

Staroangielski to język używany przez mieszkańców Wysp Brytyjskich z terenów współczesnej Anglii między inwazją anglosaską w V w. n.e. a inwazją normańską w XI w. n.e., charakteryzujący się dość swobodnym szykiem zdania i strukturą przypominającą wiele współczesnych języków germańskich, takich jak niemiecki czy islandzki. Współczesna angielszczyzna różni się zasadniczo od innych języków z rodziny germańskiej, ale w okresie anglosaskim język mieszkańców Brytanii był wolny od wpływów francuskich i miał czysto germański charakter. Badania nad językiem staroangielskim są możliwe dzięki niezwykłemu bogactwu literatury anglosaskiej, która zachowała się do naszych czasów. Poza najśłynniejszymi przykładami poezji staroangielskiej takimi jak Beowulf istnieje też wiele tekstów zapisanych prozą, z których duża część to mniej lub bardziej swobodne tłumaczenia z łaciny (m.in. fragmenty Biblii, żywoty świętych i liczne homilie).

Badania składni, czyli struktury zdań, zazwyczaj opierają się na tekstach pisanych prozą, ponieważ wykorzystywany w poezji rytm i rym często powoduje użycie nietypowych struktur. Z tego względu najpopularniejszym obecnie narzędziem do badania składni staroangielskiej jest korpus YCOE (The York-Toronto-Helsinki Parsed Corpus of Old English Prose), czyli elektroniczny zbiór tekstów pisanych prozą, w których każde słowo ma przypisaną kategorię gramatyczną, a słowa są powiązane ze sobą we frazy i zdania. To narzędzie umożliwia automatyczne wyodrębnianie interesujących badacza konstrukcji czy rodzajów zdań, co pozwala zaoszczędzić czas i prowadzić przekrojowe badania nad składnią staroangielską przy użyciu stosunkowo dużej bazy danych (korpus zawiera łącznie ok. 1,5 mln słów).

Dzięki korpusowi YCOE udało się przeprowadzić wiele interesujących badań nad językiem staroangielskim, jednak jak dotąd nikt nie przeanalizował różnic pomiędzy tekstami pisanymi prozą. Od dawna wiadomo, że składnia staroangielskiej poezji i prozy różnią się od siebie, ale większość specjalistów z góry zakłada, że w obrębie prozy nie ma istotnych różnic składniowych. Istnieje jednak wiele powodów, bo podać to założenie w wątpliwość. Po pierwsze najstarsze teksty staroangielskie pochodzą z VIII w., a najmlodsze z XI, a przez 300 lat język mógł ulec zasadniczym zmianom. Po drugie korpus YCOE obejmuje teksty z różnych obszarów dialektalnych, więc nie można też wykluczyć istnienia różnic regionalnych. Po trzecie znaczna część tekstów to tłumaczenia z łaciny, więc niektóre konstrukcje mogą stanowić przykłady kalek językowych. Po czwarte nie można też wykluczyć różnic stylistycznych, których nieuwzględnienie w opisie języka staroangielskiego może zaburzać naszą wiedzę na jego temat.

Celem proponowanego projektu jest szczegółowa analiza prozy staroangielskiej w zakresie różnic składniowych i frazeologicznych między tekstami. Istnieje wiele badań na temat składni staroangielskiej, ale nasz projekt będzie nowym spojrzeniem na składnię z punktu widzenia pojedynczego tekstu. W ramach projektu zamierzamy zidentyfikować jak najwięcej konstrukcji składniowych uznawanych dotąd za właściwe dla języka staroangielskiego, a w rzeczywistości ograniczonych wyłącznie do jednego czy kilku tekstów. To pozwoli na zweryfikowanie naszej wiedzy dotyczącej składni tego języka. Z kolei analiza frazeologii, czyli powtarzalnych połączeń międzywyrazowych, jest całkowitą nowością w badaniach nad językiem staroangielskim. Korpus YCOE nie umożliwia prowadzenia badań frazeologicznych, ponieważ słowa w korpusie nie są zlematyzowane (tj. sprowadzone do swojej formy podstawowej), co stanowi zasadniczą przeszkodę w badaniach frazeologii ze względu na złożoność fleksji staroangielskiej (rzeczowniki, przymiotniki i zaimki odmieniane przez 4 przypadki, końcówki osobowe czasowników itd.) i różnice ortograficzne. Z tego powodu jednym z głównych zadań projektu będzie ręczna lematyzacja części korpusu YCOE, co umożliwi realizację zakładanego celu.