supplement by H. D. Meritt (Cambridge, 1960).

Custom labels were created to identify each dictionary ('Som', 'LM', 'Bos', 'BT', 'HM') and used to mark *TOE* word senses in Evoke to reveal which were recorded in the aforementioned dictionaries, thereby facilitating comparisons per semantic domain.

In her case study, Fletcher focused on two *TOE* categories: “13.02.10.01.01 A warrior, fighter” and “14.01.04 A legal right”. Figure 8.1 shows an overview of Old English words denoting the former, including the established custom labels as assigned to each. A diachronic visualization of these findings is shown in Figure 8.2. The results foreground that, compared to the current state of Old English lexicography, a disproportionate number of words denoting “A warrior, fighter” were not yet included in earlier dictionaries of Old English. Notably absent in these dictionaries are so-called kennings and other compounds found in poetic diction (e.g., *scildwiga* [lit. “shield warrior”] and *heoruwulf* [lit. “sword wolf”]). Fletcher posited that these findings reflect the historical interest in Old English lexicography, which focused initially on legal and historical sources (e.g., laws, chronicles) and expanded later to poetic texts.<sup>19</sup> In the future, Fletcher concluded, further developments in Old English lexicography may change this picture. Ongoing work on Toronto's *Dictionary of Old English*,

<sup>18</sup>Citation taken from the paper abstract submitted to the workshop.

<sup>19</sup>Fletcher indicated there are various reasons why the poetic lexicon received less attention in early scholarship, including the inaccessibility of some poetic texts (e.g., Vercelli MS was only rediscovered in the nineteenth century) and a higher difficulty in interpreting them (due to their freedom in syntax, lack of Latin parallel texts, and, importantly, use of hapax legomena, words that occur in only a single text throughout the corpus).

is improving scholarly understanding, and surfacing of new material will influence how researchers perceive and study Old English language and culture.

• A warrior, fighter		i	☰	📄	☁		
noun	æscwiga	poetry	BT	Sos	LM	Old English	
noun	beadurinc	poetry	BT	Sos	LM	Old English	
noun	burgwigend	poetry	rare	BT		Old English	
noun	cempa	BT	Sos	LM	Som	Old English	
noun	cumbolwiga	poetry	BT	Sos	LM	Old English	
noun	dreng	BT	Sos	LM		Old English	
noun	feohtend	BT				Old English	
noun	feohtere	glosses	rare	BT	Sos	LM	Old English
noun	feohtling	BT	Sos	LM	Som	Old English	
noun	folcwiga	poetry	rare	BT		Old English	
noun	fyrdesne	rare	BT	Sos	LM	Som	Old English

Figure 8.1.: List in Evoke of Old English words denoting “A warrior, fighter”, annotated with custom labels indicating dictionaries in which each word has been recorded.

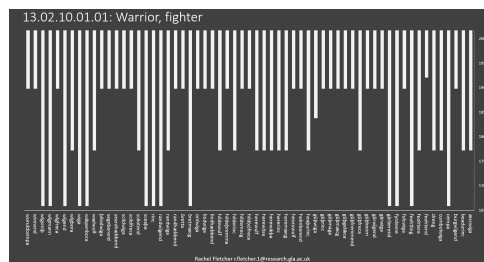


Figure 8.2.: Bar chart of Old English words denoting “A warrior, fighter” and the dates from which they were first recorded in an Old English dictionary.

Fletcher’s work foregrounds the value of custom labels for exploring aspects besides stylistics or diachronic developments of a language. Her case study illustrates the use of thesauri for the study of lexicographic practices, although such explorations are not without challenges. Variation in spelling amongst the dictionaries, differences in the granularity of sense definitions, and reinterpretation of words as belonging to another semantic domain can hamper straightforward application of labels and alter what can be deduced from subsequent analysis. Even so, Fletcher’s case study demonstrated that *TOE* can provide a semantic framework under which concepts and entire domains can be scrutinized in terms of their lexicographic developments. Annotation functionality in Evoke enables viewing and sharing such findings. The chart in Figure 8.2 indicates a need for further visualizations beyond those currently available in Evoke, and potentially including temporal information, that may inform future development of the web application.

## 8.4. Stylistics of Old English writing

Historical language thesauri can also be valuable resources for exploring stylistics of specific texts, authors from the historical period, or entire genres. Two EEMEE case studies demonstrated this avenue of research for the Old English context through enrichment of the *TOE* dataset with custom labels. Their subsequent analyses in Evoke have led to new insights into the language use in specific texts



(*Beowulf*, *Andreas*, the *Old English Martyrology*) or by a specific author (Ælfric of Eynsham). These case studies are discussed in separate subsections below.

### 8.4.1. Onomasiological profiles of Old English texts

In his case study, Thijs Porck demonstrated how Evoke can be used to investigate the lexis used in individual Old English texts.<sup>20</sup> By tagging all words that occur in *Beowulf*, *Andreas*, and the *Old English Martyrology*, Porck expanded the Linguistic Linked Data version of *TOE* with information that allows users to navigate the vocabulary in these specific texts through the semantic hierarchy of the thesaurus. In effect, this process has created three workable prototypes of thesauri, specific to the selected texts, which have been made publicly available in Evoke for anyone to browse and analyse (see Appendix 8.A).

By combining the *TOE* dataset in Evoke with additional information on individual texts, Porck was able to navigate and analyse the vocabulary use in these medieval texts. Figure 8.3, for example, lists Old English words denoting “A man, warrior” and the labels assigned to them. Overviews such as these allow differences in word use between Old English texts to be discerned. The word *ceorl*, for instance, was shown to occur in the poem *Beowulf* but not in *Andreas* or the *Old English Martyrology*, whereas *þegn* occurs in all three texts.

• A man, warrior		i	≡	↶	☁	★
noun	ceorl	beowulf	Old English			
noun	dædfruma	andreas	beowulf	poetry	Old English	
noun	hildestapa	andreas	hapax	poetry	rare	Old English
noun	hyse	andreas	beowulf	Old English		
noun	magorinc	beowulf	poetry	Old English		
noun	mann	andreas	beowulf	oemartyrology	Old English	
noun	swān	questionable	Old English			
noun	þegn	andreas	beowulf	oemartyrology	Old English	
noun	wiga	andreas	beowulf	Old English		
noun	wigend	andreas	beowulf	Old English		

Figure 8.3.: List in Evoke of Old English words (and their labels) denoting “A man, warrior”.

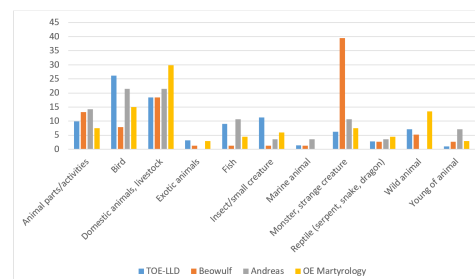


Figure 8.4.: Onomasiological profiles of various Old English texts for the semantic field of “Animal”.

Additionally, the combined information in this case study can be used to create semantic fingerprints, or ‘onomasiological profiles’, of individual Old English texts through the statistics provided by the interface of Evoke. In contrasting such profiles for individual Old English texts, Porck showed how new and distinctive patterns of vocabulary use can be brought to light. To illustrate, Figure 8.4 captures vocabulary use in the semantic field of “Animal” across the entire corpus of Old English and within the three different Old English texts,

<sup>20</sup>Porck, ‘Onomasiological Profiles of Old English Texts’.

specifically. Each text, Porck argues, “shows distinctive patterns of vocabulary use that can be related to the roles that animals play in the individual texts”.<sup>21</sup> For instance, most of the words used in the *Old English Martyrology* within this semantic field are found in the categories of “Domestic animals, livestock”, “Bird” and “Wild animal”, which matches scholarly observations on medieval hagiography. As Porck explains, “typically, saints are served by domestic animals and birds, while wild animals either miraculously come to their aid or represent devils in disguise”.<sup>22</sup> *Beowulf* displays a vastly different picture. By far, the category of “Monster, strange creature” contains the majority of the words used within this field, which is unsurprising given the prominent role of monsters within the poem.

On a higher, more abstract level, Porck employed onomasiological profiles of the three Old English texts to obtain insights into aspects that include their poetic characterization and degree of ambiguity. On the first aspect, Porck remarks that 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”.<sup>23</sup> On the degree of ambiguity found in the texts, which is based on the number of polysemous senses attributed to each lexical entry, he notes that texts marked with a higher level of words found solely in poetic contexts (i.e., *Beowulf* and *Andreas*) are more likely to be low in polysemous senses than other texts (i.e., the *Old English Martyrology*) and yield a lower degree of ambiguity.<sup>24</sup>

Porck’s use of Evoke and *TOE* exemplifies the value of linking resources and how combining sets of information allows researchers to obtain a picture hitherto unavailable. In fact, an onomasiological analysis of these Old English texts would not go amiss in the introductions to their respective scholarly editions. A case in point is the fourth edition of *Klaeber’s Beowulf*, which contains information on, amongst others, the diction, orthographic characteristics, and narrative structure. This material positions the poem in its literary and linguistic context — a positioning which would benefit from insights into the semantic domains present in this text. Contrasts with contemporary texts, across genres, and diachronic developments could, over time, be added to such sections on onomasiological characteristics. Indeed, much remains to be explored. Only a few selected fields have been scrutinized, thus far, using the analysis functionality of Evoke. The textual thesauri fashioned by Porck, publicly available in Evoke, will undoubtedly lead to a deeper understanding of these and other texts in the future.

---

<sup>21</sup>Ibid., p. 378.

<sup>22</sup>Ibid., p. 379.

<sup>23</sup>Ibid., p. 370.

<sup>24</sup>Ibid., pp. 371-2.

### 8.4.2. Exploring ‘Ælfrician’ vocabulary

Amos van Baalen used Evoke in a manner similar to that of Porck.<sup>25</sup> However, rather than creating onomasiological profiles of texts, van Baalen fashioned one for an individual author: Ælfric of Eynsham (c.955x957–c.1010). This abbot was one of the most prominent and prolific writers of the period. His vocabulary – and its influence on the surviving Old English lexis – is therefore important for our understanding of the language and culture of early medieval England. In researching Ælfric’s vocabulary, van Baalen employed the *Dictionary of Old English* and prior scholarship to identify and categorize the lexis that is characteristic for the works of Ælfric. The results were used to establish an onomasiological profile in Evoke of the abbot’s characteristic vocabulary.

Rather than tagging all words known to have been used by Ælfric, van Baalen first established and categorized those words that had been identified as being representative of ‘Ælfrician’ vocabulary. Words that are predominantly used by Ælfric (such as *cāsus* ‘case’ and *forþearle* ‘very much, greatly’) were marked in Evoke and subsequently analysed in order to create an onomasiological profile, which allowed van Baalen to comment on such aspects as the ambiguity, synonymy, specificity, and semantic distribution of the words characteristic of Ælfric’s writing (see Figure 8.5 and 8.6).<sup>26</sup> The use of custom labels in Evoke ensured that the various categories established by van Baalen could be represented that convey the frequency of a word’s use by Ælfric, as recorded in *DOEC*, compared to its frequency in surviving Old English texts by other authors.

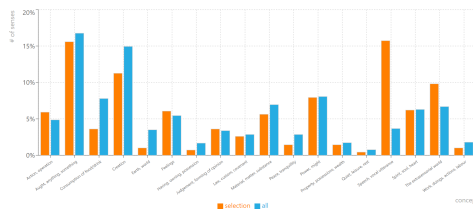


Figure 8.5.: Semantic distribution of Ælfrician vocabulary contrasted with that of all Old English lexis.

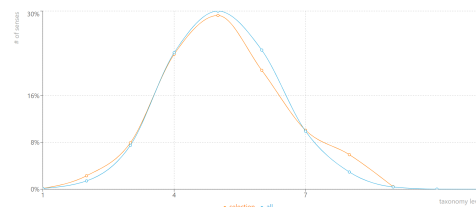


Figure 8.6.: Specificity of Ælfrician vocabulary contrasted with that of all Old English lexis.

Van Baalen’s work with Evoke establishes a working method to quantify and pinpoint distinctive authorial preferences within semantic domains. His approach demonstrates that the annotation system of Evoke is an effective means for scholars to add information to a thesaurus and to view as well as query the results. An exciting aspect of this case study is that the categorizations for Ælfrician vocabulary, added by the scholar as labels to the lexis recorded in *TOE*, hint towards future directions in which thesauri and corpora are more

<sup>25</sup>Van Baalen, ‘Identifying, Categorising and Exploring “Ælfrician” Vocabulary using the *Dictionary of Old English*, *A Thesaurus of Old English* and Evoke’.

<sup>26</sup>A list of all words found to be typical of Ælfric’s writings, listed under the various categories established in this case study, is included as an appendix to the article by van Baalen.

closely integrated. Once word attestations are linked to their corresponding entries in thesauri, onomasiological profiles can consider matters of frequency and predominant usage.<sup>27</sup> Moreover, such an integration would, effectively, allow onomasiological profiles to be generated for authorial preferences, specific texts, and even entire genres.

## 8.5. Diachronic developments of Old English

In its current, traditional form, *TOE* does not allow researchers to study changes across the Old English period, since, as Richard Dance points out, the thesaurus treats its items as “a single geographically and temporally indistinguishable mass”, ignoring diachronic differences between, e.g., Early West Saxon and Late West Saxon.<sup>28</sup> However, Evoke’s functionality to add custom labels to words does allow for diachronic investigations by incorporating such distinctions. Information on the origins of words can be added, too, in a similar manner. Jane Roberts asserts that such practice, e.g., flagging Latin loan words or those with an Old Norse origin, “could lead to the examination of loan translation and word formation, dialect vocabulary, etc.; and Evoke provides a welcome platform for such undertakings”.<sup>29</sup> The following subsection discusses one approach towards diachronic investigations, undertaken by Khan et al., who have mapped conceptual variation in *TOE* through the web application Evoke.

### 8.5.1. Mapping conceptual variation and diachronic changes

The case study by Khan et al. explored how Evoke may be used to visualize, navigate, and investigate conceptual mappings within the Old English vocabulary of SHAME.<sup>30</sup> The authors, after having established a list of 28 words and expressions for SHAME based on *TOE* and other lexicographic resources, charted conceptual variation in this vocabulary through two classifications. Firstly, each word or expression was classified as a metaphorical mapping. For instance, the verb *ārēodian* (“to blush, be ashamed”, lit. “to turn red”) belongs to the mapping EMOTION IS REDNESS IN THE FACE. Secondly, the development of each word or expression for SHAME was classified into one of four diachronic scenarios based on whether or not a sense shift must have taken place between Proto-Germanic and Old English.

The results of their classifications were captured in Evoke using its annotation system. The Old English SHAME words and expressions in *TOE* were tagged with their metaphor mapping (including their source and target) and the diachronic scenario to which they belong (see Figure 8.7). Khan et al. showed that these annotations facilitate the discoverability and further analysis of conceptual

---

<sup>27</sup>Standardizing the modelling of frequency and attestations in Linguistic Linked Data is currently being pursued (Chiarcos et al., ‘Modelling Frequency and Attestations for OntoLex-Lemon’).

<sup>28</sup>Dance, Review of *TOE1* (p. 313).

<sup>29</sup>Roberts, ‘*A Thesaurus of Old English: The Pilot Study for the Glasgow Historical Thesaurus*’, p. 312.

<sup>30</sup>Khan et al., ‘Mapping Conceptual Variation through *A Thesaurus of Old English* and Evoke’.

variation in Old English figurative language. The web application generated overviews of entire mappings or specific diachronic developments on demand (see Figure 8.8). The authors indicated that they intend to explore these research avenues further in future work.

The screenshot shows the Evoke interface for the verb 'ārēodian'. It displays the lexical sense 'ārēodian' with its type 'verb' and a note that it belongs to the entry 'ārēodian'. Below this, there are several semantic mappings listed under the heading 'Annotation'. These mappings include:

- #MAPPING\_Emotion\_is\_nakedness\_in\_face
- #SEMANTICS\_resultative\_metonymy
- #ETYMOLOGICAL\_THEME\_ārod
- #TARGET\_Shame
- #SOURCE\_Redness\_in\_face

Each mapping is accompanied by a small icon and a link to 'Old English'.

Figure 8.7.: A lexical sense annotated in Evoke with semantic mappings.

The screenshot shows an overview of mappings for the sense 'EMOTION IS NAKEDNESS'. It displays a list of mappings for various Old English words, including 'æppen', 'æppenes', 'æwicc(e)', 'æwicnes', 'æwisc', 'æwiscberend', 'æwisclic', and 'æwiscmōd'. Each mapping is shown with its source and target terms, along with a small icon and a link to 'Old English'.

Figure 8.8.: Overview in Evoke of senses with the metaphorical use EMOTION IS NAKEDNESS.

As Khan et al. have demonstrated, thesauri are valuable resources for research on metaphors. Indeed, the value of these lexicographic resources for identifying and charting metaphorical relations between semantic fields has been reported in other research projects, too.<sup>31</sup> Additionally, Khan et al. revealed advantages to their having used Evoke. The annotation system of the application and the overviews generated offered the means to interact more closely with the source material. A thesaurus edition can thus not only be a source of information but also act as a canvas for information produced. Such interaction with lexicographic resources may well facilitate feedback loops to provide the original editors with useful comments and suggestions that can be incorporated in revisions.

## 8.6. Comparative analyses of Old Germanic languages

Two further EEMEE case studies demonstrated the value of thesauri for comparative analyses of related languages. The first, by Rita van de Poel and Sander Stolk, linked Old Frisian lexis with that of Old English captured in the *TOE* dataset. The second case study, by Katrien Depuydt and Jesse de Does, connected Old Dutch to Old English. The resulting data connections provide an avenue of research for contrasting kindred languages through the overarching, onomasiological framework of the thesaurus. Thus, commonalities and differences between languages can emerge based on their representation of specific semantic fields.

<sup>31</sup> *Mapping English Metaphor through Time.*

### 8.6.1. Old Frisian and Old English KINSHIP

The case study by Rita van de Poel and Sander Stolk used Evoke to link Old Frisian to Old English words captured by the *TOE* dataset.<sup>32</sup> In their work, the authors focused on the semantic field of KINSHIP in the two related languages. In connecting Old Frisian lexis, drawn from a dictionary of Old Frisian, to the overarching structure of *TOE*, the researchers created a dataset that positions Old English and Old Frisian lexis in the same semantic framework. The connected resources were shared and analysed using Evoke (see Figure 8.9).

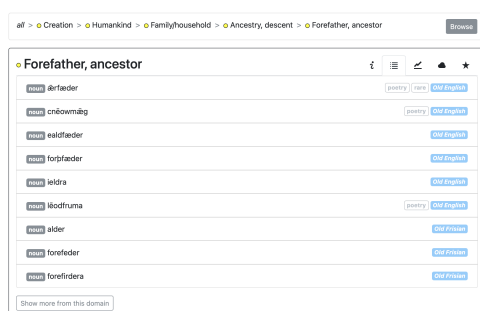


Figure 8.9.: An overview in Evoke of Old English and Old Frisian words denoting “Forefather, ancestor”.

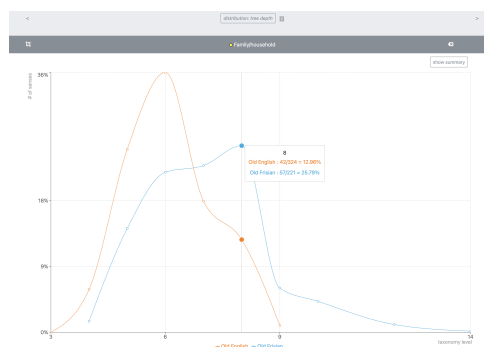


Figure 8.10.: Levels of specificity for Old Frisian and Old English KINSHIP.

Van de Poel and Stolk demonstrated how Evoke facilitates comparative analyses between the two languages and argue that onomasiological analyses (such as shown in Figure 8.10) can uncover new linguistic and cultural insights. Their case study revealed how the two language communities conceptualized aspects relating to family, ancestry, descent, and adoption.<sup>33</sup> Additionally, van de Poel and Stolk signaled that new thesauri can be created through adoption (and expansion) of the taxonomy of an existing thesaurus of a kindred language — one semantic field at a time. Evoke and the Linguistic Linked Data paradigms support such endeavours, as they allow researchers to access and extend *TOE* data whilst abiding by the license of the thesaurus, which restricts users from extracting or downloading substantial portions of the dataset.

In addition to presenting their findings, van de Poel and Stolk shed light on the major hurdles and limitations in connecting material from two lexicographic resources and analysing the results.<sup>34</sup> Comparisons such as theirs, they note, are influenced by differences in lexicographic practices and editorial choices. The editors of *TOE* and those of the dictionary of Old Frisian will have had, or worked with, different approaches towards the granularity of word senses to include.<sup>35</sup>

<sup>32</sup>Van de Poel and Stolk, ‘A Case of Kinship’. The article in its entirety is included in this dissertation and follows the current chapter.

<sup>33</sup>To illustrate, “Adoption” lacks Old Frisian senses entirely, whereas the concept of “Siblings” is lexicalized in this language but not in Old English (Ibid., p. 475).

<sup>34</sup>Ibid., p. 477.

<sup>35</sup>The practice of maintaining only generic word senses is not uncommonly referred to as

Additionally, positioning the Old Frisian lexis in *TOE* categories was complicated by the use of different languages for the sense definitions (i.e., English in *TOE* versus German for the Old Frisian lexis). Nevertheless, the researchers found ways to overcome these issues and succeeded in aligning the two lexicographic resources, thereby adopting the overarching structure of *TOE* as a comparative framework for the two related historical languages. Thus, the scholars could avoid fashioning an onomasiological structure for Old Frisian from the ground up and subsequently, prior to analysis, attempting to reconcile it with that of *TOE*. The resulting, combined dataset allows researchers of both Old Frisian and Old English to adopt the same structure and to identify and discuss notable differences in conceptualizations between the kindred languages through shared semantic domains.

### 8.6.2. Linking Old Dutch to Old English

Katrien Depuydt and Jesse de Does took a more technical route than Rita van de Poel and Sander Stolk towards linking lexis from another language to *TOE*.<sup>36</sup> In their case study, they reflected on their experiments with manual and automatic techniques in order to establish connections between entries from the *Dictionary of Old Dutch* and the *TOE* dataset available in Evoke. In doing so, they investigated whether the existing macrostructure of *TOE* can be reused to create a thesaurus of Old Dutch.

Their research compared two approaches to linking the two lexicographic resources. The first approach was a manual one, in which the tool Lex'it was used to view entries of Old Dutch words and record appropriate categories from the *TOE* hierarchy for their registered senses (see Figure 8.11). Additional matches were obtained by drawing on cognate information. Cognate relations of Old Dutch entries were traversed to Old Frisian ones that had already been linked to *TOE* categories (based on the work by van de Poel and Stolk described in section 8.6.1). The second approach explored by the authors was semi-automated linking, where five different matching algorithms were contrasted and evaluated. Depuydt and de Does concluded that further directions for linking such lexicographical resources may be found in improving automated linking techniques, crowdsourcing part of the process, and improving visualization of candidate links for verification purposes.

The case study demonstrated two key advantages of the Linguistic Linked Data form in which *TOE* has been made available. Firstly, connections between Old Dutch lexis and *TOE* categories could be established and stored separately from the *TOE* dataset. This possibility ensures that researchers other than the editors of *TOE* can extend the semantic framework formed by the thesaurus, even when they do not have direct access to the raw data of the thesaurus or to the database in which it is stored. Secondly, information linked in previous

---

lumping, whereas distinguishing fine-grained senses is termed splitting (see Fontenelle, 'Bilingual Dictionaries', pp. 45-6; Kay and Alexander, 'Diachronic and Synchronic Thesauruses', p. 370).

<sup>36</sup>Depuydt and de Does, 'Linking the *Dictionary of Old Dutch* to *A Thesaurus of Old English*'.

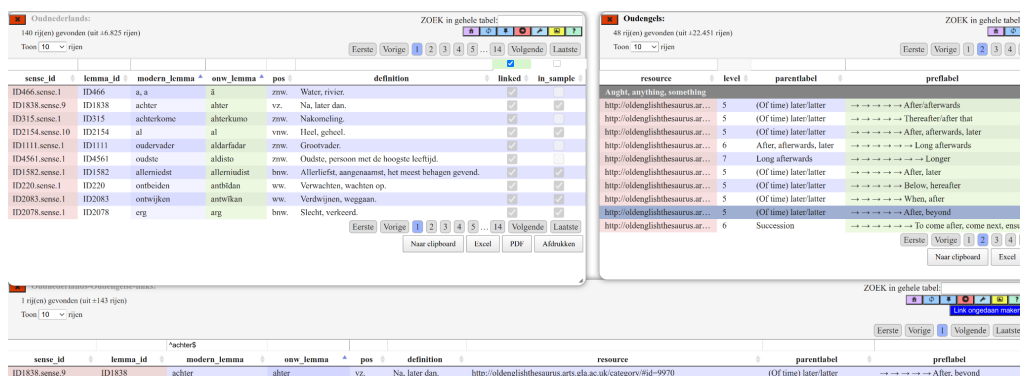


Figure 8.11.: The Lex'it-based linking tool; linking the entry for Old Dutch *achter* to the *TOE* category “05.11.07.04.02|13 (Of time) later/latter: After, beyond”.

research could be queried alongside the original thesaurus material and, thus, drawn on for analyses. In this case, the Old Frisian lexis for KINSHIP, positioned in thesaurus categories by van de Poel and Stolk (see 8.6.1), could be combined with cognate information and leveraged for positioning Old Dutch lexis within the onomasiological structure of the thesaurus.

A dataset containing the established connections between the Old Dutch dictionary material and *TOE* has been made available on a webpage dedicated to the case study.<sup>37</sup> The availability of this dataset enables researchers to perform comparative analyses of Old Dutch and Old English. In their article, Depuydt and de Does do not yet include such an analysis. Their insights on the reuse of an existing onomasiological framework – i.e., that of *TOE* – are therefore limited to the difficulties encountered in creating suitable connections. The language-specific nature of *TOE*, built as it was for “fitting the intricacies of Old English lexis”, is one of the reasons the authors put forward for the complexity of matching items from other languages.<sup>38</sup> Additionally, an unfamiliarity with the structure of the thesaurus can impede locating the categories that are of interest. Indeed, Depuydt and de Does indicate that manually assigning Old Dutch lexis to suitable categories within the onomasiological structure of *TOE* was hampered for this reason.<sup>39</sup> Therefore, studies that are sizable in scope may warrant gaining an intimacy with the topical system of a thesaurus before adopting it for other languages and comparative analyses.

## 8.7. Teaching Old English language and culture

The last aspect of historical language thesauri explored in EEMEE was their didactic potential. Kees Dekker, at University of Groningen, and Thijs Porck and Sander Stolk, at Leiden University, employed *TOE* and *Evoke* to introduce students to notable aspects of Old English language and culture. In both cases,

<sup>37</sup><https://ivdnt.org/corpora-lexica/diamant/onw-toe-linking/>

<sup>38</sup>Depuydt and de Does, ‘Linking the *Dictionary of Old Dutch* to *A Thesaurus of Old English*’, p. 509.

<sup>39</sup>*Ibid.*, p. 510.



Bachelor students were asked to discern nuances within semantic domains and theorize as to how these might reflect the culture of the Anglo-Saxons. The outcome of the experiment was that students were shown to benefit from extensive use of these resources throughout a course, as is the case with University of Groningen, but also from more concentrated use, such as in workshops at Leiden University. Experiences of the two universities are treated below in separate sections.

### 8.7.1. The Old English classroom at University of Groningen

In his case study, Kees Dekker employed *TOE* and *Evoke* in a classroom context through a series of increasingly complicated search assignments.<sup>40</sup> These assignments, intended to interest, challenge, and instruct students, were preceded by an introduction into lexical semantics to inform students of such notions as semantic fields, hyponymy, and synonymy — notions at the core of *TOE* and other thesauri. During a seven-week module, students received assignments on *TOE* and *Evoke* that were related to the Old English texts they were asked to read and translate. The assignments ranged from, at the start of the module, navigating the thesaurus to take note of the lexis and distinctions available for bodies of water (see Figure 8.12) to, at a more advanced stage, questioning the notion of ‘foreign’ in Old English (see Figure 8.13) by discerning what the early medieval speakers of this language considered foreign and what connotations their words for this concept had. Through his discussion of both the assignments for students and their responses, Dekker demonstrated that *Evoke* is a valuable addition to a teacher’s tool set, as it invites students and researchers alike to explore Old English in novel and interesting ways.

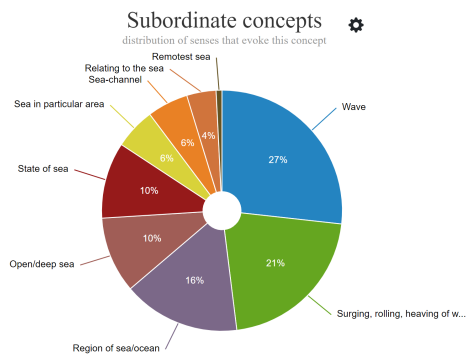


Figure 8.12.: Pie chart in *Evoke* showing the distribution across the subordinate concepts of the *TOE* category “Sea/ocean”.

Word	Part of Speech
efoelendisc	adjective
efoende	adjective
efoetidig	adjective
efoetidiglic	adjective
efoetidisc	adjective
foende	adjective
focumende	adjective
fencip	adjective
foacumen	adjective
foacund	adjective
foancumen	adjective
foera	adjective
foende	adjective
foendisc	adjective
foetic	adjective
foelisc	adjective

Figure 8.13.: List of Old English words in *Evoke* denoting “foreign”.

Although the requirements for *Evoke* were formulated to support research first and foremost,<sup>41</sup> Dekker’s article demonstrates that the design of the web

<sup>40</sup>Dekker, ‘*Evoke* and *A Thesaurus of Old English* in the Old English classroom’.

<sup>41</sup>Stolk, ‘*Evoke*’.

application enables students, too, to explore a thesaurus and engage with the lexical material presented. The functionality to navigate and view thesaurus material supports users without expertise or prior knowledge of Linguistic Linked Data or thesauri. Moreover, the many visualizations included in Evoke allow students to engage with the material in a playful manner: Next to pie charts, which provide statistics beneficial to research, wordcloud visualizations of entire semantic fields (see Figure 8.14) have led to students calling Evoke “fun” as well as “useful” (see Appendix 8.C, Table 8.C.6).

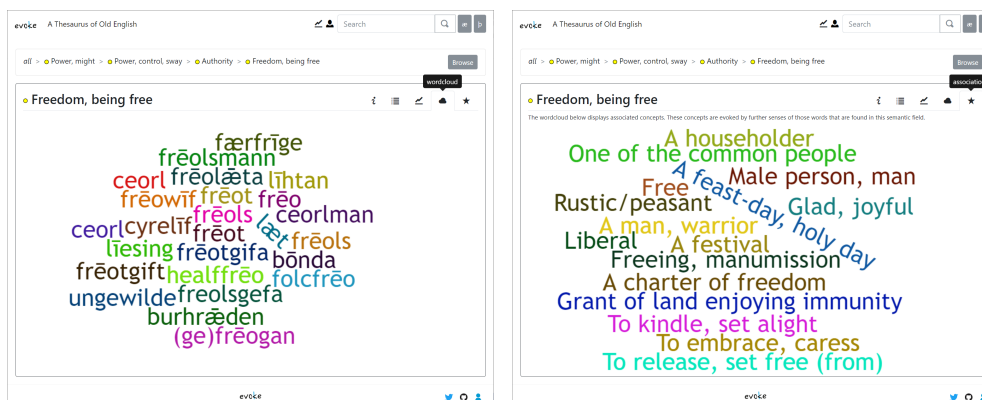


Figure 8.14.: Wordcloud in Evoke for “Freedom, being free”.  
 Figure 8.15.: Associations in Evoke for “Freedom, being free”.

For future use of Evoke in educational settings, Dekker points out that, although already valuable, the web application could be improved by incorporating “a repository of questions and assignments shared by instructors interested in the use of Evoke for curricular and research purposes”.<sup>42</sup> Such an addition is indeed already planned. Existing teaching material will be incorporated in the Evoke website. Preliminary work has been done towards its realisation. By courtesy of Prof. Carole Hough (University of Glasgow), the exercises will include units from the module *Learning with the Online Thesaurus of Old English*.

### 8.7.2. Workshops on Old English resources at Leiden University

Dekker’s findings on the use of Evoke in education are reinforced by those at Leiden University, where the web application has been used in a classroom setting, annually, since 2018. Third-year Bachelor students perform a number of assignments in Evoke during a 2-hour workshop on digital tools and resources for studying Old English, led by Thijs Porck and Sander Stolk. The first of these workshops, which took place on 13 November 2018, included an evaluation by students of the application (see Appendix 8.C). Students considered navigation in Evoke to be “clear” and “easy” to understand. Moreover, views and analyses offered by the application enabled them to obtain results to the following questions, amongst others, in an “efficient” manner:

<sup>42</sup>Dekker, ‘Evoke and A Thesaurus of Old English in the Old English classroom’, p. 526.

How many words does Old English have for Hell and its subcategories? How about for Heaven?

The statistics tab in Evoke provides exactly such figures at a glance, avoiding the need for students and researchers to manually count the words listed under various categories.

A recurring element in assignments for students at both Leiden University and University of Groningen has been to establish connotations of words or semantic fields in their entirety. A case in point is the following illustrative question:

The *TOE* lists several words for old age, including *gelefed* ‘old’, *har* ‘old’ and *frod* ‘old’. Find out what else these words mean. Search for “gelefed”, “har” and “frod”. What, according to your findings, did the Anglo-Saxons associate with old age?

Students viewed each word listed separately to ascertain their various senses and, thus, indications of connotations of words used in a certain meaning. During the second EEMEE workshop, it became clear that obtaining connotations of an entire semantic field is deemed valuable for education purposes but, with the then available functionality of Evoke, still required students to engage with content in this piecemeal fashion. A new visualization was therefore introduced. The ‘associations’ wordcloud, available since version 1.4.1 of Evoke, provides an overview of all senses in a semantic field related through polysemy (see Figure 8.15). In short, feedback on the use of Evoke and *TOE* in education has been valuable in informing the development of Evoke.

## 8.8. Discussion

This section discusses the usefulness of the digital resources at the heart of the case studies presented above: the Linguistic Linked Data form of *TOE* and the web application Evoke. Both resources have been developed specifically for facilitating research, which raises the question in what manner, if any, they have enabled new lines of scholarly enquiry compared to existing editions of historical language thesauri.

### 8.8.1. A *Thesaurus of Old English* as Linguistic Linked Data

The Linguistic Linked Data form of *TOE* (or *TOE*-LLD), as made available in Evoke, enabled researchers to extend the original dataset of the thesaurus with information salient to their inquiries. In contrast to many existing editions of historical language thesauri of Scots and English, this digital form includes unique identifiers (IRIs) for lexical items as well as categories.<sup>43</sup> These identifiers

<sup>43</sup>To illustrate, *HTE* does not issue IRIs for a specific lexical sense. The search hit *strength* as found at category “02.02.06.01|06 (n.) Strength of evidence”, for example, refers to the following subpath on the *HTE* website: </category/?id=123117&qsearch=strength>. This web address indicates that category 123117 is to be shown, i.e., “Strength of evidence”, whilst highlighting any text that matches the string ‘strength’. The web address is therefore not an exact reference to the specific item requested, but a visual aid that potentially foregrounds

have facilitated explicit references and, in the case of *TOE*-LLD, are unaffected by potential modifications to the spelling of the headword. Being able to reference, view, and link information to lexical items has proven valuable for research in the case studies described above. Moreover, the inclusion of lexical entries in *TOE*-LLD allows researchers to mark items of interest more broadly than per individual lexical sense (cf. the case study by Thijs Porck). These entries, deduced automatically from contextual information available in the original *TOE* dataset, may still need manual revision in places: some homonyms have been lumped together and senses of lexemes remain ungrouped when their headwords use inconsistent spelling or grammatical forms.<sup>44</sup> To eliminate the inherent caveats of such an automated process, thesaurus editors may wish to maintain lexical entries explicitly in order to facilitate these research needs.

Another important advantage in adopting IRIs is that their use simplified extending *TOE* content without having access to the raw data of the thesaurus. Scholars were able to capture knowledge and, in their data, include explicit references to information found in *TOE*-LLD. This capability is advantageous for research, since the license of *TOE* prohibits users from downloading large amounts of thesaurus data.<sup>45</sup> Licenses that restrict access in such a manner are not uncommon for lexicographic resources available on the Web.<sup>46</sup> The *TOE*-LLD version thus lowers the barrier for scholars to engage with the available material.

In addition to the usefulness of IRIs for linking data, even when restrictive licenses are in place, Christian Chiarcos et al. claim two further benefits to use of Linguistic Linked Data.<sup>47</sup> The first is an increased level of interoperability owing to the reuse of established data vocabularies (i.e., Lemon-OntoLex and SKOS).<sup>48</sup> Improved interoperability manifests itself in the shape of software other than Evoke being able to operate on the same data and/or data form. Examples of tools used in the various case studies include custom alignment tools, Microsoft Excel, and Lex'it.<sup>49</sup> Compatible software beyond those employed in the case studies exists, too. Cases in point are VocBench 3 and LexO, with which users can fashion and maintain such data.<sup>50</sup> Current endeavours with Linguistic Linked Data suggest an uptake of this digital form and tools that can operate on it over the next few years. These include the H2020 projects ELEXIS (2018–22), Prêt-à-LLOD (2019–22), and the COST Action NexusLinguarum (2019–23).<sup>51</sup>

---

other entries, too, that include the aforementioned string. For further information on IRIs, see Chapter 3.

<sup>44</sup>See Porck, 'Onomasiological Profiles of Old English Texts', p.366.

<sup>45</sup>See *TOE4*, section 'Copyright and Conditions of Use'.

<sup>46</sup>See, for instance, *Oxford English Dictionary Online*, *HTE*, and the historical Dutch dictionaries part of the *Geïntegreerde Taalbank*.

<sup>47</sup>Chiarcos et al., 'Towards Open Data for Linguistics: Lexical Linked Data'.

<sup>48</sup>See Chapter 3.

<sup>49</sup>These tools are described in Chapter 6, section 4.3.2, 'Linking Data'. The 'List of source code' in the back matter of the dissertation includes relevant links to code developed for software applications and data transformations, which includes the custom alignment tools.

<sup>50</sup>Stellato et al., 'VocBench 3'; Bellandi and Giovannetti, 'Involving Lexicographers in the LLOD Cloud with LexO'.

<sup>51</sup>ELEXIS: <https://cordis.europa.eu/project/id/731015>, 2018–2022; Prêt-à-LLOD: <https://>

Moreover, the form should allow thesauri other than *TOE* (and its extensions) to be made accessible in Evoke in the future without reconfiguration of the database or alterations to the source code to accommodate the new thesaurus.

The second advantage of LLD over traditional Web-based storage, Chiarcos et al., argue, is the ability to merge distinct datasets and thereby obtain a validly formatted, combined set of information. The characteristic of LLD in merging data, thus, allows for complementary data to be queried in unison or different perspectives to be related. The case studies above displayed the use of this functionality in complementing *TOE* with Old Frisian lexis and cultural concepts (see van de Poel and Stolk) or Old Dutch lexis (see Depuydt and De Does). Similarly, annotations and labels can exist in a dataset separate from *TOE* and nonetheless be viewed and analysed cohesively (see Porck; van Baalen). Achieving these ends with the traditional database of *TOE* would be possible too, albeit with a higher effort. Modifications to the database structure can add support for other languages, annotations, and custom labels, but such changes will need to be integrated into the source code underlying the presentation. In short, the proclaimed advantages of LLD have facilitated the case studies described in this chapter and may benefit future studies in a similar manner.

The advantages of the LLD form of *TOE* notwithstanding, use of this resource for research is not necessarily unproblematic. The EEMEE case studies revealed a number of hurdles in connecting additional data to *TOE*. Firstly, differences can be expected between the lexicographic choices that underlie the thesaurus and additional data stemming from another lexicographic resource.<sup>52</sup> Secondly, *TOE* contains conceptualisations, in the form of categories positioned in its topical system, which may work well for the Old English lexis for which the thesaurus has been built from the ground up, but may not be wholly adequate for organizing material from other languages: Concepts required for the additional material may be absent in the existing onomasiological framework (e.g., “Siblings” in the case of Old Frisian) or structured differently from what the additional material would find most useful.<sup>53</sup> Lastly, even when an appropriate category is available in which to position lexis, locating that category is not always straightforward when one is not intimate with the thesaurus structure.<sup>54</sup> Although automation of aligning such resources may help speed up the process, existing techniques thereto are not perfect.<sup>55</sup> These challenges are, of course, not brought about by the use of LLD as digital form of historical language thesauri, but are nonetheless relevant for employing *TOE*-LLD for research.

---

cordis.europa.eu/project/id/825182, 2019–2022; NexusLingarum: <https://www.cost.eu/actions/CA18209,2019/T1/textendash2023>.

<sup>52</sup>E.g., Porck, ‘Onomasiological Profiles’, pp. 364-9; van de Poel and Stolk, ‘A Case of Kinship’, pp. 477-8.

<sup>53</sup>Ibid.

<sup>54</sup>Depuydt and de Does, ‘Linking the Dictionary of Old Dutch to A Thesaurus of Old English’, pp. 509-10.

<sup>55</sup>Ibid., p. 508.

### 8.8.2. Evoke and its functionalities

Chapter 2 identified functionalities for editions of historical language thesauri that would benefit research. Some of these functionalities – i.e., mostly those surrounding navigation and resource views – could already be found in existing editions of thesauri of Scots and English. In contrast, nearly all functionality geared towards extension of original thesaurus content, statistical analyses, and data management is novel.<sup>56</sup> All five groups of functionality were employed in the case studies reviewed in this chapter. The prospect of novel features becoming available to them encouraged researchers to formulate new lines of questions, ones that hinged on either the ability to insert custom labels or the means to position lexis from languages other than Old English in *TOE*. A suite of advanced analysis options, which utilize the topical structure of the thesaurus, provided researchers with insights on their own information complementing the original thesaurus content, effectively creating onomasiological profiles.

Not all functionality listed in Chapter 2 is currently available in Evoke. Most notably, the ability to filter words when browsing, based on user-specified word features, has not been implemented yet. Such functionality would bring about focused views of subthesauri, including the ‘Beowulf Thesaurus’, ‘Ælfrician Vocabulary’ and ‘Old Frisian: Kinship’. Users of these datasets (created in the case studies by Porck, van Baalen, and van de Poel and Stolk) would benefit from not having to inspect the characteristics of words found at a specific category (e.g., their labels or language) in order to determine whether they belong to the subthesaurus in question.

Additionally, exploration of thesaurus content could be enhanced further. To illustrate, a top 10 of categories containing lexical items conforming to a selection of features salient to the user would facilitate localizing points of interest in the thesaurus structure. Similarly, graphs generated by analyses in Evoke could be supplemented with samples of words or word senses that constitute the figures presented in that graph, thereby making results more tangible and assisting the user in identifying inclusion of any undesired material that warrants refinement of the initial query.

Beyond improvements for exploration, the case study by van Baalen demonstrates another future direction for work on thesaurus editions. Vocabulary characteristic of Ælfric could be identified with higher precision by linking attestations of lexical items in corpora and determining their frequency of occurrence. Such insights would, of course, not be advantageous solely in the context of Ælfric’s writings. A selection of single texts or entire genres would be made possible, too. Previous research has explored how such links between *TOE* and the corpus *DOEC* can be represented as Linguistic Linked Data.<sup>57</sup> With the availability of such links, analyses that traverse them and obtain relevant figures would be possible to realize.

The set of functionality identified in Chapter 2, alongside the features put

---

<sup>56</sup>Only functionality for basic analyses was provided by some editions of historical language thesauri (see Chapter 2).

<sup>57</sup>Chiarcos et al., ‘Modelling Frequency and Attestations for OntoLex-Lemon’, pp. 7-8.

forward in the previous two paragraphs, is by no means exhaustive. Additional workshops and case studies, increasing the diversity of disciplines and interests represented, are likely to elicit further functionality desired for research and education. Digital Humanities analyses of historical sources could, for example, be given more attention in future research programmes involving thesauri such as *TOE*. Through iterations of software use and development, tools available to researchers can improve incrementally and expedite expansion of our knowledge of historical languages and cultures.

## 8.9. Conclusion

This chapter has provided an overview of case studies, in research and education, that take advantage of the Linguistic Linked Data form of *A Thesaurus of Old English* and the web application Evoke. The combination of these resources facilitates enrichment of the original *TOE* dataset with information regarding, amongst others, lexicographical history, stylistics, diachronic developments, and kindred languages. Onomasiological analyses, possible through the user interface of Evoke, have led to new insights into Old English language and culture for both students and researchers. This is not to say that all research possible with *TOE* would benefit equally from the Linguistic Linked Data format and the functionality identified in Chapter 2. Smaller scale studies that focus on a single category, for instance, may not directly benefit from the availability of features that perform advanced analyses over the entire thesaurus. Functionality already covered by existing software may necessitate resorting to other, supported data formats. More importantly, the case studies and the resulting materials discussed in this chapter show that the development of digital research tools for historical language thesauri can be a powerful impetus for future research. The new resources at the centre of these studies – *TOE*-LLD and Evoke – have introduced new instruments, embraced by researchers and students participating in EEMEE, that should prove useful beyond the scope of this research project.

## References

- ‘About the Dictionary of Old English’, *The Dictionary of Old English*, ed. C. Monahan (Toronto, 2022). [https://doe.artsci.utoronto.ca/?page\\_id=11](https://doe.artsci.utoronto.ca/?page_id=11).
- Bellandi, A. and E. Giovannetti, ‘Involving Lexicographers in the LLOD Cloud with LexO, an Easy-to-Use Editor of Lemon Lexical Resources’, Proceedings of the 7th Workshop on Linked Data in Linguistics, Marseille, 22-3 June 2020, pp. 70–4. <https://aclanthology.org/2020.ldl-1.10.pdf>.
- Benedek, J. and T. Miner, ‘Measuring Desirability: New Methods for Evaluating Desirability in a Usability Lab Setting’, Proceedings of Usability Professionals Association 2002, Orlando, 8-12 July 2002, pp. 8–12.
- Bosworth, J., *A Dictionary of the Anglo-Saxon Language* (London, 1838).
- Bosworth, J. and T. N. Toller, *An Anglo-Saxon Dictionary Based on the*

- Manuscript Collections of the Late Joseph Bosworth* (London, 1898), *Supplement* by T. N. Toller (Oxford, 1921), with *Enlarged Addenda and Corrigenda* by A. Campbell (Oxford, 1972).
- Bouwer, H., *Studien zum Wortfeld um eald und niwe im Altenglischen* (Heidelberg, 2004).
- Bremmer Jr, R. H., 'Treasure Digging in the Old English Lexicon', Review of *TOE2*, *NOWELE* 40 (2002), 109–14.
- 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. Larratt Keefer, K. L. Jolly, and C. E. Karkov, Medieval European Studies XI (Morgantown, 2010), 204–32.
- Chiarcos, C. et al., 'Towards Open Data for Linguistics: Lexical Linked Data', in *New Trends of Research in Ontologies and Lexical Resources: Ideas, Projects, Systems*, eds. A. Oltramari, P. Vossen, L. Qin, and E. Hovy (Heidelberg, 2013), pp. 7–25.
- Chiarcos, C. et al., 'Modelling Frequency and Attestations for OntoLex-Lemon', Proceedings of the 2020 Globalex Workshop on Linked Lexicography, Marseille, 20 June 2020, pp. 1–9. <https://aclanthology.org/2020.globalex-1.1.pdf>.
- Clark Hall, J. R., *A Concise Anglo-Saxon Dictionary for the Use of Students* (London, 1894).
- Clark Hall, J. R., *A Concise Anglo-Saxon Dictionary*, 4th edn, with a supplement by H. D. Meritt (Cambridge, 1960).
- Dance, R., Review of *TOE1*, *Medium Ævum* 66.2 (1997), 312–13.
- David, A. and J. Simpson, 'The Middle Ages to ca. 1485: Introduction', *Norton Anthology of English Literature*, ed. S. Greenblatt, 2 vols, 8th edn (London, 2006), I, 1–21.
- Dekker, K., 'Evoke and *A Thesaurus of Old English* in the Old English classroom', *Amsterdamer Beiträge zur älteren Germanistik* 81.3-4 (2021), 514–29. doi: [10.1163/18756719-12340241](https://doi.org/10.1163/18756719-12340241).
- Depuydt, K. and J. de Does, 'Linking the *Dictionary of Old Dutch* to *A Thesaurus of Old English*: A First Exploration', *Amsterdamer Beiträge zur älteren Germanistik* 81.3-4 (2021), 493–513. doi: [10.1163/18756719-12340240](https://doi.org/10.1163/18756719-12340240).
- DOE = Dictionary of Old English: A to I Online*, eds. A. Cameron et al. (Toronto, 2018). <https://tapor.library.utoronto.ca/doe/>.
- DOEC = Dictionary of Old English Web Corpus*, eds. A. diPaolo Healey et al. (Toronto, 2009). <https://tapor.library.utoronto.ca/doecorpus/>.
- 'Europeana Impact Playbook', *Europeana*. <https://pro.europeana.eu/page/europeana-impact-playbook>.
- Fontenelle, T., 'Bilingual Dictionaries: History and Development; Current Issues', in *The Oxford Handbook of Lexicography*, ed. P. Durkin (Oxford, 2016), pp. 44–61.
- Geïntegreerde Taalbank* (Leiden, 2018). <https://gtb.ivdnt.org>.
- Görlach, M., Review of *TOE1*, *Anglia* 116.3 (1998), 398–401.
- Hall, A., *Elves in Anglo-Saxon England: Matters of Belief, Health, Gender and*



- Identity*, Anglo-Saxon Studies 8 (Woodbridge, 2007).
- HTE* = *The Historical Thesaurus of English*, eds. C. Kay et al. (Glasgow, 2014).  
<http://historicalthesaurus.arts.gla.ac.uk>.
- Kay, C. and M. Alexander, 'Diachronic and Synchronic Thesauruses', in *The Oxford Handbook of Lexicography*, ed. P. Durkin (Oxford, 2016), pp. 367–80.
- Kestemont, M. et al., 'Forgotten Books: The Application of Unseen Species Models to the Survival of Culture', *Science* 375.6582 (2022), 765–9.
- Khan, F. et al, 'Mapping Conceptual Variation through *A Thesaurus of Old English* and Evoke: Towards a Topical Thesaurus of Old English Emotional Expressions', *Amsterdamer Beiträge zur älteren Germanistik* 81.3-4 (2021), 442–56. doi: [10.1163/18756719-12340238](https://doi.org/10.1163/18756719-12340238).
- Learning with the Online Thesaurus of Old English (TOE)*, eds. C. Hough and C. Kay (Glasgow, 2017). <http://oldenglishteaching.arts.gla.ac.uk/>.
- Lye, E., *Dictionarium Saxonico et Gothico-Latinum* (London, 1772).
- Mapping English Metaphor through Time*, eds. W. Anderson, E. Bramwell, and C. Hough, (Oxford, 2016).
- Moriyama, H., 'Synonyms and Synonymous Expressions in the Old English Semantic Field "Hospitality, Harbours, and Entertaining"', *Waseda Global Forum* 7 (2010), 163-92.
- Oxford English Dictionary Online* (Oxford, 2000). <http://www.oed.com/>.
- Porck, T., 'Onomasiological Profiles of Old English Texts: Analysing the Vocabulary of *Beowulf*, *Andreas* and the *Old English Martyrology* through Linguistic Linked Data', *Amsterdamer Beiträge zur älteren Germanistik* 81.3-4 (2021), 359–83. doi: [10.1163/18756719-12340236](https://doi.org/10.1163/18756719-12340236).
- Roberts, J., 'The Old English Vocabulary of Nobility', in *Nobles and Nobility*, ed. A. J. Duggan (Cambridge, 2000), pp. 69-84.
- Roberts, J., 'A *Thesaurus of Old English*: The Pilot Study for the Glasgow *Historical Thesaurus*', *Amsterdamer Beiträge zur älteren Germanistik* 81.3-4 (2021), 297–317. doi: [10.1163/18756719-12340234](https://doi.org/10.1163/18756719-12340234).
- ScT* = *The Scots Thesaurus*, eds. I. Macleod et al. (Aberdeen, 1990).
- Selected Writings of Edward Sapir in Language, Culture and Personality*, ed. D. G. Mendelbaum (Berkeley, 1963).
- Sharifian, F., 'Cultural Linguistics and World Englishes', *World Englishes* 34.4 (2015), 515–32.
- Sommer, W., *Dictionarium Saxonico-Latino-Anglicum* (Oxford, 1659).
- Stellato, A., et al., 'VocBench 3: A Collaborative Semantic Web Editor for Ontologies, Thesauri and Lexicons', *Semantic Web* 11.5 (2020), 855–81.
- Stolk, S., 'Evoke: Exploring and Extending *A Thesaurus of Old English* using a Linked Data Approach', *Amsterdamer Beiträge zur älteren Germanistik* 81.3-4 (2021), 318–58. doi: [10.1163/18756719-12340235](https://doi.org/10.1163/18756719-12340235).
- TOE* = *A Thesaurus of Old English*; see *TOE1*, *TOE2*, *TOE3*, *TOE4*.
- TOE1* = *A Thesaurus of Old English*, eds. J. Roberts et al., 2 vols., King's College London Medieval Studies XI (London, 1995).
- TOE2* = *A Thesaurus of Old English*, eds. J. Roberts et al., 2 vols., Costerus New Series 131, 2nd edn (Amsterdam, 2000).

- TOE3* = *A Thesaurus of Old English*, eds. J. Roberts et al. (Glasgow, 2005).  
<http://libra.englant.arts.gla.ac.uk>. Accessed: 1 February 2013.
- TOE4* = *A Thesaurus of Old English*, eds. J. Roberts et al. (Glasgow, 2015).  
<http://oldenglishthesaurus.arts.gla.ac.uk>. Accessed: 15 March 2022.
- Van Baalen, A., ‘Identifying, Categorising and Exploring “Ælfrician” Vocabulary using the *Dictionary of Old English*, *A Thesaurus of Old English* and *Evoque*’, *Amsterdamer Beiträge zur älteren Germanistik* 81.3-4 (2021), 384–441. doi: [10.1163/18756719-12340237](https://doi.org/10.1163/18756719-12340237).
- Van de Poel, R. and S. Stolk, ‘A Case of Kinship: Onomasiological Explorations of KINSHIP in Old Frisian and Old English’, *Amsterdamer Beiträge zur älteren Germanistik* 81.3-4 (2021), 457–92. doi: [10.1163/18756719-12340239](https://doi.org/10.1163/18756719-12340239).
- Wierzbicka, A., *Understanding Cultures through Their Key Words: English, Russian, Polish, German, and Japanese*, Oxford Studies in Anthropological Linguistics 8 (New York, 1997).

## Appendix 8.A: EEMEE datasets publicly available

Many of the datasets created in the EEMEE research project have been made available publicly alongside the Linguistic Linked Data version of *A Thesaurus of Old English*. Figure 8.A.1 show the overview of content available in Evoke. Information on the new datasets, including their licenses and availability, is provided below.

The screenshot displays a grid of six dataset cards on the Evoke website. Each card features a circular icon, a title, author information, a brief description, and an 'Explore' button. The datasets are:

- A Thesaurus of Old English**: by Jane Roberts & Christian Kay with Lynne Grundy (version: 26 May 2017). This historical thesaurus captures the vocabulary of early medieval English, spoken between roughly 500 and 1100. Words that are considered similar in meaning are grouped together. Those that are special, because they are found only in poetry or glosses, are marked with a label.
- Andreas Thesaurus**: by Thijs Poek (version: 1.0, published Nov 2, 2021). This thesaurus contains words and phrases found in the Old English text *Andreas*. These words and phrases have been located in *A Thesaurus of Old English* and are tagged with the label 'andreas'.
- Ælfrician Vocabulary**: by Anos van Balen (version: 1.0, published Nov 11, 2021). This dataset contains words and phrases that are characteristic of the language use of Ælfric of Eynsham (c.955-1010), a prolific writer of Old English texts. These words and phrases have been located in *A Thesaurus of Old English* and are tagged with the label 'Ælfrician'.
- Beowulf Thesaurus**: by Thijs Poek (version: 1.0, published Nov 2, 2021). This thesaurus contains words and phrases found in the Old English text *Beowulf*. These words and phrases have been located in *A Thesaurus of Old English* and are tagged with the label 'beowulf'.
- Old English Martyrology Thesaurus**: by Thijs Poek (version: 1.0, published Nov 2, 2021). This thesaurus contains words and phrases found in the Old English text called the *Old English Martyrology*. These words and phrases have been located in *A Thesaurus of Old English* and are tagged with the label 'oemartyrology'.
- Old Frisian: Kinship**: by Rita van de Poel & Sander Stek (version: 1.0, published Nov 17, 2021). This dataset adds Old Frisian words and phrases pertaining to the semantic domain of Kinship to *A Thesaurus of Old English*, expanding the hierarchy of the thesaurus with specialized concepts specific to the medieval Frisian culture where necessary.

Figure 8.A.1.: Overview of content available in Evoke at <http://evoke.ullet.net/content/>.

### **Andreas Thesaurus**

*Creator:* Thijs Porck

*Version:* 1.0

*Released:* 2 November 2021

*License:* CC-BY-SA

*Availability:*

- in Evoke: <http://evoke.ullet.net/content/andreas/>
- in DataverseNL: <https://doi.org/10.34894/IHVH0Z>

### **Ælfrician Vocabulary**

*Creator:* Amos van Baalen

*Version:* 1.0

*Released:* 11 November 2021

*License:* CC-BY-SA

*Availability:*

- in Evoke: <http://evoke.ullet.net/content/aelfric/>
- in DataverseNL: <https://doi.org/10.34894/4BQ3ER>

### **Beowulf Thesaurus**

*Creator:* Thijs Porck

*Version:* 1.0

*Released:* 2 November 2021

*License:* CC-BY-SA

*Availability:*

- in Evoke: <http://evoke.ullet.net/content/beowulf/>
- in DataverseNL: <https://doi.org/10.34894/TOTFGZ>

### **Old English Martyrology Thesaurus**

*Creator:* Thijs Porck

*Version:* 1.0

*Released:* 2 November 2021

*License:* CC-BY-SA

*Availability:*

- in Evoke: <http://evoke.ullet.net/content/oemartyrology/>
- in DataverseNL: <https://doi.org/10.34894/QZCNW1>

### **Old Frisian: Kinship**

*Creators:* Rita van de Poel and Sander Stolk

*Version:* 1.0

*Released:* 17 November 2021

*License:* CC-BY-SA

*Availability:*

- in Evoke: <http://evoke.ullet.net/content/ofris-kinship/>
- in DataverseNL: <https://doi.org/10.34894/SOLVNU>

## Appendix 8.B: Exploring Early Medieval English Eloquence workshops

This appendix details the three workshops within the project ‘Exploring Early Medieval English Eloquence’ (EEMEE), initially titled ‘Exploring Anglo-Saxon Eloquence’ (EASE). The project centred around the use of the web application Evoke and *A Thesaurus of Old English* (*TOE*) for use in research and education. Further information on the workshops and other events surrounding EEMEE and Evoke can be found at <http://evoke.ullet.net/events>. EEMEE was supported by the LUCAS Extra Resources Open Call-II Grant 2020, awarded by the Leiden University Centre for the Arts in Society, and the LUCDH Small Grant 2018, awarded by the Leiden University Centre for Digital Humanities.

251

8

### 8.9.1. First workshop

The first workshop in the series was held at Leiden University on February 1st, 2019. The full programme is shown in Figure 8.B.2. The workshop drew attention to the web application Evoke and explored its value for various fields of research. Researchers convened to formulate potential lines of enquiry that utilise *TOE* in combination with this software. The aim of the first workshop was twofold: (1) to formulate powerful and evocative case studies in research using the thesaurus, and (2) to elicit feature requests to improve the Evoke platform for research purposes. Three talks at the start of the workshop introduced participants to the two resources. The full programmes of the workshops are included on the last three pages of the appendix.

After the introductory talks, the workshop continued with brainstorming on the use of the two resources in research. This activity was facilitated through the Change Pathway from the Europeana Impact Playbook.<sup>58</sup> This framework allowed participants to formulate lines of enquiry in terms of the intended outcomes, any resources required besides *TOE* and Evoke, and the activities involved. Figure 8.B.1 depicts the Change Pathway as filled out for Thijs Porck’s proposed research concerning onomasiological profiling of the Old English text *Beowulf*. During the workshop, groups of researchers reflected on each other’s ideas, which were sketched in this manner, and suggested improvements. Subsequently, feature requests for Evoke were elicited based on these ideas. On handouts, researchers could describe functionality desired of tooling to benefit research. The results have been incorporated into the development roadmap of Evoke and can be found in Appendix 2.A.

### 8.9.2. Second workshop

On October 17th, 2020, twelve researchers presented preliminary results of their research utilizing Evoke and *TOE* in the second EEMEE workshop. Its programme is shown in Figure 8.B.3. Old English language and culture, developments of metaphors, lexis used in specific texts or by specific authors,

<sup>58</sup>See ‘Europeana Impact Playbook’.

lexicographical practices, ways in which Old Germanic languages can be contrasted with each other, and the use of these digital resources in a classroom setting — the great variation of these talks demonstrated the value of historical language thesauri and their potential in academic research. Discussion of the approaches and results thus far supplied participants with useful notes for refining, and advancing, their work over the next few months.

### 8.9.3. Workshop at ICEHL-21

A full-day workshop on working *TOE* and Evoke was held at the 21st International Conference on English Historical Linguistics (ICEHL-21) on June 7th, 2021. The papers presented in this workshop demonstrated ways in which these two digital resources, at times complemented with additional data, can be used to explore exciting new aspects of Old English language and culture. Open to all ICEHL-21 attendants, the workshop discussed the results of the EEMEE case studies. The full programme is listed in Figure 8.B.4.

After the workshop, the majority of the case studies were developed into articles and submitted for inclusion in *Amsterdamer Beiträge zur älteren Germanistik* 81.3-4. This special issue of the journal, published in November 2021, is titled ‘Exploring Early Medieval English Eloquence: A Digital Humanities Approach with *A Thesaurus of Old English* and *Evoke*’ and was edited by Thijs Porck and Sander Stolk.

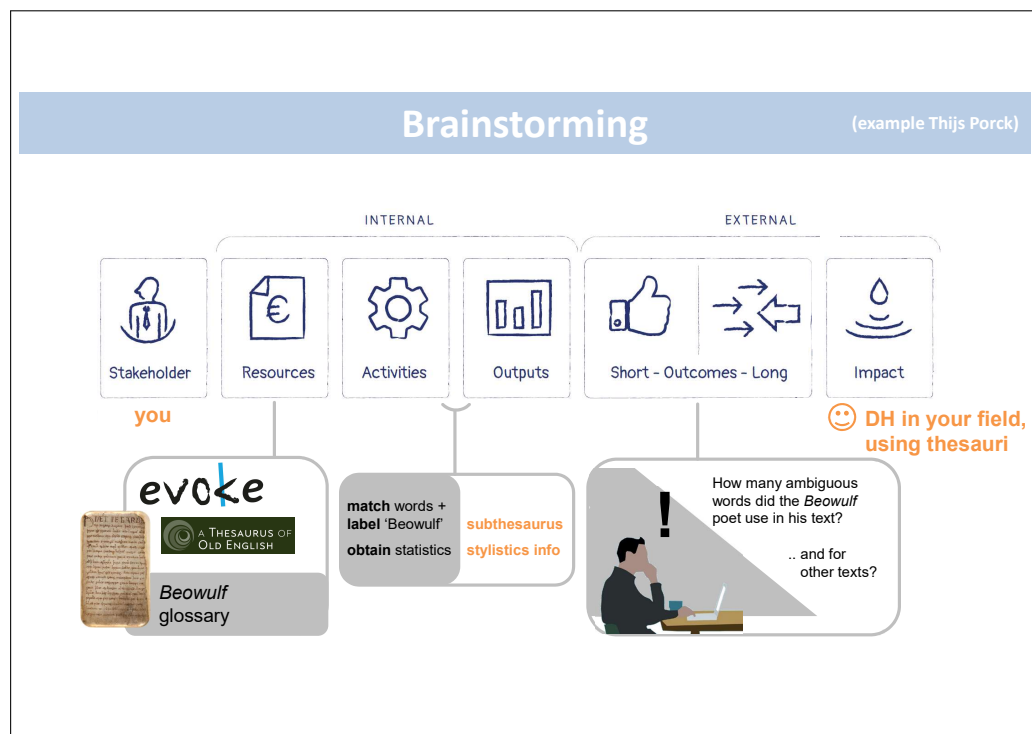


Figure 8.B.1.: Change Pathway towards onomasiological profiling of *Beowulf*.



## EASE: Exploring Anglo-Saxon Eloquence

First workshop: 1 February 2019, Leiden University.

The aim of this workshop series is to draw attention to the new digital platform Evoke, created at Leiden University, and to explore its value for various fields of research. This platform is meant to digitally open up historical thesauri for research. The first thesaurus that Evoke supports is *A Thesaurus of Old English*, which allows scholars to take a closer look at Anglo-Saxon language and culture. This workshop series aims to explore the kinds of inquiries that researchers would want to make using these resources and to ensure that such efforts can be accomplished with relative ease.

### First workshop

This first workshop invites researchers to become familiar with the online platform Evoke and *A Thesaurus of Old English* in order to explore the value of these resources for research in their fields. In discussing the use of the platform in combination with this thesaurus, as envisioned by those attending, the aim is twofold: (1) to brainstorm about powerful and evocative case studies in research using the thesaurus, and (2) to elicit feature requests to improve the Evoke platform for research purposes.

**When:** Friday 1<sup>st</sup> of February 2019 (13:00-16:00h)

**Where:** Leiden University, Lipsius building, room 2.17

#### Agenda:

[ 12:40 ] Reception with coffee and tea

[ 13:00 ] Welcome and introduction

[ 13:15 ] *The Thesaurus of Old English* (TOE) and the Evoke platform – from 3 perspectives

1. editor On the thesaurus TOE  
*presenter: J. Roberts – Prof. of English Language, editor of TOE*
2. developer On the software Evoke  
*presenter: S. Stolk – PhD student in DigHum, creator of Evoke*
3. researcher On possible research cases  
*presenter: M.H. Porck – Assistant Professor of Medieval English*

[ 14:15 ] *(Coffee break)*

[ 14:30 ] Group discussions on value for research

- Formulating case studies
- Formulating requirements on data and tooling

[ 16:00 ] Closing

[ 16:20 ] Informal drinks

Figure 8.B.2.: Programme of the first workshop.

## **ease** Virtual Workshop 17-10-2020

### **Working with *A Thesaurus of Old English* and the digital platform *Evoke***

Organisers: Thijs Porck (Leiden University) & Sander Stolk (Leiden University)

Virtual conf.-room: <https://smart.newrow.com/#/room/ldt-698> (open in Chrome, Edge or Firefox)

#### **11:00 CEST – Welcome by organisers**

#### **11:10 CEST - Session 1: *A Thesaurus of Old English*: Past, present and future**

11:10-11:30 – “The pilot study for the Glasgow Historical Thesaurus”

Jane Roberts (King’s College London)

11:30-11:50 – “Children and cousins: Projects arising from the Historical Thesaurus”

Marc Alexander + Fraser Dallachy (University of Glasgow)

11:50-12:15 – “*Evoke* and *A Thesaurus of Old English*: Exploring language and culture”

Sander Stolk (Leiden University)

12:15- 12:30 – *Short break*

#### **12:30 CEST - Session 2: Creating your own Old English thesaurus with *Evoke***

12:30-12:55 – “*By their words you shall know them*: A *Beowulf* thesaurus and onomasiological profiling of Old English texts with *Evoke*”

Thijs Porck (Leiden University)

12:55-13:15 – “Mapping conceptual variation through *A Thesaurus of Old English* and *Evoke*: Towards a topical thesaurus of Old English emotional expressions”

Anas F. Khan (Institute for Computational Linguistics “A. Zampolli”, CNR, Pisa) +

Javier E. Díaz-Vera + Francisco J. Minaya Gómez (University of Castilla-La Mancha) + Monica

Monachini (Institute for Computational Linguistics “A. Zampolli”, CNR, Pisa)

13:15-14:00 – *Lunch break (Bring your own lunch...)*

#### **14:00 CEST - Session 3: *Evoke* and Old English Studies**

14:00-14:25 – “*Evoke* and the history of Old English lexicography: Preliminary explorations”

Rachel Fletcher (University of Glasgow)

14:25-14:50 – “*Unhal, unmiht ond wanhal*: Exploring Old English semantic fields of disease and disability”

Berber Bossenbroek (Leiden University)

14:50-15:15 – “*Evoke* and *A Thesaurus of Old English* in the Old English classroom”

Kees Dekker (University of Groningen)

15:15-15:30 – *Short break*

#### **15:30 CEST – Session 4: Ælfrician English, Old Frisian and Old Dutch**

15:30-15:55 – “Using *A Thesaurus of Old English* and *Evoke* to research ‘Ælfrician’ vocabulary”

Amos van Baalen (Leiden University)

15:55-16:20 – “Contrasting Old Frisian with Old English: An exploration using *A Thesaurus of Old English*, the *Alfriesisches Handwörterbuch*, and the *Evoke* platform”

Rita van de Poel (Leiden University)

16:20-16:55 – “Linking vocabulary of the *Dictionary of Old Dutch* in DiaMaNT to *A Thesaurus of Old English*: An exploration of the possibilities and challenges”

Katrien Depuydt + Jesse De Does (Dutch Language Institute)

#### **16:55 CEST – Closing by organisers + possibility for final discussion**

Figure 8.B.3.: Programme of the second workshop.



## eASE Workshop @ ICEHL-21

Working with *A Thesaurus of Old English* and the digital platform Evoke  
organised by Thijs Porck + Sander Stolk (Leiden University)

**Welcome [10:00-10:15 CEST]**

**Session 1: A Thesaurus of Old English: Past, present and future [10:15-11:45]**

“The pilot study for the Glasgow Historical Thesaurus”  
Jane Roberts (King’s College London)

“Children and cousins: Projects arising from the Historical Thesaurus”  
Marc Alexander + Fraser Dallachy (University of Glasgow)

“Evoke and *A Thesaurus of Old English*: Exploring language and culture”  
Sander Stolk (Leiden University)

**Session 2: Creating your own Old English thesaurus with Evoke [12:00-13:00]**

“*By their words you shall know them*: A *Beowulf* thesaurus and onomasiological profiling of Old English texts with Evoke”  
Thijs Porck (Leiden University)

“Using *A Thesaurus of Old English* and Evoke to research ‘Ælfrician’ vocabulary in two twelfth-century texts”  
Amos van Baalen (Leiden University)

**- Lunch -**

**Session 3: Evoke and Old English Studies [14:30-15:30]**

“Mapping conceptual variation through *A Thesaurus of Old English* and Evoke: Towards a topical thesaurus of Old English emotional expressions”

Anas F. Khan (Institute for Computational Linguistics “A. Zampolli”, CNR, Pisa) +  
Javier E. Díaz-Vera + Francisco J. Minaya Gómez (University of Castilla-La Mancha) +  
Monica Monachini (Institute for Computational Linguistics “A. Zampolli”, CNR, Pisa)

“Evoke and *A Thesaurus of Old English* in the Old English classroom”  
Kees Dekker (University of Groningen)

**Session 4: Old English, Old Frisian, Old Dutch –  
Comparing Old Germanic lexis with Evoke [15:45-16:45]**

“Contrasting Old Frisian with Old English: An exploration using *A Thesaurus of Old English*, the *Alfriesisches Handwörterbuch*, and the Evoke platform”  
Rita van de Poel (Leiden University)

“Linking vocabulary of the *Dictionary of Old Dutch* in DiaMaNT to *A Thesaurus of Old English*: An exploration of the possibilities and challenges”  
Katrien Depuydt + Jesse De Does (Dutch Language Institute)

Reflection/Closing

Figure 8.B.4.: Programme of the third workshop.

## Appendix 8.C: Student-evaluation of Evoke (Nov. 2018)

On 13 November 2018, a small evaluation was carried out on the usefulness of Evoke in education. This evaluation took place in a 2-hour workshop, as part of a third-year Bachelor course on Old English at Leiden University, on digital resources for studying the language and culture. In the workshop, twenty-two students worked on assignments that revolved around the information available in *A Thesaurus of Old English* to study Old English language and culture. The evaluation contrasted the usefulness of Evoke with that of the existent *TOE* website, made available by the University of Glasgow (UoG).<sup>59</sup> The intention was to ensure that the typical classroom setting was disturbed as little as possible (and therefore to avoid the introduction of foreign elements, such as cameras and microphones) whilst still yielding valuable information on the user experience in education.

During the workshop, students were asked to form pairs; one student would work with Evoke, the other with UoG. In the workshop section devoted to *TOE*, each student first got acquainted with the website assigned to them over the time span of five minutes. They were asked to explore how these websites could be navigated and what the options were that they offered, informing each other of their findings afterwards. The subsequent thirty minutes were spent on worksheets with assignments on the *TOE* data. These were to be solved individually, using the assigned website, with the possibility to obtain help from the paired student when needed. After these assignments, both pieces of software were evaluated by the students.

In order to assess the usefulness of Evoke in a classroom setting, fundamental aspects were evaluated that, combined, establish a set of metrics on usefulness. Usefulness can be described as a combination of utility (i.e., the extent to which the application provides the necessary features to support users' needs) and usability (i.e., the ease with which the user interface can be used).<sup>60</sup> Usability, in turn, can be broken down further into various aspects to facilitate a more fine-grained evaluation. Table 8.C.1 lists five key aspects of usability. The classroom setting allowed evaluation of three of the five aspects on usability: learnability, efficiency, and satisfaction. Memorability and errors were not measured, since students would, during this workshop, be getting acquainted with the user interface of Evoke. The three evaluated aspects of usability were included alongside utility and an overall impression on usefulness.

The evaluation of Evoke and UoG, across various aspects of usefulness, was performed through a short poll on each aspect. In order to remove bias from these polls, the evaluation drew on the Microsoft Desirability Toolkit,<sup>61</sup> which consists of a list of 118 words or phrases for possible reactions (e.g., "convenient", "difficult", "boring"). The students were asked to select the words

---

<sup>59</sup>The evaluation was performed on Evoke v1.2.0 and UoG's *TOE4*.

<sup>60</sup>See, for instance, Nielsen Norman Group's methodologies for evaluating user experience. <https://www.nngroup.com/articles/usability-101-introduction-to-usability/>

<sup>61</sup>Benedek and Miner, 'Measuring Desirability'.

Aspect	Description
learnability	the ease with which first-time users can accomplish basic tasks
efficiency	the speed with which users can perform tasks
memorability	the extent to which users remember how to work with the interface after not having used it for an extended period
errors	the number and severity of mistakes users make
satisfaction	the pleasantness of the design and its use

Table 8.C.1.: Key aspects of usability

that best described their stance towards the aspect under consideration, which was introduced by a short phrase (e.g., “The look/visuals of that website” for the aspect of satisfaction). Their possible answers, in the form of these reaction words, were narrowed down to a maximum of ten that were suitable for the aspect in question (such as “fast” and “slow” for efficiency).

The results of the evaluation are shown in the tables below for both users of Evoke (ten out the twenty-two students) and UoG (twelve students).<sup>62</sup> A wordcloud next to each table visualizes the results for Evoke, specifically, with scale and darkness of a word or phrase representing the relative number of users that selected it. These results show that both Evoke and UoG were received positively by the students, on both matters of utility and usability. When contrasting the two websites in the results, the most striking differences include that Evoke was more often considered to offer “desirable” functionality, which students later indicated was mostly owing to the statistics generated by the application, and “fun” to use in the assignments.

Word	Value	Evoke users	UoG users
Desirable	+	50 %	17 %
Ineffective	-	10 %	17 %
Powerful	+	20 %	8 %
Helpful	+	100 %	100 %
Dated	-	0 %	0 %
Cutting edge	+	30 %	8 %
Irrelevant	-	0 %	0 %
Not valuable	-	0 %	0 %
Poor quality	-	0 %	0 %
Useful	+	100 %	92 %

Table 8.C.2.: Results for Evoke on utility  
 (“The functionality offered by that website”)

<sup>62</sup>The discrepancy with the number of students working with Evoke and with UoG is the result of one pair of students, out of the eleven pairs, having misunderstood their distinctive roles and worked with UoG both.

Word	Value	Evoke users	UoG users
Effortless	+	40 %	17 %
Annoying	-	0 %	25 %
Fast	+	70 %	42 %
Slow	-	10 %	25 %
Disruptive	-	10 %	25 %
Efficient	+	70 %	50 %

Effortless  
**Efficient**  
 Fast

Table 8.C.3.: Results for Evoke on efficiency  
 (“The efficiency with which it allowed me to perform the tasks”)

Word	Value	Evoke users	UoG users
Difficult	-	0 %	0 %
Straightforward	+	60 %	50 %
Confusing	-	20 %	33 %
Too technical	-	0 %	0 %
Clear	+	80 %	42 %
Incomprehensible	-	0 %	0 %
Accessible	+	60 %	58 %
Understandable	+	50 %	75 %
Easy	+	70 %	33 %
Stressful	-	0 %	8 %

Easy  
 Straightforward  
**Clear**  
 Understandable  
 Accessible

Table 8.C.4.: Results for Evoke on learnability  
 (“The process of learning to use that site”)

Word	Value	Evoke users	UoG users
Attractive	+	50 %	25 %
Boring	-	10 %	0 %
Clean	+	90 %	58 %
Overwhelming	-	0 %	0 %
Calm	+	60 %	58 %
New	+	20 %	0 %
Cutting edge	+	0 %	0 %
Unattractive	-	0 %	0 %
Patronizing	-	20 %	0 %
Old	-	0 %	8 %
Organized	+	80 %	92 %
Satisfying	+	50 %	25 %

Satisfying  
Calm  
Patronizing  
Clean  
Attractive  
Organized

Table 8.C.5.: Results for Evoke on satisfaction  
("The look/visuals of that website")

Word	Value	Evoke users	UoG users
Convenient	+	70 %	67 %
Frustrating	-	10 %	33 %
Valuable	+	20 %	25 %
Useful	+	80 %	100 %
Poor quality	-	0 %	0 %
Essential	+	20 %	8 %
High quality	+	30 %	8 %
Dated	-	0 %	0 %
Fun	+	90 %	25 %
Professional	+	20 %	50 %

Useful  
Essential  
Fun  
Professional Valuable  
Convenient  
High quality

Table 8.C.6.: Results for Evoke on overall perception  
("My feeling of that website overall...")