

Old English was a language used by the inhabitants of the British Isles occupying the area of present-day England, in the period between the Anglo-Saxon invasion of the 5th century and the Norman Conquest of the 11th century. The language was characterised by a relatively free word order and its structure resembled that of many modern Germanic languages, such as German or Icelandic. Modern English is considerably different from other Germanic languages, but in the Anglo-Saxon period, the language of the inhabitants of Britain was free from French influence and its character was decidedly Germanic. Researching Old English is possible thanks to an immense richness of Anglo-Saxon literature, which has survived to date. Apart from the most celebrated specimens of Anglo-Saxon poetry, such as Beowulf, there exists a number of prose texts, the best part of which are more or less free translations from Latin (e.g. fragments of the Bible, lives of saints and numerous homilies).

Studies in syntax, i.e. the structure of clauses, are typically conducted on prose texts because poetic rhymes and rhythm often force the use of unusual structures. Therefore, the most popular tool for researching Old English syntax is the YCOE corpus (The York-Toronto-Helsinki Parsed Corpus of Old English Prose), i.e. a digital collection of prose texts, where every word is assigned to a grammatical class and words are joined in phrases and clauses. This tool allows for an automatic extraction of constructions and clause types which are of interest to the researcher and saves time, as well as facilitates a comprehensive study of Old English syntax with the use of a relatively large database (the corpus numbers c. 1.5 million words).

Thanks to the YCOE corpus, many compelling studies in Old English syntax have appeared, but so far, nobody has analysed the differences between prose texts themselves. It is a long-recognised fact that the respective syntaxes of poetry and prose are different from each other, but most experts assume that there is no significant syntactic variation between individual prose texts. However, there is a number of reasons to question this assumption. Firstly, the oldest Old English texts date back to the 8th century, while the most recent were written in the 11th century, and it is not inconceivable that the language had undergone some major changes over 300 years. Secondly, the YCOE corpus includes texts from different dialectal areas so regional differences cannot be ruled out. Thirdly, a large portion of texts are translations from Latin so certain constructions may be linguistic calques. Finally, the existence of stylistic differences should be allowed for, and disregarding them in general descriptions of Old English may distort the perception of what the language looked like.

The aim of the project is to conduct a detailed analysis of Old English prose as regards syntactic and phraseological differences between texts. There are numerous studies in Old English syntax, but our project will approach syntax from a different angle, i.e. that of an individual text. In the course of the project, we intend to identify those syntactic constructions which so far have been considered representative of Old English, but which, in reality, are concentrated in one or few texts. This will allow us to verify our knowledge of Old English syntax. Then, the analysis of phraseology, i.e. recurrent word combinations, is a clear novelty in studies of Old English. The YCOE corpus does not allow for studying phraseology because the words included are not lemmatised (i.e. simplified to their base forms), which proves a major obstacle for analysing phraseology due to the complexity of Old English inflection (nouns, adjectives and pronouns are declined for 4 cases, verbs take personal endings etc.) and spelling inconsistencies. Hence, one of the main tasks will be to manually lemmatise a portion of the corpus, which will help meet the goals set for the project.