

Sebastian Jan Dunat

ACADEMIC LANGUAGE FUNCTIONS IN CONFERENCE ABSTRACTS CORPUS ANALYSIS OF ACADEMIC LANGUAGE FUNCTIONS

STRESZCZENIE

3.05.2025r.

Słowa kluczowe: język akademicki, funkcje języka akademickiego, abstrakty, analiza korpusowa, analiza korpusowa abstraktów, badania korpusowe, badania korpusowe oparte na płci autorów

Istnieją różne sposoby przedstawiania, definiowania i badania struktury i funkcji retorycznych abstraktów, co ilustrują przytoczone w tej rozprawie prace naukowe innych badaczy. Pomimo tego, że poprzednie badania odnosiły się do wielu celów badań nad abstraktami, nie tylko w zakresie językoznawstwa korpusowego, żadne z nich nie koncentrowało się na funkcjach języka akademickiego w abstraktach. Zwłaszcza z wykorzystaniem metod lingwistyki komputerowej do celów badawczych. Niniejsza rozprawa została rozplanowana następująco: materiał z przygotowanego przez badacza korpusu został oceniony i opisany z pomocą metod lingwistyki komputerowej, w celu porównania słownictwa akademickiego używanego w abstraktach z konferencji naukowych, przez obie płcie w wybranych do badania dziedzinach naukowych. Podstawą do określenia najczęściej używanego słownictwa była metodologia Averil Coxhead (2000). Celem analizy było wykrycie ewentualnych interdyscyplinarnych różnic w częstości występowania funkcji języka akademickiego. Jako wiodący parametr badania przyjęto płć. Przedstawione dane korpusowe, które posłużyły do analizy leksykalnej, zostały wpierw podzielone na części mowy, a następnie na trzy kategorie: Słownictwo w procesie badawczym, Słownictwo w procesie analizy oraz Słownictwo w procesie oceny. Kategorie te zostały wybrane na podstawie metodologii przyjętej przez Anne V. Martin (1976). Dane do korpusu zostały wybrane wyłącznie z pięciu ksiąg abstraktów, z pięciu konferencji, z których każda odbywała się na innym kontynencie. Dane wybierane były ręcznie przez badacza, aby wyrównać liczbę przykładów, zgodnie z parametrem płci.

Dane korpusu zostały podzielone na trzy dziedziny nauki humanistyczne, nauki biologiczne i zdrowotne oraz nauki wykorzystujące technologie informatyczne. Korpus użyty w niniejszej rozprawie został zaprojektowany w celu określenia ilościowych różnic w wykorzystaniu funkcji języka akademickiego w różnych dziedzinach pisanego języka akademickiego. Dane korpusu zawierają następujące elementy: Źródło; Numer referencyjny i numer strony; Autora, Płć, Typ mówcy, Instytucję i Pochodzenie mówcy; Dyscyplinę, Tytuł,

Abstrakt; Abstrakt oznaczony częściami mowy; Ilość rzeczowników, czasowników, przymiotników i przysłówków użytych w tekstach korpusu. Zakłada się, że w grupach badawczych, oraz między nimi, wystąpią różnice w realizacji funkcji językowych i użytku słownictwa akademickiego.

Praca podzielona jest na pięć rozdziałów. Pierwszy rozdział przedstawia podstawy teoretyczne dotyczące: komunikacji w dyskursie akademickim, gatunku abstraktu, elementów lingwistyki korpusowej i komputerowej, aspektu płci w badaniach językoznawczych, elementów oceny danych w badaniach statystycznych oraz słownictwa i funkcji języka akademickiego; informacje zawarte w pięciu podrozdziałach pierwszego rozdziału tworzą podstawę najważniejszych pojęć użytych w tej rozprawie. Drugi rozdział zawiera opis praktycznych podejść do analizy abstraktów, nie tylko w badaniach korpusowych. Siedem punktów przedstawia streszczenia prezentowanych artykułów z czasopism, które koncentrują się na perspektywach międzyjęzykowych, dyscyplinarnych, międzykulturowych i diachronicznych, a także na strukturze retorycznej i elementach studiów korpusowych. W rozdziale trzecim przedstawiono metody badań oraz przybliżono programy komputerowe użyte do analizy, tj. programy wykorzystywane w rozprawie (R, AntConc, ProtAnt, TagAnt). Centralna część pracy, rozdział czwarty, prezentuje dane korpusowe, przebadane pod kątem płci i pochodzenia mówców, aby lepiej zobrazować różnice i podobieństwa występujące w badanym materiale. Ilościowa analiza komputerowa była podstawową metodą badania wybranych próbek zebranego materiału. Interpretacje przedstawionych tabel, zawierające dane odnośnie użycia określonej leksyki oraz ilościowego użycia określonych funkcji języka akademickiego przez kobiety i mężczyzn, są weryfikowane za pomocą testu chi-kwadrat Pearsona przeprowadzonego w programie R i oprogramowaniu Ant. Ostatnia część pracy dyplomowej zawiera wnioski odpowiadające interpretacji danych przedstawionych w poszczególnych częściach rozdziału czwartego oraz propozycje dalszych badań w tych zakresach. Poszczególne sekcje tego rozdziału przywołują podobieństwa i różnice danych badawczych, dotyczących płci i pochodzenia autorów (native/non-native). Wyniki badań omówiono w odniesieniu do wybranych źródeł literatury. Wnioski poparte są testami chi-kwadrat Pearsona (w stosownych przypadkach), wyodrębnionymi z oprogramowania R, przedstawiającymi liczbowe różnice między określonymi danymi badanego materiału.

Przeprowadzone testy chi-kwadrat potwierdziły statystyczne kilka zależności. Po pierwsze, użytek funkcji F7, związanej z ewaluacją, jest zróżnicowany pomiędzy badanymi dziedzinami i wpływa na niego płeć ($p=0,02$). Po drugie, istotną wartość testu chi-kwadrat ($p=0,01$) wykazał użytek funkcji F9, odnoszącej się do organizowania kolejności

przedstawianych informacji, w odniesieniu do ogólnej liczby rodzimych i nierodzimych pisarzy stosujących tą funkcję w abstraktach w badanych dziedzinach. Po trzecie, funkcja F1, związana z procesami badawczymi, analizowana w odniesieniu do połączonych parametrów dziedziny akademickiej i pochodzenia pisarzy dała wartość p równą 0,009. Następnie, funkcje F3.1, F3.2 oraz F3.3, związane z prezentowaniem metodologii, badane w odniesieniu do połączonych parametrów płci oraz dziedziny akademickiej, wykazały wartości testu na poziomach 0,008, 0,03 oraz 0,004. Kolejno, funkcja F4, związana z przedstawianiem wyników badań, jest używana częściej przez pisarzy rodzimych w dziedzinie Biologii i Nauk o zdrowiu oraz przez nierodzimych w pozostałych badanych dziedzinach ($p=0,03$). Ostatecznie, funkcja F5, odnosząca się do przedstawiania wniosków, uzyskała w badaniu chi-kwadrat wartość 0,002, co przekłada się na częstszy statystyczny użytek owej funkcji przez nierodzimych pisarzy w dziedzinach Statystyki komputerowej oraz Humanistyki w badanym materiale.

Sebastian Jurek