

Opole, 23 lipca 2020 r.

dr hab. Łukasz Grabowski

Instytut Językoznawstwa

Uniwersytet Opolski

Recenzja rozprawy doktorskiej mgr Aleksandry Beaty Makowskiej pt. *Terminologia i frazeologia w artykułach naukowych z dziedzin nowych technologii. Studium korpusowe w kognitywnym ujęciu porównawczym angielsko-polskim* (ss. 432) przygotowanej pod kierunkiem prof. UŁ. dra hab. Stanisława Goździa-Roszkowskiego

Uwagi wstępne

Recenzowana rozprawa doktorska pt. *Terminologia i frazeologia w artykułach naukowych z dziedzin nowych technologii. Studium korpusowe w kognitywnym ujęciu porównawczym angielsko-polskim* to w istocie badanie korpusowe nad najczęściej używaną leksyką i frazeologią w specjalnie opracowanym przez Autorkę korpusie tekstów obejmującym artykuły naukowe z dziedziny mikroelektroniki, telekomunikacji i informatyki. W dysertacji, opatrzonej rozbudowanym wstępem teoretycznym poświęconym syntezie stanu badań nad dyskursem naukowym, Autorka podejmuje niełatwą próbę charakterystyki cech gatunkowych, a także leksyki, frazeologii i terminologii w ujęciu porównawczym (angielsko-polskim). Ze względu na własne zainteresowania naukowe w zakresie językoznawstwa korpusowego w mojej recenzji skupię się przede wszystkim na ocenie samej metodologii badania, a także interpretacji wyników analizy. W trakcie lektury rozprawy starałem się zwrócić uwagę na jej mocne, jak i słabe strony, które poniżej dokładniej skomentuję.

Struktura rozprawy

Recenzowana dysertacja liczy 432 strony; składa się ze wstępu, 9 rozdziałów (5 pierwszych rozdziałów teoretycznych, rozdział 6. metodologiczny, analityczne rozdziały 7-9), podsumowania, obszernej bibliografii, streszczeń w języku angielskim i polskim, a także 2

pokaźnych załączników składających się na opis materiału badawczego. Struktura pracy zyskałaby na przejrzystości, gdyby wyraźniej wyodrębniono w niej część teoretyczną, metodologiczną i analityczną. Ponadto w spisie treści rzuca się w oczy wyraźna dysproporcja pomiędzy bardzo rozbudowaną częścią teoretyczną (o czym będzie jeszcze w recenzji mowa) liczącą prawie 240 stron, krótką i bardzo pobieżną częścią metodologiczną (25 stron), a częścią analityczną liczącą 110 stron. Nieprzypadkowo w trakcie lektury pracy czytelnik odnosi wrażenie, że jej najważniejszą częścią i jednocześnie celem jest synteza różnych podejść teoretycznych do analizy dyskursu naukowego, a nie samo badanie empiryczne, co sugerowałby tytuł rozprawy.

Wstęp

Lektura wstępu na stronach 7-8 pozwala sądzić, że cele postawione w pracy są bardzo szerokie, jednak nadrzędnym celem wydaje się być wyodrębnienie wyróżniających się, tudzież dystynktywnych, cech językowych na przykładzie angielskich i polskich tekstów naukowych z dziedziny nowych technologii, z uwzględnieniem charakterystyki pojęć, terminów złożonych, zapożyczeń, skrótowców, neologizmów, które w zamierzeniu zoperacjonalizowano jako słowa kluczowe, zbitki wielowyrazowe (*lexical bundles*)¹ i kolokacje, a także opisano z wykorzystaniem takich procedur badawczych, jak analiza konkordancji i list frekwencyjnych. Oprócz tego Autorka stawia sobie za cel przestudiowanie zjawisk polisemii i synonimii w pozyskanej terminologii, określenia stopnia standaryzacji terminologii, określenia statusu teoretycznego i terminologicznego kolokacji, określenie funkcji tekstowych pozyskanych frazeologizmów, pozycjonowania autora w tekście, a wszystko to również w perspektywie porównawczej angielsko-polskiej.

Część teoretyczna rozprawy

W mojej ocenie najmocniejszą stroną recenzowanej rozprawy jest jej część teoretyczna, obejmująca aż 5 rozdziałów. Doktorantka wychodzi od ogólnych spostrzeżeń dotyczących powiązań między nauką, wiedzą i komunikacją, gdzie jako spoiwo jawi się społeczność naukowa traktowana jako grupa praktyków, wspólnota interesów, a także grupa zawodowa. Następnie omówiono zarówno cechy językowe dyskursu naukowego, tj. specjalistycznego

¹ Jak wyjaśnię później, w praktyce Autorka analizuje n-gramy, a nie zbitki wielowyrazowe/leksykalnej w ujęciu Bibera i in. (1999).

języka naukowego i technicznego, rozbudowaną typologię tekstów konstytuujących rzeczony dyskurs, a także pozycję i rolę języka angielskiego w globalnej komunikacji naukowej. W tym miejscu Autorka zawarła wiele cennych spostrzeżeń, dotyczących m.in. uprzywilejowanej pozycji rodzimych użytkowników tego języka (str. 58-64); rozróżnień pojęciowych osadzonych w literaturze z dziedziny socjolingwistyki (język, odmiana języka, technolekt, rejestr, żargon, styl itd.); rozróżnienia pomiędzy językiem i komunikacją naukową a techniczną; charakterystyki artykułu naukowego jako funkcjonalnego gatunku tekstu, podkreślającej różnice w strukturze i konwencjach stylistycznych takich tekstów w zależności od dyscypliny i kultury naukowej (w krajach anglojęzycznych i w Polsce, ss. 118-126). Brakuje tutaj nawiązania do kilku badań empirycznych (m.in. Bielski & Bielska 2008; Donesch-Jeżo 2013, 2016) traktujących o różnicach leksykalnych i retorycznych między abstraktami artykułów naukowych napisanych w języku polskim i angielskim (z dziedziny medycyny i językoznawstwa), choć Autorkę niejako usprawiedliwia fakt, iż wspomniane badania nie dotyczyły nauk technicznych. Dalej szeroko i wyczerpująco omówiono kluczowe pojęcia dotyczące jednostek analizy, tj. terminologię (w tym zwrócono uwagę na neologizmy i zapożyczenia terminologiczne) i frazeologię², a także różne podejścia badawcze w tym zakresie (tj. bardziej tradycyjne - formalne, funkcjonalne, jak i dystrybucyjne), po czym zwrócono uwagę (Rozdział 4) na rolę korpusów językowych (porównywalnych i równoległych) i metodologii korpusowej w takich badaniach. Można tam znaleźć wiele cennych spostrzeżeń (np. dotyczących kryteriów kompilacji korpusów porównywalnych), jednak niektóre sformułowania są nazbyt kategoryczne, np. na str. 208 Autorka pisze, że „analiza słów kluczowych jest jedną z najbardziej efektywnych metod określenia dziedzinowego słownictwa”, choć w zależności od testu statystycznego, miary siły efektu, a także korpusu referencyjnego wyniki takiej analizy mogą również uwypuklić wyrazy funkcyjne, determinujące styl i gramatykę tekstu naukowego (na str. 212 Autorka sama przytacza przykłady tego typu). Autorka jednak słusznie zauważa, że w celu wyodrębnienia terminów spośród wykazu słów kluczowych niezbędna jest konsultacja ze specjalistą z danej dziedziny. Wszystkie te rozdziały są bardzo rozbudowane, a Doktorantka możliwie głęboko czerpie z literatury anglojęzycznej, niemieckojęzycznej, jak i polonistycznej opublikowanej

² Na str. 199 Autorka podkreśla trudność w rozróżnieniu pomiędzy terminami złożonymi a kolokacjami wskazując na fakt, iż kolokacje mogą stanowić składowe takich terminów. Trudno się z tym nie zgodzić, jednak wyraźnie należy tutaj podkreślić, że zależy to w dużej mierze od tego, jak zdefiniujemy i zoperacjonalizujemy samą kolokację (np. stosując jedynie kryteria dystrybucyjne tudzież również kryteria formalne).

zarówno kilka dekad temu, jak i współcześnie. Za tę wyczerpującą część pracy i trud włożony w jej przygotowanie należą się Autorce niewątpliwie słowa pochwały, jednak zastanawiam się jednocześnie, na ile – zważywszy na cele pracy i materiał badawczy (próbki jednego gatunku tekstu: artykułów naukowych) – niezbędna jest aż tak rozbudowana synteza teoretyczna, która dotyczy przede wszystkim bardzo szeroko rozumianych problemów komunikacji naukowej. Kolejny rozdział (5.) traktuje o problemach związanych z tłumaczeniem tekstów naukowych i technicznych, a także podejmuje dyskusję nad niezbędnymi kompetencjami tłumacza w tym zakresie. Zważywszy jednak na cele pracy wyartykułowane we wstępie (a także w późniejszym rozdziale metodologicznym, o czym będzie jeszcze mowa), nie do końca widzę uzasadnienie dla podejmowania tej problematyki, wszak praca nie traktuje bezpośrednio o tłumaczeniu. Jedyne wytłumaczenie, które przychodzi mi do głowy (ale nie wynika wprost z tekstu) to założenie, że wyniki analizy porównawczej mogą okazać się użyteczne w praktyce przekładu specjalistycznego.

Metodologia

Lektura rozdziału 6., metodologicznego, przynosi wiele wątpliwości wynikających z braku precyzji. Opisując materiał badawczy, Doktorantka wskazuje na źródła pozyskania tekstów naukowych, jednak nie wyjaśniła dokładnie, czy owe artykuły pochodzą z czasopism naukowych, czy też z materiałów konferencyjnych. Na przykład źródło opisane jako MIXDES 2013, które jest bliżej niewyjaśnione, to konferencja „20th International Conference Mixed Design of Integrated Circuits and Systems”, która odbyła się 20-22 czerwca 2013 w Gdyni.³ W odniesieniu do innych źródeł również można mieć wątpliwości, np. co oznacza nazwa „Elektronika⁴ & other papers”. Ma to o tyle istotne znaczenie, że artykuł naukowy opublikowany w czasopiśmie może istotnie różnić się – pod względem rozmiaru, struktury itp. – od tekstu w materiałach konferencyjnych. Ponadto dziwi brak bliższego wyjaśnienia źródeł, z których Doktorantka pozyskała materiał badawczy, ponieważ czytelnik-językoznawca, a także nauczyciel języka specjalistycznego, może zwyczajnie nie wiedzieć, o jakie źródła chodzi (np. Ślesin papers, Warsaw papers), a także nie jest w stanie ocenić jednorodności zebranego materiału. Ma to również istotne znaczenie jeśli idzie o ewentualną replikowalność tudzież powtarzalność badania w przyszłości. Autorka nie

³ <https://www.mixdes.org/Mixdes3/yearSummary/viewConf/2013>

⁴ Zapewne chodzi tutaj o czasopismo branżowe dostępne pod adresem internetowym: <https://elektronika.orf.pl/>

udostępnia zebranego materiału badawczego w domenie publicznej (np. w repozytorium danych osf.io), a dane w tabeli 6.1 pozostawiają pewne niejasności, np. „558 (?)” przy liczbie wyrazów pochodzących z czasopisma *Elektronika*. Nie do końca wiadomo również, ile łącznie zebrano polskich tekstów (na str. 248 pojawia się enigmatyczna liczba 103), a pod rzezoną tabelą Autorka przyznaje, że nie wszystkie teksty zostały umiejętnie skompilowane. Taka sytuacja nie powinna mieć miejsca. Ponadto niejasne jest stwierdzenie, że artykuły napisała „wystarczająca liczba autorów” (str. 248). Trudno zgodzić się ze wcześniejszym stwierdzeniem Doktorantki na str. 247, że większa liczba zebranych tekstów w języku angielskim świadczy o dominującej pozycji języka angielskiego. W rzeczy samej to przecież osoba kompilująca porównywalny korpus badawczy odpowiednio „ustawia” parametry takich kryteriów, jak liczba tekstów, próbkowanie, reprezentatywność, zbilansowanie i rozmiar korpusu. Kilka stron wcześniej (str. 204) Doktorantka cytuje jednak stanowisko Belindy Maia (2003): “There must be a balance among samples, i.e. there must be a comparable quantity of texts meeting the predefined criteria (...)” Trudno się również zgodzić, że różnice w odniesieniu do liczby wyrazów w korpusach wynikają z tego, że „język polski nie posiada wyrazów gramatycznych” (str. 247). Trudno zasłaniać się tutaj ograniczeniami oprogramowania SketchEngine (Kilgarriff i in. 2014), bo nawet jeśli niektóre pliki źródłowe były w formacie pdf to Doktorantka winna była – w związku z zaobserwowanymi problemami – wykonać procedurę OCR i przetworzyć je do formatu edytowalnego (np. txt), przy okazji dokonując niezbędnych korekt i zapisując każdy tekst jako oddzielny plik. Widać to wyraźnie na str. 263 (Tabela 6.8), gdzie rzucają się w oczy problemy z segmentacją wyrazów i kodowaniem polskich znaków. Więcej informacji na temat korpusu przynoszą str. 248-261: widać, jak wysoce niejednorodny jest korpus badawczy, np. w przypadku angielskiego podkorpusu MTCS (str. 259) większość autorów to Polacy (369), podczas gdy rodzimi użytkownicy języka angielskiego to jedynie 5 osób (opierając się na kryterium narodowości, które nie musi w pełni oddawać rzeczywistego języka rodzimego). Tym samym stworzony korpus umożliwia raczej porównanie międzynarodowej (nienatywnej) odmiany angielskiego języka naukowego z trzech wybranych dziedzin z językiem polskim.

Do analizy Doktorantka zdecydowała się wykorzystać komercyjne narzędzie SketchEngine, jednak nie wspomina nigdzie o autorach oprogramowania⁵. Warto również wskazać, że

⁵ Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P. & Suchomel, V. (2014). The Sketch Engine: ten years on. *Lexicography* 1(1), 7-36.

oprogramowanie nie narzuca korpusu referencyjnego, ponieważ to sam użytkownik dokonuje wyboru w tym zakresie spośród wielu dostępnych korpusów. Niemniej jednak wybór Doktorantki jest w tym miejscu odpowiedni i uzasadniony. Wydaje się, że pożądanym byłby komentarz wyjaśniający Autorki, dlaczego prócz korpusów języka ogólnego (BNC i NKJP) zdecydowała się użyć jako korpusu referencyjnego Cambridge Academic English? O podobnym korpusie akademickim w odniesieniu do języka polskiego nie wspomniano w pracy, co może rodzić wątpliwości w odniesieniu do porównywalności procedur badawczych w odniesieniu do obu podkorpusów.

Na str. 264-265 Doktorantka opisuje cele badania, które dotyczą m.in. porównania sposobów konceptualizacji terminologii specjalistycznej, określenia statusu terminologicznego kolokacji, wskazania różnic w zakresie wykorzystanych frazeologizmów, ocenienia wpływu interdyscyplinarności na terminologię i kolokacje w analizowanych tekstach, porównania struktury analizowanych tekstów w języku polskim i angielskim, a także obecności Autora w tekstach. Wspomniane cele są bardzo szerokie, różnią się znacznie od tych opisanych we wstępie pracy (gdzie były bardziej precyzyjne i częściowo zoperacjonalizowane) – w tym miejscu jawią się one jako bardzo ogólne i czytelnik nie wie do końca, w jaki sposób mogą zostać osiągnięte. Częściową odpowiedź na to pytanie znajdujemy na str. 266-267, gdzie precyzując jednostki analizy Autorka pisze o słowach kluczowych i terminach, a także n-gramach. Niestety nie dowiadujemy się z tej części pracy, jaką metodę ekstrakcji słów kluczowych wykorzystuje program, a także dlaczego Autorka zdecydowała się na słowa kluczowe usytuowane między rzadkimi i średnio-częstymi (i co taki wybór oznacza), o czym dowiadujemy się jedynie z ryciny na str. 6.11 (zrzut ekranu). Autorka pisze dalej o analizie tzw. n-gramów (czyli ciągów n-wyrazów), które traktuje jako równoważne zbitkom wielowyrazowym/leksykalnym (*lexical bundles*) w ujęciu Bibera i in. (1999, 2003, 2004). Trudno się z tym zgodzić, ponieważ w wypadku zbitek leksykalnych kryterium ekstrakcji to prócz długości i frekwencji również dystrybucja w poszczególnych tekstach, natomiast w recenzowanym badaniu nie może być o tym mowy, gdyż Doktorantka traktuje poszczególne podkorpusy jako całość, a nie jako zbiór poszczególnych dokumentów. Ponadto biorąc pod uwagę różnice typologiczne między językiem angielskim i polskim dość kontrowersyjne wydaje się porównywanie n-gramów o podobnej długości (np. 4-wyrazowych), np. n-gram *the end of the* może odpowiadać polskiemu pojedynczemu wyrazowi *koniec*. Ponadto

Autorka pisze na str. 270, że porównanie jednostek frazeologicznych (czyli w praktyce n-gramów) będzie wykonane na podstawie wytycznych da Silvy i in, (2017), którzy badali teksty akademickie, jednak mam w tym miejscu wątpliwości wymagające wyjaśnienia, czy przytoczeni Autorzy również badali wyłącznie artykuły naukowe w czasopismach bądź materiałach konferencyjnych. Wyjaśnienie w tej materii byłoby wysoce pożądane. Na koniec warto zaznaczyć, że Doktorantka nie wspomina w tym miejscu m.in. o neologizmach i zapożyczeniach, które wyłaniają się z późniejszych analiz, więc czytelnik domyśla się jedynie, że takowe zostaną określone przy okazji analiz słów kluczowych, terminów i n-gramów. Szkoda, że nie jest to eksplicitnie wyjaśnione.

W odniesieniu do metodologii mam jeszcze dwa ważne zastrzeżenia: z tytułu dysertacji wynika jasno, że Doktorantka przeprowadziła analizę porównawczą, jednak w treści rozdziału brak precyzyjnie określonego *tertium comparationis*, tj. swoistego wspólnego mianownika względem którego porównane będą teksty angielskie i polskie. Ponadto prócz opisu celów badania, w treści krótkiego i dość powierzchownego rozdziału metodologicznego nie znajdujemy precyzyjnie określonych pytań badawczych, na które oczekujemy szczegółowych odpowiedzi w oparciu o zoperacjonalizowane procedury badawcze zastosowane względem wyraźnie określonych jednostek analizy. Brakuje również eksplicitnie wyrażonych hipotez podlegających weryfikacji, jak ma to zazwyczaj miejsce w badaniach empirycznych. Rodzi to pewne wątpliwości, ponieważ wszelkie wnioski i rekomendacje wynikające z pracy nieuniknienie będzie cechować niska precyzja i wysoki stopień ogólnikowości, co widać również w konkluzjach i podsumowaniu (np. str. 388 w odniesieniu do różnic w zakresie konceptualizacji pojęć). Wydaje się, że sama Doktorantka nie jest zadowolona z niektórych decyzji metodologicznych i ich wykonania, zwłaszcza w odniesieniu do kompilacji korpusu badawczego, określając swoje dokonania mianem połowicznego sukcesu (str. 271: „the final effect is, however, only a partial success”).

Wyniki badania

Ogólnie rzecz ujmując (i nie zważając na opisane niedociągnięcia metodologiczne), wyniki badania przedstawione w rozdziałach empirycznych dostarczają zarówno nowych, interesujących danych językowych w zakresie opisu prac naukowych z dziedziny mikroelektroniki, telekomunikacji i informatyki, jak i wielu interesujących spostrzeżeń w

odniesieniu do różnic między angielszczyzną i polszczyzną. Na przykład w odniesieniu do analizy makrostruktury (Rozdział 7) Autorka⁶ zauważa, że autorzy polskich prace nie stosują konsekwentnie tzw. modelu IMRAD (str. 275), natomiast abstrakty angielskich i polskich prac rzadko można scharakteryzować w oparciu o pełen model CARS (Swales 1990). Zastanawiam się w tym miejscu, na ile owe wyniki można generalizować do innych tekstów dziedzinowych, gdyż w moim odczuciu mogą one być zwyczajnie wypadkową specyficznego materiału badawczego zebranego przez Autorkę (wszakże przygniatająca większość autorów tekstów angielskich to nie rodzimi użytkownicy tego języka). Przeprowadzona analiza, polegająca na liczeniu tzw. kroków retorycznych, musiała być bardzo czasochłonna, stąd szkoda, że Doktorantka nie przybliżyła czytelnikowi jakichś wskazówek odnośnie narzędzi tudzież makr, z których zapewne korzystała (a może się mylę i wszystkie owe kroki zliczano ręcznie). Do porównawczej analizy multimodalności Autorka – całkiem zresztą słusznie, choć opisuje to dopiero w tym miejscu – wykorzystuje tzw. wordsketches w odniesieniu do par wyrazów, takich jak np. *rysunek* i *figure*, *table* i *tabela* (str. 284), która w praktyce sprowadza się do porównania wybranych profili kolokacyjnych.

Rozdział 8. poświęcono analizie terminologii, również wykonanej przy pomocy oprogramowania SketchEngine, które – korzystając z wcześniej otagowanego materiału tekstowego – pozwala wyodrębnić z tekstów głównie frazy nominalne i przymiotnikowe. Pożądanym byłby komentarz Doktorantki w odniesieniu do wpływu tematyki analizowanych tekstów na wyniki badania, tj. na wysoki stopień zróżnicowania terminologicznego. Jeśli idzie o analizę neologizmów (str. 297 i dalej) to brakuje mi wyjaśnienia, czy Autorka wykorzystowała również kryterium słownikowe (tj. czy dany termin zarejestrowano wcześniej w słownikach specjalistycznych), aby upewnić się co do statusu neologizmu. Mimo tej uwagi, wyniki analizy dostarczają ciekawego materiału językowego, który Autorka przeanalizowała również pod kątem etymologicznym. W podobny sposób przeanalizowano również zapożyczenia (str. 304-307). Równie ciekawych danych dostarcza analiza akronimów, skrótowców i innych symboli, których powstanie i częste użycie wynika z mechanizmu kompresji informacji w tekstach – być może w tym miejscu warto by również

⁶ Niektóre fragmenty pracy zostały częściowo opublikowane przez Autorkę wcześniej, np. w artykule pt. „Research Articles as a Means of Communicating Science: Polish and Global Conventions” (Makowska 2018, w tomie pt. *University Writing in Central and Eastern Europe: Tradition, Transition, and Innovation*, red. M. Chitez, C. I. Doroholschi, O. Kruse, Ł. Salski i D. Tucan, wyd. Springer). Zaleca się przytaczanie takich faktów, jeśli rzeczywiście miały miejsce.

wspomnieć o dobrze znanej zasadzie relewancji, tj. maksymalizacji przekazu informacji przy minimalizacji wysiłku kognitywnego (Sperber i Wilson 1986). Komponent kognitywny w pracy jest potraktowany dość pobieżnie (odnosi się do opisu terminów, które wynikają z metaforycznego postrzegania zjawisk i bytów, a same metafory i tzw. rozszerzenia metaforyczne zaczerpnięto z wyników analizy kolokacji), jednak wyniki tej części analizy są również bardzo interesujące poznawczo (np. str. 313-314). Podobnie rzecz się ma w wypadku wyników analizy synonimów oraz terminów polisemicznych (str. 317-324). Niemniej jednak mam w tym miejscu pewne zastrzeżenie: otóż Doktorantka używa szerokiego wachlarza terminów, np. synonimy, pseudo synonimy tudzież synonimy fałszywe (str. 320), jednak brakuje w treści rozdziału ich precyzyjnych definicji. Ponadto brak eksplicytnie wyrażonego *tertium comparationis* powoduje, że często analiza porównawcza sprowadza się do omówienia przypadkowo wybranych ekwiwalentów formalnych (np. *acquisiton/akwizycja*). Z drugiej strony bardzo oryginalnym rozwiązaniem jest wykorzystanie danych pozyskanych przy pomocy programu SketchEngine do opracowania tudzież doprecyzowania ontologii terminów specjalistycznych (str. 336-340).

Rozdział 9. przynosi wyniki analizy kolokacji i odtwarzalnych n-gramów. Warto w tym miejscu nadmienić, że obecnie w odniesieniu do kolokacji używa się częściej terminów *core* (ośrodek kolokacji) i *collocate* (kolokat/wyraz współwystępujący) aniżeli *base* i *collocator*. Analiza kolokacji została przeprowadzona w sposób poprawny: w pierw Doktorantka wygenerowała listę terminów jednowyrazowych dla tekstów angielskich (40) i polskich (39) w oparciu o progi frekwencji i określone okno kolokacji, a następnie wyodrębniła najczęstsze ich kolokaty. Wyjaśnienia wymaga w tym miejscu, dlaczego lista polskich terminów nie została poddana lematyzacji (np. *filtru, filtracji* jako ośrodki kolokacji)⁷. Niemniej jednak sama analiza dostarcza bogatych danych językowych i stanowi chyba najcenniejszą część recenzowanej pracy, podobnie jak reszta omawianego rozdziału, gdzie znajdujemy również

⁷ Można do tego celu w przyszłości wykorzystać narzędzie MEWEX opracowane w ramach konsorcjum CLARIN-PL i służące do automatycznego wyodrębniania kolokacji i dłuższych ciągów wielowyrazowych z tekstów napisanych w języku polskim (<https://ws.clarin-pl.eu/mewex.shtml>). Więcej: Piasecki, M. Wendelberger, M., Maziarz, M. (2015). *Extraction of the multi-word lexical units in the perspective of the wordnet expansion. Proceedings of the International Conference Recent Advances in Natural Language Processing*, 512-520. Innym narzędziem umożliwiającym lematyzację jest TermoPL, które jest obecnie zintegrowane z dostępnym w chmurze narzędziem Korpusomat (www.korpusomat.pl). Więcej: Kieraś, W., Kobyliński, Ł. & Ogrodniczuk, M. (2018). Korpusomat — a tool for creating searchable morphosyntactically tagged corpora. *Computational Methods in Science and Technology*, 24(1):21–27. Pozwoliłoby to zwiększyć jakość polskich n-gramów, wyszczególnionych w załączniku do ocenianej rozprawy (w istocie wiele z nich to niepełne struktury składniowe, zazębiające się ze sobą ciągi wyrazów itp.).

wyniki analizy terminów wielowyrazowych i n-gramów, tych ostatnich również w aspekcie funkcjonalnym (podejście Bibera i in. 1999, 2003, 2004). W nawiązaniu do wcześniej sygnalizowanego przeze mnie problemu w rozdziale metodologicznym, wyniki analizy n-gramów dość jednoznacznie wskazują, że w przypadku tak różnych typologicznie (w odniesieniu do morfologii) języków, porównywanie zbitek o tej samej długości obarczone jest utratą cennych informacji, np. (str. 366) angielskiemu n-gramowi *it is possible to* może w polszczyźnie odpowiadać czasownik modalny *można*. Widać również (Tabela 9.13), że właśnie z tego względu polskie n-gramy w zdecydowanie mniejszym stopniu stanowią jedność formy i znaczenia (funkcji dyskursywnej), np. *można podzielić na dwie, można wykonać posługując się, zgadza się z lokalną* i wiele innych. Być może można było temu zaradzić stosując dodatkowo – jak zaleca się w literaturze – kryterium dystrybucyjne. Dodatkowy rozdział 10. to porównanie pozycjonowania się autorów analizowanych tekstów, który dostarcza ciekawych danych odnośnie preferowanych modeli narracji, również w zestawieniu do wcześniejszych badań Dontchevej-Navratilovej (2013). Wyniki badania pokazują wyraźnie m.in., że polscy autorzy unikają narracji w 1. os. l. poj. (str. 371).

Całość rozprawy wieńczy uwagi końcowe, gdzie Autorka podsumowuje i syntezuje uzyskane wyniki, a także wyszczególnia zaobserwowane wzorce leksykalne i terminologiczne. Treść tej części pracy jest niezwykle cenna, ale należy również dodać, że brak odniesienia do precyzyjnych pytań badawczych i zoperacjonalizowanych hipotez sprawia, że czytelnik ma wrażenie dużej ogólnikowości wywołu, np. „In addition, definitions are better depicted in English than Polish.” (str. 389); „Equivalent or nearly equivalent collocators are in fact form completely different phrases in English and Polish. On the other hand, there is a degree of overlap between generated collocations and multiword key terminology, which confirms previous studies on the subject but their structure is more diversified” (str. 387); “In addition, the predominant majority of scalar and construction collocators confirm the empirical and practical nature of the MTCS disciplines.” (str. 388); „The phraseological units together with the structure of the research paper contribute to the success of research, i.e. promoting the solutions as the most relevant alternative in the scientific arena.” (str. 390). Trudno się zgodzić z dość zaskakującym stwierdzeniem na str. 384, że oprogramowanie SketchEngine nie potrafi przetwarzać tekstów w językach fleksyjnych i syntetycznych, takich jak polski („it can be concluded that the program cannot process texts in synthetic, inflectional languages,

like Polish.”). Ów program dobrze radzi sobie z tagowaniem i parsowaniem tekstów polskich (a także napisanych w innych językach fleksyjnych), a zaobserwowane problemy wynikają ze sposobu przygotowania materiału badawczego, o czym wspominałem w swoich uwagach odnoszących się do rozdziału metodologicznego.

Ocena strony językowej i edycyjnej dysertacji

Biorąc pod uwagę dużą ilość tabel, zestawień i wykresów ocena strony edycyjnej pracy jest zdecydowanie pozytywna, a zastrzeżenia można mieć jedynie do przygotowania spisu treści., gdzie rzucają się w oczy niewyrównane numery stron, a także wyboldowana sekcja 8.13.

Ogólnie rzecz ujmując, praca jest napisana dojrzałą angielszczyzną i w miarę dobrze się ją czyta, jednak czasem razi brak pewnej konsekwencji stylistyczno-terminologicznej, np. raz Doktorantka używa przymiotnika *Saxon* (np. str. 84 i 118), a innym razem *saxonic* (np. str. 124 i 216) w odniesieniu do stylów pisarstwa. Zdarzają się również błędy w nazwiskach przytaczanych autorów, np. Hundston, 2008 (str. 203 i dalej), zamiast Hunston.

Niestety w trakcie lektury pracy widać wyraźnie, że językowo jest ona dość nierówna, tj. bardziej dopracowana w części teoretycznej i empirycznej, a mniej w części metodologicznej i podsumowaniu. Rzucają się również w oczy dość liczne błędy gramatyczne i literówki. Z recenzenckiego obowiązku przytaczam tylko niektóre przykłady (w nawiasach podaję numery stron):

- because the latter **are often are** constituents of the most salient terms (199)
- This computational program is designed for **linguistists**, translators or terminologists (261)
- characteristic **for** the research articles (261)
- **Additinally**, the author decided (270)
- scrutiny of **intersisciplinarity**, due to the **interdepent** character (271)
- For **thise** reason, **thee** author performed a (271)
- which is in alliance with Kozłowska’s (2009) findings **in terms the** behaviour of terminology in hard sciences (327)

- All things considered, the domains of new technologies **are relatively new empirical and practical**, thus borrowing terminology from other, more traditional subject fields. (330)
- Despite the fact that these categories originally **pertained English terminology** (358)
- **On the end**, cognitive mechanisms of terminology creation (389)

Mocne strony recenzowanej dysertacji

Reasumując, recenzowana dysertacja ma wiele zalet, które nadają jej twórczy i oryginalny charakter:

- oryginalny pomysł na porównanie leksyki, frazeologii i terminologii w pracach naukowych ze specjalistycznych dziedzin, takich jak mikroelektronika, telekomunikacja i informatyka, również w aspekcie porównawczym; dysertacja wypełnia ważną lukę w tym zakresie;
- zaprojektowanie i kompilacja korpusu badawczego;
- rozbudowana, wszechstronna i dopracowana część teoretyczna, wykraczająca poza zakres przeprowadzonego badania;
- cenne poznawczo wyniki badania, dostarczające nowych danych empirycznych, potencjalnie interesujące dla językoznawców, a także nauczycieli i badaczy języków specjalistycznych.

Słabe strony recenzowanej dysertacji

Wśród słabszych stron recenzowanej dysertacji należy podkreślić m.in.:

- brak równowagi pomiędzy częścią teoretyczną, metodologiczną i empiryczną;
- dość powierzchowna część metodologiczna, w której brakuje eksplicytnie wyrażonych i zoperacjonalizowanych pytań badawczych i hipotez;
- brak eksplicytnie wyrażonego *tertium comparationis*, co czasem skutkuje dość przypadkowym zestawieniem danych angielskich i polskich;

- niezrównoważony korpus badawczy, w którym nie mamy do czynienia z tekstami napisanymi przez rodzimych użytkowników języka angielskiego (za wyjątkiem kilku autorów); brak precyzyjnego opisu źródeł, z których pozyskano materiał badawczy (czytelnik-językoznawca może nie wiedzieć, czy mamy do czynienia z artykułami naukowymi opublikowanymi w czasopiśmie, czy też w materiałach konferencyjnych);
- liczne usterki, jeśli idzie o stronę językową analizowanej rozprawy;
- niektóre wnioski cechują się zbyt dużym poziomem ogólnikowości.

Konkluzja

Recenzowana dysertacja doktorska posiada słabsze strony, wynikające głównie z dość powierzchownego opisu metodologii, a także rozmytych, licznych i niezbyt precyzyjnie określonych celów badania. Niemniej jednak należy jednocześnie zauważyć szereg mocnych stron recenzowanego badania, którego wyniki dostarczają nowych i interesujących danych językowych, zwłaszcza dla nauczyciela tudzież badacza języków specjalistycznych. Moim zdaniem mocne strony pracy przeważają nad jej metodologicznymi niedociągnięciami, które z recenzenckiego obowiązku wypunktowałem w niniejszej opinii i które z całą pewnością nie dyskwalifikują dysertacji jako całości, tj. jako przemyślanego i pokaźnego wysiłku naukowego. W związku z powyższym stwierdzam, że rozprawa doktorska mgr Aleksandry Beaty Makowskiej pt. *Terminologia i frazeologia w artykułach naukowych z dziedzin nowych technologii. Studium korpusowe w kognitywnym ujęciu porównawczym angielsko-polskim* **spełnia wymagania minimum stawiane dysertacjom doktorskim i wnioskuję o dopuszczenie Doktorantki do dalszych etapów postępowania doktorskiego.**

Bibliografia

Biber, D., Johansson, S., Leech, G., Conrad, S. & Finegan, E. (1999). *The Longman grammar of spoken and written English*. London: Longman.

Biber, D., Conrad, S. & Cortes, V. (2003). "Lexical bundles in speech and writing: An initial taxonomy". [w:] A. Wilson, P. Rayson, & T. McEnery (Eds.), *Corpus Linguistics by the Lune: A Festschrift for Geoffrey Leech*. Frankfurt/Main: Peter Lang, 71-92.

Biber, D., Conrad, S. & Cortes, V. (2004). "If you look at...: Lexical bundles in university teaching and textbooks". *Applied Linguistics*, 25 (3), 371-405.

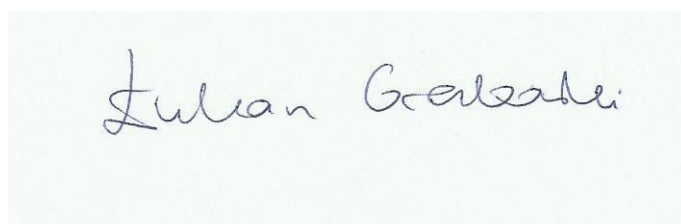
Bielski, M. & Bielska, J. (2008): "Analysing medical language: A study of Polish/English abstract translations". [w:] Burgess, S. & P. Martín-Martín (red.) *English as an Additional Language in Research Publication and Communication*. Bern: Peter Lang.

Donesch-Jeżo, E. (2013). *Akademicki dyskurs medyczny w językach angielskim i polskim w ujęciu lingwistycznym i metodycznym: artykuł naukowy i jego abstrakt*. Kraków: Fundacja Zdrowia Publicznego. Uniwersyteckie Wydawnictwo Medyczne "Vesalius".

Donesch-Jeżo, E. (2016). „Cross-cultural variability of research article abstracts from different discourse communities”. *Państwo i Społeczeństwo* XVI nr 3, 33-60.

Grabowski, Ł. (2014). "On Lexical Bundles in Polish Patient Information Leaflets: A Corpus-Driven Study". *Studies in Polish Linguistics*, 19 (1): 21–43.

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P. & Suchomel, V. (2014). "The Sketch Engine: ten years on". *Lexicography* 1(1), 7-36.

A handwritten signature in blue ink on a light green background. The signature reads "Julian Grzeszalski" in a cursive script.