

Celem projektu jest przeanalizowanie zmian w angielskiej frazeologii od najwcześniejszego okresu istnienia tego języka, tzw. staroangielszczyzny (V-XI w.), aż do początków XX wieku. Badania frazeologiczne stają się obecnie coraz popularniejsze dzięki dostępności dużych elektronicznych zbiorów językowych (tzw. korpusów), których analiza pozwala stwierdzić, że rodzimi użytkownicy wszystkich języków regularnie korzystają z różnego rodzaju jednostek frazeologicznych, a skala tego zjawiska jest nadszpiewanie duża. Okazuje się, że bardzo duża część tego, co mówimy i piszemy, jest formułkowa, czyli nie jest tworzona za każdym razem od zera poprzez łączenie ze sobą pojedynczych słów, lecz bazuje na tzw. prefabrykacjach językowych, czyli różnego rodzaju ciągach słów o różnej długości jak *no i wiesz, nie ma mowy, to jest moje ostatnie słowo*, które nasz mózg przechowuje w całości, dzięki czemu są od razu gotowe do użycia. Tego rodzaju jednostki frazeologiczne są bardzo zróżnicowane i obejmują m.in. kolokacje swobodne (całkowicie przejrzyste znaczeniowo kombinacje słów, np. *grać na pianinie czy iść do domu*), kolokacje ograniczone (dość przejrzyste znaczeniowo, ale z mniejszą swobodą wyboru jednego z elementów, np. *strzelić focha*), idiomy (*ręce opadają*), formuły o ściśle określonej funkcji (*pozdrawiam serdecznie*), ale również tzw. zbitki leksykalne (ang. *lexical bundles*), które nie stanowią kompletnych jednostek składniowych, np. *nie wiem czy, tak że są, i wtedy on*. Nasza wiedza na temat skali użycia różnych jednostek frazeologicznych i ich funkcji we współczesnych językach, zwłaszcza angielskim, jest dość duża, ale badania frazeologiczne języków historycznych należą wciąż do rzadkości. Nasz projekt stawia sobie za cel kompleksową analizę zmian frazeologicznych w angielszczyźnie, a główne pytania badawcze projektu to:

1. jaka jest rola prefabrykacji językowej w języku staroangielskim, średnioangielskim oraz wczesnym i późnym nowoangielskim? jakie są najistotniejsze różnice w ilości i rodzaju jednostek frazeologicznych pomiędzy tymi okresami?
2. które rodzaje jednostek frazeologicznych wykazują największą dynamikę zmian i jakie są przyczyny tak dużej zmienności?
3. które jednostki frazeologiczne były obecne w angielszczyźnie od jej najwcześniejszego okresu? jakie są ich funkcje i jakie czynniki spowodowały ich utrzymanie?
4. które gatunki prozy (np. teksty medyczne, homilie, traktaty naukowe, kroniki) w największym stopniu korzystają z prefabrykatów językowych i czy ta tendencja utrzymuje się w czasie?
5. które ze zmian frazeologicznych mogły być motywowane czynnikami pozajęzykowymi (zmiany społeczne i kulturowe, wydarzenia historyczne, nowe konwencje pisarskie) i co może nam to powiedzieć o mechanizmach zmiany frazeologicznej?

Aby zrealizować cele projektu, konieczne jest zastosowanie nowoczesnych narzędzi korpusowych, w tym automatycznych lematyzatorów bazujących na metodach NLP (ang. *natural language processing*), pozwalających na powiązanie każdej słowoformy z formą bazową (np. mianownika liczby pojedynczej dla rzeczowników czy bezokolicznika dla czasowników, w standardowej formie ortograficznej). Algorytmy NLP, stworzone z myślą o językach współczesnych, po odpowiednim treningu mogą być stosowane w przetwarzaniu danych historycznych, a nasz zespół zamierza w pełni wykorzystać potencjał tej metodologii. Dzięki tym działaniom i po zlematyzowaniu korpusów reprezentujących wszystkie cztery okresy rozwoju angielszczyzny (między 1 a 2 miliony słów na każdy okres) będziemy w stanie stworzyć internetowe słowniki kolokacji w formacie opracowanego przez nas zespół słownika VARIOE (<http://varioe.pelcra.pl/collocations>). Oznacza to, że nasz projekt w nowatorski sposób łączy najnowocześniejsze technologie językowe z celami naukowymi językoznawstwa historycznego, gdzie badania prowadzone są zazwyczaj w bardziej tradycyjny sposób. Wykryte przez nas jednostki frazeologiczne zostaną poddane wielowymiarowej analizie uwzględniającej aspekty takie jak stabilność w czasie, funkcja, powtarzalność użycia w kontekście i specjalizacja gatunkowa. Naszym celem jest wychwycenie głównych różnic jakościowych i ilościowych między poszczególnymi okresami rozwoju angielszczyzny i interpretacja tych zmian w szerokim kontekście zmian społeczno-kulturowych w krajach anglojęzycznych. Efektem projektu będzie przekrojowa i głęboka analiza dynamiki zmian w obszarze frazeologii angielskiej, pozwalająca frazeologii historycznej na zyskanie odpowiedniego statusu w ramach badań nad historią języka angielskiego.