

Wrocław/Opole, 20.11.2023 r.

Prof. dr hab. Łukasz Grabowski

Instytut Językoznawstwa

Uniwersytet Opolski

Pl. Kopernika 11

45-040 Opole

Recenzja rozprawy doktorskiej mgr Pauliny Halczak pt. *Collocations in legal language in online monolingual English learners' dictionaries* (tytuł w j. polskim: *Kolokacje w języku prawa w jednojęzycznych słownikach online dla uczących się języka angielskiego*) (ss. 139) przygotowanej na Uniwersytecie im. Adama Mickiewicza w Poznaniu pod kierunkiem prof. dra hab. Roberta Lwa oraz prof. dra Thomasa Herbsta

Niniejsza recenzja została przygotowana na podstawie Uchwały nr 2/RND/SNJL/2023/2024 Rady Naukowej Dyscyplin Językoznawstwa i Literaturoznawstwa Uniwersytetu im. Adama Mickiewicza w Poznaniu z dnia 19.10.2023 r.

Uwagi wstępne

Przedstawiona do recenzji rozprawa doktorska z obszaru leksykografii stanowi ważny wkład w dotychczasowe badania nad używalnością elektronicznych słowników pedagogicznych języka angielskiego w zakresie wyodrębniania zeń – dla potrzeb tłumaczenia z języka polskiego na angielski – frazeologii specjalistycznej (kolokacji typowych dla języka prawnego i prawniczego). Badanie wykonano na próbie 90 studentów (82 osoby z Poznania i 8 osób z Łodzi) kierunku filologia angielska. W dysertacji przedstawiono także cenną propozycję spójnego modelu metodologicznego do przeprowadzania takich i podobnych badań.

Struktura dysertacji, cele badania i wyniki

Recenzowana rozprawa doktorska liczy 139 stron; składa się ze wstępu (poprzedzonego wykazami tabel i wykresów), syntetycznego rozdziału teoretycznego, rozdziału metodologicznego oraz rozdziału empirycznego, prezentującego wyniki badań. Całość

wieńczą konkluzje, bibliografia (16 stron) oraz obszerne załączniki opisujące materiał badawczy.

Celem studium przedstawionego w dysertacji była odpowiedź na trzy pytania badawcze. Pierwsze z nich dotyczyło określenia sposobów prezentacji kolokacji specjalistycznych w mikrostrukturze przebadanych elektronicznych jednojęzycznych słowników pedagogicznych języka angielskiego (grupa słowników znanych jako tzw. Big Five). Jako próbkę wybrano 16 kolokacji (połączenia rzeczownika z przydawką oraz rzeczownika z czasownikiem), gdzie ośrodek kolokacji wykazywał polisemię, tj. posiadał przynajmniej trzy sensy w artykule hasłowym. Skupiono się na takich elementach artykułów hasłowych, jak prymarne przykłady użycia, dodatkowe przykłady użycia oraz zbiory kolokacji. Drugie pytanie badawcze dotyczyło określenia statystycznie istotnych zmiennych niezależnych spośród zaproponowanych w badaniu i mających wpływ na poprawność tłumaczenia kolokacji (tj. na zmienną zależną/nominalną o wartości *Correct*). W tym celu przeprowadzono regresję logistyczną z modelowaniem efektów mieszanych (*mixed effects logistic regression models*). Ostatnie, trzecie pytanie dotyczyło określenia zmiennych niezależnych mających wpływ na szybkość (zmienna ilorazowa *Time*) wykonania zadania tłumaczeniowego, do czego wykorzystano regresję liniową.

Ogólnie rzecz ujmując, Autorka przeprowadziła swoje badanie w sposób metodologicznie poprawny, a także udzieliła przekonujących odpowiedzi na wszystkie trzy pytania badawcze. Po pierwsze wskazano najbardziej typowe sposoby egzemplifikacji kolokacji w prymarnych przykładach użycia, przykładach dodatkowych oraz zbiorach kolokacji w przeanalizowanych słownikach, zwracając uwagę na niektóre nieścisłości typograficzne. Po drugie wykazano, że dostęp do informacji o kolokacyjności (przez kliknięcie bądź bez kliknięcia w link), struktura składniowa kolokacji (tzw. koligacja) oraz interakcja między obiema zmiennymi mają wpływ na poprawność wyboru, z zastrzeżeniem, że dostęp bez kliknięcia w link zwiększa znacząco poprawność wyboru. Po trzecie wykazano, że opcja dostępu do informacji o kolokacji wpływa w sposób statystycznie istotny na czas wykonania zadania, a mówiąc ściślej, że dostęp przez kliknięcie w link znacząco skraca czas wykonania zadań, których łącznie było 16

Oryginalność recenzowanej dysertacji doktorskiej

W niniejszej recenzji chciałbym podkreślić kilka aspektów ocenianej rozprawy, które nadają jej twórczy i oryginalny charakter:

- Dysertacja wypełnia zauważalną lukę w badaniach empirycznych nad używalnością jednojęzycznych słowników elektronicznych w kontekście poszukiwania ekwiwalentów przekładowych połączeń wyrazowych. Jak wskazuje Autorka we wstępie (str. 2), powołując się na wyniki badania Laufer z 2011 r.¹, wyszukiwanie kolokacji w słownikach jednojęzycznych i poprawne ich tworzenie w mowie i piśmie to poważne wyzwanie dla uczących się języka angielskiego (poprawność kolokacji wynosi mniej niż 50%). Wyniki badania przeprowadzonego przez Autorkę rzucają zatem nowe światło na sposób korzystania ze słowników jednojęzycznych (ogólnych) w sytuacji tłumaczenia tekstu specjalistycznego (prawnego/prawniczego) i w moim odczuciu mogą okazać się pomocne dla konstruktorów leksykalnych baz danych (w tym relacyjnych) zawierających frazeologię specjalistyczną. Wyniki i wnioski zachęcają także do przeprowadzenia dalszych badań porównawczych z wykorzystaniem zasobów innych niż słowniki jednojęzyczne – mamy tu na myśli leksykalne bazy danych, a także zasoby tekstowe, np. korpusy równoległe, porównywalne lub nowe rodzaje korpusów mieszanych (tzw. *comparable corpora*, cf. Bernardini 2023²).
- Na pochwałę zasługuje syntetyczny rozdział teoretyczny traktujący o różnych spojrzeniach na zjawisko kolokacyjności i samej kolokacji, która jest jednostką analizy przyjętą w badaniu. Autorka cytuje najważniejsze prace wywodzące się – co zrozumiałe – z tradycji anglosaskiej. Odrobinę niedosytu pozostawia brak szerszych odniesień do polskojęzycznej tradycji badań frazeologicznych, wszak w badaniu empirycznym Autorka zestawia polskie i angielskie połączenia wyrazów. Walorem tej części pracy jest synteza badań nad aktywną i pasywną znajomością kolokacji oraz ich rolę w przyswajaniu języka obcego (uwzględniając zjawisko transferu językowego). Ze względu na specyfikę badania obszerna część rozdziału teoretycznego ukierunkowana jest na kolokacyjność w kontekście leksykografii, gdzie Autorka nawiązuje również m.in. do wcześniejszych wyników badań nad używalnością słowników (str. 16-26) i umiejętnie wykorzystuje je przy podejmowaniu decyzji metodologicznych w odniesieniu do własnego badania;

¹ Cyt. w dysertacji.

² Bernardini, S. (2023). „Translation is changing, and so should corpora: complex parallel corpus designs for translation research”. Plenary lecture delivered at the 4th International Symposium on Parallel Corpora, Leon, Spain, 22-24 November 2023.

- Autorka badania stawia nie tylko hipotezę opisową (pytanie badawcze nr 1), ale również wyjaśniającą (pytanie badawcze nr 2 i 3), co koresponduje z dzisiejszymi trendami w prowadzeniu badań empirycznych w językoznawstwie.
- Poprawnie zaprojektowane badanie empiryczne, którego metodologia umożliwia udzielenie odpowiedzi na postawione w pracy pytania badawcze. Szczególnie chciałbym pochwalić Autorkę za opracowanie zadań testowych do wykonania w trybie online, a także wykorzystanie testu LexTALE do pomiaru poziomu kompetencji językowej studentów-filologów, którzy zgodzili się uczestniczyć w badaniu.
- Zastosowanie metod regresji i przeprowadzenie analiz statystycznych w środowisku programistycznym R, z wykorzystaniem odpowiednich bibliotek programistycznych, np. lme4 (Bates i in. 2015)³. Co jeszcze ważniejsze: Autorka wykazała również umiejętność przygotowania danych do analizy (przygotowanie danych, czyszczenie danych, kodowanie zmiennych), interpretacji uzyskanych wyników i metryk statystycznych; Autorka potrafi spośród wielu modeli wybrać ten najprostszy i optymalnie dopasowany do danych. Jest to niewątpliwie ogromna wartość dodana pracy.

Uwagi do rozprawy doktorskiej

W trakcie lektury rozprawy zauważyłem nieliczne fragmenty, w odniesieniu do których chciałbym wyrazić swoje uwagi krytyczne. Chciałbym zaznaczyć, że większość z nich ma jedynie charakter polemiczny i nie rzutuje na moją wysoką ocenę recenzowanej rozprawy.

- W części teoretycznej zabrakło mi nawiązania do koncepcji ‘rozszerzonej jednostki znaczenia’ (*extended unit of meaning*) zaproponowanej przez Johna Sinclaira w 1996⁴, obejmującej pięć elementów, tj. koligację, kolokację, preferencję semantyczną, prozodię semantyczną oraz niezmienny rdzeń takiej jednostki. Autorka implicytnie czerpie z czterech elementów tej koncepcji, za wyjątkiem prozodii semantycznej, zwłaszcza na etapie wyboru jednostek analizy. Jednocześnie ciekawi mnie zdanie Autorki, czy prozodia semantyczna (tj. nacechowanie kolokatów względem ośrodka kolokacji i odwrotnie) może mieć wpływ na poprawny wybór kolokacji w tłumaczeniu, również specjalistycznym⁵.

³ Cyt. w dysertacji.

⁴ Sinclair J. 1996. “The Search for Units of Meaning”. TEXTUS IX (1), 71-106.

⁵ Moja uwaga jest nieprzypadkowa, gdyż zauważyłem podjąłem próbę zdiagnozowania takiej sytuacji na przykładzie nienormatywnego połączenia wyrazów napotkanego w tłumaczeniu (Grabowski, Ł. (2022).

- Jeśli idzie o miary siły kolokacji wykorzystane w badaniu to można w przyszłości rozważyć wykorzystanie wskaźnika PMI (*pointwise mutual information score*) dla połączeń polskich, z zastrzeżeniem, że jest on mocno skorelowany z kompozycyjnością połączeń rzeczownika z przydawką⁶.
- Autorka jest świadoma licznych ograniczeń przeprowadzonego badania, które opisuje na str. 78-79. Zastanawiające jest to, że jedynie 29% uczestników skorzystało z opcji rozwinięcia informacji o połączeniach wyrazowych po kliknięciu w przygotowany link, a wtedy tylko ok. 30% decyzji tłumaczeniowych można uznać za poprawną kolokację w angielszczyźnie. Na str. 79 Autorka podejrzewa również, że niektórzy studenci mogli podjąć próbę tłumaczenia fragmentów zdań z kolokacjami bez konsultacji słownikowej. Można zatem przyjąć, że wyniki badania mówią wiele o nawykach (nie)korzystania ze słowników przez (próbę) studentów kierunków filologicznych⁷. Zastanawiam się jednak, czy wyniki byłyby zupełnie lub znacząco inne, gdyby poproszono studentów o skorzystanie z autentycznych słowników internetowych zamiast spreparowanych artykułów hasłowych. Wymagałoby to zapewne wideo-rejestracji pracy studentów podczas wykonywania zadań. Rodzi się też kolejne pytanie dotyczące tzw. trafności ekologicznej (*ecological validity*) przeprowadzonego eksperymentu, tj. czy korzystanie z jednojęzycznych słowników pedagogicznych języka ogólnego jest obecnie typowym (naturalnym) scenariuszem podczas tłumaczenia tekstu specjalistycznego (frazologii specjalistycznej)? Być może nawet na zajęciach z przekładu specjalistycznego studenci nie posiłkują się w takich sytuacjach z tego typu słowników, a zamiast tego korzystają ze słowników specjalistycznych, baz tłumaczeń online (np. ReversoContext), korpusów równoległych (za taki można również uznać również bazę EurLex) itp.
- Na str. 79 Autorka zauważa, że niektórzy studenci mogli uznać, że wymagano od nich znajomości wybranych kolokacji i stąd uznali, że zadania można wykonać bez konsultacji ze słownikiem tudzież mogli konsultować się z innymi źródłami informacji. Ponieważ mogło to wpłynąć w jakiejś mierze na wynik to chciałbym zapytać, czy zebrano takie informacje w ankiecie po przeprowadzeniu badania?

“Provoke or encourage improvements? On semantic prosody in English-to-Polish translation”. *Perspectives: Studies in Translation Theory and Practice* 30, 120-136).

⁶ Maziarz, M., Grabowski, Ł., Piotrowski, T., Rudnicka, E., Piasecki, M. (2023). “Lexicalisation of Polish and English word combinations: an empirical study”. *Poznan Studies in Contemporary Linguistics*.

⁷ Jest to niewątpliwie cenna informacja również dla dydaktyków na kierunkach filologicznych.

- Na str. 78 Autorka wspomina o teorii obciążenia kognitywnego (Sweller 1988) w kontekście wyniku pokazującego, że w wypadku, gdy informacja o wzorcach kolokacyjnych była podana w artykule hasłowym od razu (tj. bez kliknięcia w link) wykonanie zadania trwało dłużej (a poprawność wyboru była wtedy wyższa). Można w tym kontekście nawiązać również do tzw. liczby Millera (7 ± 2)⁸, którą ludzki mózg jest w stanie jednorazowo i bezpośrednio przetworzyć.
- W trakcie lektury pracy nie zauważyłem nigdzie linka do danych badawczych oraz metadanych. Zachęcam – być może przy okazji publikacji badania – do umieszczenia zarówno danych źródłowych, jak i kodu programistycznego w jednym z otwartych repozytoriów, co umożliwi m.in. replikację badania.

Ocena strony językowej i edycyjnej rozprawy

Strona językowa i redakcyjna recenzowanej dysertacji zasługuje na wysoką ocenę – Autorka posługuje się dojrzałą angielszczyzną na poziomie akademickim. Rozprawa jest również doskonale przygotowana pod względem edycyjnym. Na szczególną pochwałę zasługuje klarowność, precyzja i logiczność wywodu naukowego. Zauważyłem nieliczne, drobne błędy edycyjne (podwajanie nawiasów), np. (Goźdź-Roszkowski (2013: 4) – na str. 29, (Wickham et al. 2021)) – str. 73.

Konkluzja

Moim zdaniem rozprawa doktorska mgr Pauliny Halczak pt. *Collocations in legal language in online monolingual English learners' dictionaries*, przygotowana pod kierunkiem prof. prof. Roberta Lwa i Thomasa Herbsta spełnia wszystkie wymagania stawiane dysertacjom doktorskim w obowiązujących przepisach prawa. Uważam, że recenzowana dysertacja winna jak najszybciej ukazać się drukiem, a dane badawcze wraz z metadanymi udostępnione w otwartym repozytorium. W związku z powyższym jednoznacznie wnioskuję o dopuszczenie Doktorantki do dalszych etapów przewodu doktorskiego.

/ prof. dr hab. Łukasz Grabowski /



⁸ Miller, G. (1956). „The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information”. *The Psychological Review* 63, 81-97.