

Evoked: Platform for Historical Thesauri on the Semantic Web ¹

Sander Stolk (Leiden University) ²

Thesauri are lexicographic resources that organize words and word senses.² One of their main uses is to look up available alternative phrasings,³ but thesauri offer a number of uses beyond that: they are veritable treasure troves for cultural, linguistic, anthropological, and literary-critical research. This is especially the case when such thesauri are arranged in a topical fashion, an ordering of all its groups of synonymous words according to their meaning.^{4,5,6} Some historical thesauri – thesauri that show the history or development of a language – already exist but their information is currently restricted to a paper form (e.g., *The Scots Thesaurus*) or to webpages that are rather limited (e.g., *Thesaurus of Old English*, *Historical Thesaurus of English*).^{7,8,9} These forms make it difficult for scholars to use them to the fullest. In making existing historical thesauri available in a more suitable form for use and reuse, bringing them to the Semantic Web as Linked Data, I aim to facilitate a wide variety of research that focusses on the vocabulary and culture of current and past times.

Which lexical items were available to the Old English authors Cynewulf and Ælfric? What concepts and notions do these words evoke? Which words did Cynewulf prefer over those that brought the same notions to mind? How many nuances did the English language possess for the notions of peace and war, and did this change throughout history? It is just such questions that historical thesauri should be able to answer in an effective manner.

In this demonstration, I will present (1) the proposed digital form for historical thesauri on the Semantic Web and (2) **Evoked**, the new web-based platform for thesauri expressed in such a form.³ The selected digital form is based on a number of best practices for data, employing standardized and interoperable data vocabularies.^{10,11} As a result, it will enable scholars to access, query, filter, and expand thesaurus content. The second part of this demonstration is a glimpse at an early version of **Evoked**. Using this novel platform, I will show how valuable thesaurus content such as that of the *Thesaurus of Old English* can be explored by users in a simple but effective manner. The *Thesaurus of Old English* is a historical thesaurus that captures the lexis of the early medieval variant of English, spoken between roughly 500 and 1100 by the Anglo-Saxons, and has been the first resource to be reshaped into the proposed Semantic Web form. The added value of this digital form becomes easier to grasp when realising that, with **Evoked**, one need not be a computer scientist in order to fully utilize such content. This platform, then, aims to unlock the potential of these semantic treasure troves for scholars – firmly integrating the digital in Humanities.

¹ *Evoked* (Leiden, 2018), <http://evoked.ullet.net> .

² s.s.stolk@hum.leidenuniv.nl

² C. Kay and J. Roberts, *The Encyclopedia of Language and Linguistics*, ed. R. E. Asher, 10 vols. (Oxford, 1994), s.v. ‘Thesaurus’.

³ See, for instance, the foreword to the first edition of Peter Mark Roget’s thesaurus: P. M. Roget, *Thesaurus of English Words and Phrases, Classified and Arranged so as to Facilitate the Expression of Ideas and Assist in Literary Composition* (London, 1852).

⁴ C. Brewer, Review of *Historical Thesaurus of the Oxford English Dictionary*, *The Review of English Studies* 61:252 (2010), 801–5 (p. 802).

⁵ A. Adamska-Sałaciak, Review of *Historical Thesaurus of the Oxford English Dictionary*, *International Journal of Lexicography* 23:2 (2010), 227–33 (p. 232).

⁶ B. Busse, ‘A Celebration of Words and Ideas: The Stylistic Potential of the *Historical Thesaurus of the Oxford English Dictionary*’, *Language and Literature* 21:1 (2012), 84–92 (p. 88).

⁷ *The Scots Thesaurus*, eds. I. Macleod et al. (Aberdeen, 1990).

⁸ *A Thesaurus of Old English*, eds. J. Roberts et al. (Glasgow, 2015), <http://oldenglishthesaurus.arts.gla.ac.uk> .

⁹ *The Historical Thesaurus of English*, eds. C. Kay et al. (Glasgow, 2014), <http://historicalthesaurus.arts.gla.ac.uk> .

¹⁰ M. D. Wilkinson et al., ‘The FAIR Guiding Principles for Scientific Data Management and Stewardship’, *Scientific Data* 3:160018 (2016).

¹¹ ‘Data on the Web Best Practices’, eds. B. F. Lóscio et al., W3C (2017), <http://www.w3.org/TR/dwbp/> , W3C Recommendation.