

MICHAŁ GRECH

Uniwersytet Wrocławski, Wydział Komunikacji Społecznej i Mediów

michal.grech@uwr.edu.pl

ORCID: 0000-0002-4242-5213

Wartościowanie i wydźwięk emocjonalny w recenzjach awansowych

Valuation and Linguistic Determinants of Value in Promotion Reviews

Abstract

This paper aims to delineate the outcomes of evaluative research and its linguistic manifestations, using the corpus of academic reviews in promotional proceedings as a case study. The analysis encompasses all the reviews from post-doctoral proceedings in the field of linguistics for the years 2012 to 2019, which were obtained from the archives of the Central Commission for Degrees and Titles. The primary dataset consists of 1,225 texts, divided into two sub-corpora: 91 reviews with negative outcomes and 1,134 with positive outcomes. The principal research question addresses the identification of key lexical features that distinguish between positive and negative reviews in promotional assessments within the discipline. The approach adopted is principally comparative corpus analysis, involving a detailed examination of collocations within the texts. The results corroborate the initial hypotheses posited by prior researchers and validate the applicability of the method for analysing a broader spectrum of texts, potentially extending to other academic disciplines. Furthermore, the burgeoning use of semi-automatic and fully automatic techniques for large-scale text corpus analysis and classification appears promising.

Keywords: promotion reviews, post-doctoral reviews, corpus linguistics, evaluation

Słowa kluczowe: recenzje awansowe, recenzje habilitacyjne, korpus językowy, ewaluacja

1. Wstęp i wcześniejsze prace

Zainteresowanie badaniami wartościowania w tekstach nie należy do zagadnień nowych w polskim językoznawstwie. Na przestrzeni ostatnich lat, szczególnie od pierwszych tekstów Jadwigi Puzyniny,

powstało wiele prac podejmujących to zagadnienie z różnych perspektyw. Ich opracowania podjął się m.in. Aleksander Kiklewicz (2017). Rośnie też liczba prac naukowych wykorzystujących automatyczne analizy wartościowania – analizy wydźwięku emocjonalnego do wielu zastosowań praktycznych, szczególnie analiz recenzji produktów (Binxuan, Yanglan, Carley 2018; Song i in. 2019) i usług czy automatycznej analizy wypowiedzi na tzw. platformach społecznościowych, takich jak Twitter (Bahrainian, Dengel 2013).

Celem niniejszego artykułu jest przedstawienie metody i wyników analizy tekstów recenzji awansowych (habilitacyjnych) ze względu na obszary i sposoby wartościowania. Ponadto zaprezentowane zostaną narzędzia i aplikacje użyteczne do ilościowych i automatycznych analiz recenzji habilitacyjnych. Szczególnie istotne jest wskazanie listy słów używanych do nadawania wydźwięku emocjonalnego, co może pozwolić w przyszłości na stworzenie stosownego narzędzia do analiz wydźwięku emocjonalnego w recenzjach naukowych.

W literaturze angielskiej nie udało się znaleźć pracy naukowej, która podejmowałaby podobne problemy. Znalaziono jedynie prace skupione na analizie wydźwięku emocjonalnego w recenzjach artykułów naukowych (Machado, Ribeiro, Oliveira 2019) oraz zgłoszeń na konferencję naukową (Chakraborty, Goyal, Mukherjee 2020) czy recenzji książek naukowych (Hamdan, Bellot, Bechet 2016). Dostępne są też prace przeglądowe, oceniające przydatność różnych podejść i specjalizacji (Muni i in. 2017). Zastanawiając się nad możliwymi przyczynami takiego stanu rzeczy, stwierdzono, że może być ich kilka. Najistotniejszą z nich wydaje się ograniczony zasięg procedury habilitacji, która nie funkcjonuje w modelu anglosaskim. Drugim istotnym czynnikiem wydaje się to, że publiczna dostępność recenzji nie jest standardem. Sprawdzono też, choć nie obejmuje to zakresu pracy, analizy recenzji doktoratów. Niestety także w tym przypadku nie udało się znaleźć żadnych artykułów naukowych.

Jedyne prace zajmujące się problematyką recenzji awansowych (habilitacyjnych), które udało się znaleźć, są napisane przez badaczy z Polski. Pierwsza z nich to praca Ewy Szkudlarek-Śmiechowicz (2015), która wskazała przy pomocy analizy jakościowej to, jakie elementy są oceniane i jakie są wyznaczniki tej oceny. Analizowanym materiałem były także recenzje awansowe – habilitacyjne w naukach humanistycznych. We wspomnianym tekście Szkudlarek-Śmiechowicz wymienia istotne i zasadniczo obligatoryjne elementy konstytucyjne recenzji naukowej awansowej, jak i główne wskazywane elementy i sposoby wartościowania. Ze względu na specyfikę tematu i zakres materiału, wspomniana powyżej praca stanowi główny punkt odniesienia dla niniejszego artykułu. Praca ta jest najbliższa podejmowanej problematyce, ponieważ w oparciu o jakościową analizę recenzji z zakresu humanistyki wskazuje wstępnie zarówno oceniane obszary, jak i językowe środki służące wartościowaniu. Niniejszy tekst stanowi rozwinięcie analiz na gruncie bardziej ilościowym niż jakościowym.

Pozostałe prace, to artykuły Łukasza Remisiewicza (2019, 2023), przy czym w pracy „Niespójność recenzji a wynik postępowania habilitacyjnego. Próba analizy w dyscyplinie socjologia” (Remisiewicz 2019) skoncentrowano się głównie na porównaniu konkluzji recenzji złożonych w postępowaniach habilitacyjnych z efektem postępowania. Choć ciekawym wątkiem w tej pracy jest nawiązanie do prac Sunny Hyon i Rogera Chena (2005) oraz samej Hyon (2008, 2011), którzy zajmowali się recenzjami awansowymi i promocyjnymi (z ang. *retention-promotion-tenure* [RPT]), które ze względu na dalszy przebieg kariery naukowej można w pewnym zakresie uznać za procedurą awansową. Więcej o podobieństwach i różnicach, szczególnie w dostępności, pisze jeszcze Remisiewicz (2019: 112–113).

Ostatnią pracą, dotyczącą częściowo problemu analiz wartościowania w recenzjach awansowych, jest niepublikowana praca doktorska Ewy Rozkosz. Dysertacja stanowi bardzo ciekawą i dość kompleksową próbę spojrzenia na problematykę ewaluacji naukowej na przykładzie habilitacji. Dla podejmowanego w niniejszym tekście problemu wartościowania szczególnie istotne wydają się wyniki analizy jakościowej recenzji z pięciu wybranych dyscyplin (Rozkosz 2017: 227–312). Autorce udało się wyróżnić katalog wątków (jak nazywa je sama autorka) używanych w analizowanych recenzjach, grupując je w czterech głównych tematach: 1) cechy formalne, 2) jakość osiągnięć naukowych, 3) społeczno-etyczny wymiar osiągnięć, 4) punkt odniesienia dla oceny (2017: 229).

Jest też nieco więcej publikacji, które z różnych perspektyw podejmują problematykę recenzji habilitacyjnych, jednak ze względu na cele i ograniczony zakres, nie zostaną one tutaj omówione (por. Szpunar 2020).

2. Wartościowanie

W niniejszej pracy przyjmuję wstępnie rozumienie wartościowania, zaproponowane przez Elżbietę Laskowską (1992: 19), jako określania, że coś jest dobre lub złe pod jakimś względem. Jednocześnie staram się rozszerzyć rozumienie tak, by obejmowało określenia czegoś jako pozytywnego lub negatywnego, docenianego lub deprecjonowanego. W ramach analizy wartościowania w recenzjach nie ograniczam się więc tylko analiz etycznych – tego, co dobre i złe (por. Fleischer 2007). Docelowo staram się uchwycić wszystkie językowe przejawy oznaczania czegoś jako warte czy też wartościowe, lepsze lub gorsze, ładne czy brzydkie. Piszę też o tym Michael Fleischer, który wskazuje, że:

pojęcie (oraz zjawisko) „wartości” skonstruowane jest tak, że posiada charakter predykatu, to znaczy określa pewien aspekt dowolnego w zasadzie obiektu, przypisywalne jest dowolnemu obiektowi itp. Wszystko może posiadać czy wykazywać wartość, jeśli tylko ktoś, temu obiektowi wartość przypisuje. (Fleischer 2010: 5)

Jednym z głównych celów niniejszej pracy jest więc wyróżnienie zarówno tych obszarów i obiektów, którym przypisuje się wartość jak i językowych wyznaczników i kryteriów tego przypisywania (o ile można je wyróżnić).

3. Materiał

Jak wspomniano wcześniej, w niniejszym artykule przedstawione zostały wyniki badań nad wartościowaniem w recenzjach habilitacyjnych. Materiałem poddanym analizie były wszystkie recenzje habilitacyjne w jednej dyscyplinie dostępne na stronie Centralnej Komisji ds. Stopni i Tytułów. Zdecydowano się na wybór dyscypliny, która zajmuje się językiem, z założeniem większej znajomości i świadomości języka oraz jego praktyk pośród recenzentów z dziedziny, a co za tym idzie, zróżnicowania słownictwa i większej różnorodności oraz złożoności tekstów recenzji. To, czy to założenie było poprawne, będzie możliwe dopiero po zaplanowanych w kolejnych krokach badaniach porównawczych. Dodatkowym powodem jest orientacja autora (jako językoznawcy) w zakresie recenzji, co pozwoli na

głębszą analizę tekstu, niż miałyby to miejsce w przypadku wybrania innej dyscypliny. Ze względu na charakter badań, dobór materiału opisano w części metodycznej tekstu razem z metodą.

4. Hipotezy i cele pracy

Przedstawiany tekst jest drugim, podejmującym problematykę wartościowania w recenzjach naukowych. W pierwszym tekście dokonano wstępnej identyfikacji słownictwa używanego do pozytywnej i negatywnej oceny wniosków habilitacyjnych, także wstępnie wskazano obszary oceniane (Grech w druku na 2024). W przypadku pierwszego badania materiał empiryczny był bardziej ograniczony: 92 recenzje pozytywne i 91 recenzji negatywnych. Metodą przyjętą w poprzedniej analizie była identyfikacja słów kluczowych (*keywords*) w obu korpusach poprzez porównanie ich wzajemnie na swoim tle: pozytywnych na tle negatywnych i odwrotnie. W wyniku analiz uzyskano listy słów charakterystycznych dla obu badanych typów recenzji. Kolejnym istotnym ustaleniem poprzednich analiz była znacząca różnica między średnią wielkością recenzji negatywnych (5 478 słów) i pozytywnych (2 406 słów).

W niniejszym badaniu zdecydowano się na sprawdzenie, czy wyniki uzyskane w pierwszym badaniu zostaną też w mocy na materiale wszystkich recenzji z dyscypliny – ponad sześć razy bardziej obszernym niż w poprzednim badaniu.

5. Metoda i przygotowanie materiału

Z uwagi na to, że problem badawczy i materiał są rzadko badane, a przyjęta metoda nie była dotychczas używana, to zdecydowano się na bardziej szczegółowe przedstawienie poszczególnych etapów pracy. Jako metodę przyjęto analizę porównawczą korpusów tekstowych.

Materiał stanowiły recenzje z postępowań awansowych z dziedziny nauk humanistycznych w dyscyplinie językoznawstwo, pobrane ze strony Centralnej Komisji ds. Stopni i Tytułów (<https://www.ck.gov.pl>). Szczegółowo procedurę tę opisano w poprzednim tekście dotyczącym tego tematu (Grech w druku na 2024).

Ostatecznie uzyskano 1225 poprawnie rozpoznanych recenzji z 413 spraw (katalogów). Ponieważ pobrane recenzje są w większości zapisane w formacie skanów (obrazów) „zaszytych” w formacie PDF, to konieczne było konwertowanie wszystkich plików do formatu czystego tekstu. Ta baza była podstawą do dalszej pracy. Po przetestowaniu kilku narzędzi, zdecydowano się na skorzystanie ze skryptu i bibliotek w języku Python. Kolejne kroki podjęte w ramach przygotowania materiału to: czyszczenie danych z drobnych błędów powstałych przy rozpoznawaniu tekstu i lematyzacja przy użyciu pakietu SpaCy w języku Python. Tak przygotowane dane poddano głównej analizie.

6. Analiza główna

W pierwszym etapie oddzielono recenzje negatywne od pozytywnych. Do wstępnego wskazania recenzji z negatywną konkluzją postanowiono posłużyć się istniejącymi narzędziami i programami oraz przede

wszystkim wstępną wiedzą o schematach konstrukcyjnych tych tekstów – szczególnie konkluzji. W pierwszym etapie wypróbowano różne dostępne narzędzia, o czym więcej napisano w poprzednim tekście (Grech 2024 w druku). Ostatecznie oparto się o pakiet Orange (Demsar, Janez, Curk, Tomaz, Erjavec Ales i in. 2013), który trafnie wskazał pierwszych 30 recenzji negatywnych (spośród wszystkich 1225). Niewystarczalność istniejących narzędzi do oceny wydźwięku emocjonalnego wskazuje na specyfikę recenzji awansowych i konieczność opracowania odpowiednich narzędzi dla recenzji naukowych.

Kolejne recenzje negatywne wyróżniono przy pomocy wyszukiwania kluczowych fraz wskazujących na konkluzję negatywną, takich jak: *nie spełnia wymogów*, *nie spełnia kryteriów*, a także: *nienadawanie*, *niedopuszczenie*, *odmowa*, *nie mogę poprzeć wniosku* itp. Ze świadomością możliwej niekompletności tego wyszukiwania zgodzono się i przyjęto jednocześnie założenie, że ewentualne pozostałe recenzje negatywne zostaną wyróżnione w kolejnych krokach, po zakończeniu analiz zawartości recenzji – będą więc swoistym testem poprawności i analiz.

Metoda ta, choć niedoskonała, pozwoliła wyróżnić w sumie 91 tekstów recenzji kończących się negatywną konkluzją, które zostały użyte do dalszych analiz (tab. 1). Pozostałe 1134 teksty to recenzje pozytywne. Jak można zobaczyć, teksty recenzji negatywnych są średnio ponad 1,5 razy dłuższe. Widoczne jest, że w badanym materiale pozytywna opinia wymaga mniej wyjaśnień niż negatywna, którą trzeba bardziej rozwinąć.

Tab. 1. Ogólne miary korpusu tekstów użytych do analizy. Źródło: badania własne.

Charakterystyki korpusów	Korpus recenzji negatywnych (91)	Korpus recenzji pozytywnych (1134)	Cały korpus (1125)	Proporcje p/n
Liczba tekstów	91	1134	1225	12,4
Types*	37 850	144 500	158 803	3,81
Tokens*	422 380	3 419 753	3 842 133	8,09
Średnia liczba słów na tekst	4 591	3 015	3 133,9	1,52

*z liczbami i słowami błędnie rozpoznanymi

7. Porównanie korpusów – listy słów wartościujących – metoda i wyniki

Kolejnym (drugim) etapem było wyróżnienie list słów charakterystycznych dla obu korpusów, ze szczególnym uwzględnieniem słownictwa używanego do wartościowania.

We wcześniejszych analizach wykorzystano do tego celu identyfikację słów kluczowych (*keywords*) w odniesieniu do korpusów: a) języka polskiego i b) wzajemnie do recenzji pozytywnych/negatywnych. Szczegółowa metoda i wyniki zostały opisane w tekście (Grech 2024, w druku). W niniejszej analizie zastosowano nieco odmienne podejście – porównano częstości i proporcje występowania kolejnych słów w obu korpusach.

Procedura dla obu podkorpusów (recenzji pozytywnych i negatywnych) była identyczna. Poniżej wymieniono w punktach główne kroki, które szczegółowo wyjaśniono w dalszej części tekstu.

1. Uzyskano listy częstości słów dla każdego z podkorpusów recenzji (pozytywnych i negatywnych) oraz zsumowano ich wartości, tworząc reprezentację całego korpusu.
2. Przy pomocy wartości średniej i odchylenia standardowego oznaczono górny próg istotnej liczebności do analizy: dla każdego podkorpusu i dla sumy łącznie 1200 najczęściej występujących wyrazów.
3. Obliczono proporcję występowania: pozytywne/negatywne i negatywne/pozytywne oraz dokonano jej normalizacji, uwzględniając dysproporcje w wielkości korpusów równą 8,1. Zapewnia to porównywalność uzyskanych wyników dla korpusów o różnych liczebnościach. Jeśli w korpusie pozytywnym wyraz nie wystąpił ani razu, dzielono przez 1, by uniknąć straty potencjalnie ważnych wyrażen. Tak uzyskano współczynnik negatywny (P_n) dla każdego z wyrazów.
4. Oznaczono wyrazy wyraźnie częściej występujące w każdym z korpusów – na podstawie wcześniejszego kroku, wartości średniej oraz odchylenia standardowego.
5. Usunięto z analizy wyrazy częste, imiona, nazwiska i nazwy własne.
6. Oznaczono wyrazy używane głównie w recenzjach negatywnych i pozytywnych – na podstawie proporcji z pkt. 4.
7. Pozostałe wyrazy skategoryzowano, mając na uwadze szczególnie kategorie wskazujące na to, czego dotyczy wartościowanie i w jaki sposób jest realizowane.
8. Oznaczono wyrazy o nacechowaniu emocjonalnym potrzebne do dalszego wykorzystania w potencjalnej analizie wydźwięku emocjonalnego (*entymment analysis*).
9. Oznaczono wyrazy definicyjnie wartościujące.

Mając na uwadze specyfikę wartościowania (pozytywne i negatywne), poniżej zaprezentowano wyniki analiz w podziale na oba korpusy.

7.1. Porównanie korpusów: recenzje negatywne na tle pozytywnych

Pierwszego porównania dokonano z zamiarem wyłonienia istotnych słów dla korpusu recenzji z negatywną konkluzją. Jak wspomniano w poprzednim punkcie artykułu, wyróżnienie list słów i charakterystycznych wyrażen w podkorpusie recenzji negatywnych odbyło się na podstawie porównania częstości występowania z podkorpusem recenzji pozytywnych.

Zdecydowano się na analizę uzyskanych wyników równoległe z kilku perspektyw:

1. istotność wyrazów ze względu na częstość w korpusach,
2. istotność wyrazów ze względu na proporcje występowania słów w korpusach recenzji negatywnych i pozytywnych,

a także ze względu na:

3. wartościowania wyrażone wprost,
4. wyrażenia emocjonalne,
5. kategorie tematyczne.

Dalsza część artykułu zawiera wyniki przedstawione i sortowane według proporcji (P_n i P_p), jako głównej kategorii porównawczej.

Prezentację wyników rozpoczęto od słów istotnych ze względu na częstość występowania w korpusie, choć niekoniecznie charakterystycznych tylko dla tego korpusu. Najczęstsze wyrazy z korpusu recenzji negatywnych nie są specyficzne dla tego korpusu, jak pokazuje analiza, słownictwo często dotyczy obszarów, które podlegają ocenie (praca, tekst, artykuł, monografia), albo o których często się wspomina: tytuły i stopnie naukowe, role osób (autorka, habilitantka). Jak pokazuje dalsza analiza, wśród najczęstszych słów nie ma wyrazów bardziej charakterystycznych dla tego podkorpusu. Stanowią one znaczną część konstrukcji teksów, jednak są wspólne dla różnych recenzji. Nie mogą więc stanowić istotnej podstawy do odróżnienia recenzji negatywnych od pozytywnych (i na odwrót). Wobec tego istotne jest, by w kolejnym kroku wskazać te słowa, które będą specyficzne dla recenzji negatywnych oraz wprost wartościujące.

W pierwszym kroku wybrano te słowa z całego subkorpusu recenzji negatywnych, które mieściły się pomiędzy wartością mediany (1,0) a sumy średniej i odchylenia standardowego (10,04). W drugim – te, które przekraczały zsumowaną wartość mediany i odchylenia standardowego (powyżej 8). Dopiero takie, mniej restrykcyjne, użycie tych miar pozwoliło na uzyskanie sensownych wyników, gdyż oparcie się tylko na sumach średniej (lub mediany) i kolejnych wartości odchylenia standardowego dawało zbyt rozczłonkowane i analitycznie nieprzydatne wyniki. W sumie udało się wyróżnić 69 wyrazów wartościujących negatywnie i charakteryzujących się jednocześnie znacząco częstszym użyciem w recenzjach zakończonych negatywną konkluzją.

W tabeli 2 przedstawiono zestawienie słów ze względu na istotność – wysokość współczynnika Pn $\text{Proporcja (Pn)} = \frac{N_{\text{pozytywnych}}}{N_{\text{negatywnych}}} / 8,1$ (jak na stronie 85). Kryterium decydującym o umieszczeniu wyrazu jak „negatywne wartościowanie” było nie jego słownikowe czy uznaniowe znaczenie, ale istotnie częstsze proporcjonalne występowanie właśnie w zbiorze recenzji negatywnych – zsumowane wartości mediany i odchylenia standardowego dla proporcji liczebności.

Tab. 2. Kategorie słów o mniejszej wartości współczynnika Pn (1,0-10,98). Źródło: badania własne.

Nr	Kategorie słów o mniejszym współczynniku Pn (2,08-2,98)	Liczba słów	Suma z N negatywne	Słowa
1	negacja	5	6156	nic, nie, brak, bez, negatywny, żaden
2	błąd	2	358	błąd, pomijać
3	zaledwie/niemal	2	91	niemal, zaledwie
4	sposób działania	1	86	poważny
5	emocja	1	180	niestety
	suma	9	6691	

Jeśli przyjrzeć się bliżej zawartości tabeli 2, można stwierdzić, że słowa te są bardziej ogólne i mniej nacechowane emocjonalnie niż słowa przedstawione w tabeli 3.

Tab. 3. Kategorie słów o większej wartości współczynnika Pn (11–81). Źródło: badania własne.

Nr	Kategorie słów o większym współczynniku Pn (8-81)	Liczba słów (N)	Suma z N negatywne	Słowa
1	emocje	9	138	kłopotliwy, lęk, przykry, przykrość, rozczarowywać, straszny, załamywać, zdenerwowanie, zdumienie
2	plagiat/kopiowanie	6	47	autoplagiat, nieoznaczony, plagiat, plagiatowanie, przekopiować, skopiować
3	negacja sensu	3	42	bezsens, niejasny, nieprzekonujący, nonsens
4	zniesztalcenie/ przekłamanie	7	37	dwuznaczność, nieprawda, nieuczciwość, nieuczciwy, oczerniać, okłamywać, zniesztalić
5	inne	7	32	czepiać, dręczyć, kosztować, podwyższać, rozwodzić, uwierzyć, wzmocnić
6	pejoratywne	4	18	prymityw, prymitywny, upośledzony, złodziejski
7	metafory	2	18	przepaść, urodzaj
8	konkluzja	3	14	niedopuszczenie, nienadanie, nienadawanie
9	inne negacja	4	11	negatywy, nieakceptowany, niedobry, niewnoszący
10	sposób działania	2	10	anarchistyczny, nonszalancki
11	nieodpowiednie	3	8	nierелеwantny, niewspółmierny, osobliwy
12	nienormalność	2	8	nienormalność, nienaturalnie
13	incydentalnie / uporczywie	2	4	incydentalnie, uporczywie
14	bezzadność	2	4	bezzadność, nieporadnie
15	inne: oryginał	1	97	oryginał
	SUMA	57	488	

W tabeli 3 widoczne jest zestawienie kategorii słów o wysokiej wartości współczynnika Pn, ułożone według liczby wystąpień słów w korpusie recenzji negatywnych.

1. Najliczniejsza kategoria: *emocje* gromadzi różne wyrażenia określające stany emocjonalne w zasadzie tylko negatywne. Uznaję to za bardzo ciekawe i znaczące, że w recenzjach pracy naukowej pojawiają się wyrażenia określające nie tylko rozczarowanie, ale też żal, gniew, smutek, strach, irytację, oburzenie czy pogardę. Wydaje się, że konieczne jest zbadanie funkcji nie tylko emotywnych, ale i retorycznych wyrażania, szczególnie tak skrajnych emocji negatywnych. Temat ten zostanie jeszcze rozwinięty w następnej części pracy, przy omawianiu wydźwięku emocjonalnego. Na tym etapie pracy zdecydowano się na ograniczenie prezentacji i omówienia wyników do wskazania wyrażen o charakterze wartościowań. Ich celowość, etyczność i dopuszczalność będą przedmiotem osobnych analiz i tekstów.

2. Druga istotna kategoria zawiera wyrazy i wyrażenia, które w dyskursie naukowym stanowią oskarżenie, orzeczenie o olbrzymiej wadze.
3. Ciekawym zabiegiem wydaje się też odmawianie *sensu*, *jasności* czy *zrozumiałości* ocenianym pracom – to z pewnością kolejny bardzo mocny zabieg retoryczny, który niekoniecznie musi świadczyć o słabościach pracy, a np. o nieodpowiednim doborze recenzentów.
4. Kolejna kategoria wskazuje na szeroki zakres zastrzeżeń: od *dwuznaczności*, przez *zniekształcenia*, po celowe *oszustwo*, *okłamywanie* czy podawanie *nieprawdy*.

Ze względu na ograniczenia wielkości artykułu kolejne kategorie nie będą już omawiane, gdyż zostały szczegółowo przedstawione w tabeli 3. Jednak wydaje się istotne, żeby wskazać, że rozszerzają one katalog zabiegów recenzentów o dość pejoratywne epitety, metafory, odniesienia do (nie)*normalności* i (nie)*naturalności*. Widoczne są także wspomniane w części metodycznej słowa klucze pojawiające się w konkluzjach recenzji (*niedopuszczenie*, *nienadawanie*, *etc.*).

7.2. Wydźwięk emocjonalny

W trzecim etapie pracy skoncentrowano się na zidentyfikowaniu wyrażen wydźwięku emocjonalnego. Podczas wstępnej analizy wydźwięku emocjonalnego recenzji negatywnych istniejącymi narzędziami udało się poprawnie zidentyfikować tylko około 30% z wszystkich recenzji negatywnych. Pozostałe recenzje były oznaczane podobnie jak teksty o konkluzji pozytywnej. Taki stan rzeczy wynika ze specyfiki recenzji naukowej, szczególnie pozytywnej, gdzie jest mniej czy wręcz brak miejsca na wyrażanie ocen w sposób mocny i nacechowany nastawieniem. Dlatego też w tej części pracy podjęto próbę (re)konstrukcji słownika wyrażen o znaczącym wydźwięku emocjonalnym, co być może pozwoli w przyszłości na stworzenie stosownego narzędzia do analiz wydźwięku emocjonalnego w recenzjach naukowych.

Zdecydowano się na umieszczenie tej części jako rozwinięcie analiz recenzji negatywnych, ponieważ tylko takie nacechowanie udało się znaleźć. W tabelach 4.1 i 4.2 przedstawiono wybrane słowa charakteryzujące się mocnym wydźwiękiem. Dotyczą dość szerokiego zakresu określeń i odniesień do emocji (negatywnych). Oznaczeń dokonano według modelu NRC (Saif, Mohammad, Turney, Peter 2013), by zapewnić przejrzystość i powtarzalność wyników podczas rozszerzania badań.

7.2.1. Wydźwięk emocjonalny – negatywny

Przegląd tabeli wskazuje na najczęstsze słowo: *lęk*, które po sprawdzeniu kolokacji okazuje się tematem jednej z prac i nie będzie w tej analizie liczony jako wyraz emocji. Jednak pozostałe wyróżnione słowa nie mają już takiego uwikłania, więc można je wstępnie potraktować jako wykładniki wydźwięku emocjonalnego. I tak kolejne często napotykanne słowo: *przykrość* zwykle używane jest w zwrocie: „z (ogromną) przykrością stwierdzam” lub też: „z (ogromną) przykrością muszę stwierdzić”.

Tab. 4.1. Słowa o negatywnym wydźwięku emocjonalnym. Źródło: badania własne.

Nr	Wydźwięk emocjonalny	Słowo (lem)	N.w rec. negat.	Proporcja neg./poz. znormaliz.	N w rec. pozyt.	Wybrane przykłady
1	irytacja, gniew	zdenerwowanie	2	16,20	1	Dotyczy tematu pracy*

Nr	Wydźwięk emocjonalny	Słowo (lem)	N.w rec. negat.	Proporcja neg./poz. znormaliz.	N w rec. pozyt.	Wybrane przykłady
2	gniew	nieuczciwość	5	40,50	1	o nie tylko nieuczciwość, ale również niechlujstwo
3	gniew	nieuczciwy	3	24,30	0	Nieuczciwe korzystanie ze źródła
4	gniew	nieprawdy	2	16,20	1	Nieprawdziwe też są informacje o (...)
5	gniew	okłamywać	2	16,20	1	Dotyczy tematu pracy
6	irytacja	uporczywie	2	16,20	1	uporczywe zaś wyliczanie udziałów
7	irytacja	incydentalnie	2	16,20	1	pojawiają się incydentalnie
8	irytacja	dręczyć	2	16,20	1	Dotyczy tematu pracy
9	irytacja	czepiać	2	16,20	1	Dotyczy tematu pracy
10	oburzenie, pogarda	prymitywny	2	16,20	1	Ćwiczenia na s. 17-20 są dość prymitywne
11	rozczarowanie	rozczarować	12	3,89	1	mimo intrygującego tytułu przynosi rozczarowanie
12	rozczarowanie	bezsens	4	32,40	1	Bezsensowny jest również zapis typu
13	smutek	przykre	2	16,20	1	Szczególnie przykre są przypadki nieprecyzyjnej interpretacji
14	smutek	przykrość	33	12,15	22	stwierdzam z przykrością
15	strach	straszny	3	24,30	1	Dotyczy tematu pracy
16	strach	załamywać	2	16,20	0	załamuje się szybciej niż chciałby tego Autor
17	zaskoczenie	kłopotliwa	5	40,50	0	kłopotliwe, jeśli nic niemożliwe
18	zaskoczenie	zdumienie	7	28,35	2	Zdumienie wywołuje pomieszanie różnych podejść
19	żał	bezzadność	2	16,20	1	Językoznawca staje bezzadny
20	żał	nieporadnie	2	16,20	1	definicje znaczeniowe często brzmią nieporadnie

*Jeśli przykłady jednoznacznie wskazują, że dotyczą one tematy pracy, a nie sposobu wyrażenia wartościowania, to zaznaczono to poprzez adnotację: „Dotyczy tematu pracy”.

W zestawieniu wydźwięku emocjonalnego (tabela 4.1), przeważają emocje takie jak rozczarowanie (33%), smutek i żal (27% +2%), uzupełniane przez: gniew (11%), irytację (8%), oburzenie/pogardę

(4%) oraz zaskoczenie, które – jak wskazują proporcje – jest prawie zawsze negatywne (9%). Wydaje się więc, że recenzenci oczekują od prac, a oczekiwania te nie zostają spełnione, co w następstwie generuje repertuar odczuć określanych jako negatywne.

7.2.2. Wydźwięk emocjonalny – pozytywny

Zakres słów o wydźwięku emocjonalnym pozytywnym w recenzjach z konkluzją negatywną jest znacznie mniejszy, bo dotyczy jedynie około 6 wyrazów przedstawionych w tabeli 4.2, z czego tylko jedno określenie dotyczy wprost nazwy emocji (*radość*), a pozostałe wskazują na podziw wobec *erudycji* i podejmowania *niełatwych* problemów i tematów oraz ich *wzorcową* realizację.

Słowo *niezwykle* jest kolejnym z grupy słów, w przypadku którego szczegółowa analiza ich konkordancji i kolokacji pozwala dość łatwo zobaczyć, co jest szczególnie doceniane przez recenzentów. Są to między innymi: *cenne źródła* (materiały, wyniki), *drobiazgowa* analiza, *interesujące* podejście do problemów, *pracochłonne* zadanie, *pracowitość*, *rzetelność*, *skrupulatność*, itd.

Tab. 4.2. Słowa o pozytywnym wydźwięku emocjonalnym. Źródło: badania własne.

Nr	słowo (lem)	N razem	N negatywne	N pozytywne	neg/poz znormalizowane	poz/neg znormalizowane
1	niezwykle	1019	60	959	0,51	1,97
2	erudycja	171	4	167	0,19	5,15
3	przyjemność	96	4	92	0,35	2,84
4	niełatwy	76	1	75	0,11	9,26
5	wzorcowy	87	8	79	0,82	1,22
6	radość	37	2	35	0,46	2,16

7.3. Porównanie korpusów: recenzje pozytywne na tle negatywnych

Podobną procedurę przeliczenia i normalizacji, jak w przypadku analiz słów kluczowych negatywnych, zastosowano też w przypadku odwrotnym – gdy szukano kluczowych słów pozytywnych:

$$\text{Proporcja (Pp)} = N_{\text{negatywnych}} / N_{\text{pozytywnych}} / 8,1$$

Także w tym przypadku otrzymaną listę oczyszczono z wyrażen w oczywisty sposób niewartościujących, jak: nazwiska, nazwy własne, wyrazy krótkie częste.

Tab. 5. Najczęstsze wyrazy z korpusu pozytywnego – fragment tabeli. Źródło: badania własne.

Nr	słowo (lem)	N razem	N negatywne	N pozytywne	Neg/poz znormaliz.	Poz/neg znormaliz.
1	naukowy	28790	2294	26496	0,70	1,43
2	język	26564	2304	24260	0,77	1,30

Nr	słowo (lem)	N razem	N negatywne	N pozytywne	Neg/poz znormaliz.	Poz/neg znormaliz.
3	doktor	20986	1612	19374	0,67	1,48
4	nie	22901	4681	18220	2,08	0,48
5	praca	19340	2155	17185	1,02	0,98
6	rok	16902	1510	15392	0,79	1,26
7	polski	15638	1235	14403	0,69	1,44
8	badanie	14112	1184	12928	0,74	1,35
9	jako	13859	1511	12348	0,99	1,01
10	habilitantka	13462	1271	12191	0,84	1,18
11	przez	13355	1345	12010	0,91	1,10
12	tekst	12429	1483	10946	1,10	0,91
13	artykuł	11405	1324	10081	1,06	0,94

Tabela 5 przedstawia najczęstsze wyrazy korpusu recenzji pozytywnych. Dotyczą głównych obszarów tematycznych recenzji: badań, prac naukowych, szczególnie tekstów, odniesień do języka i językoznawstwa, tytułów i stopni naukowych. O ile dość dobrze wskazują na główne tematy, o tyle trudno się tu doszukać samych słów służących wyrażeniu nacechowania i oceny.

W tabeli 6 pokazano słowa wybrane według kryterium najwyższego poziomu współczynnika Pp – obserwacje odstające (powyżej sumy średniej i 3 odchyłeń standardowych). Zawiera ona zarówno słowa wyrażające nienaturalne cechy i oceny, jak i te, które wskazują, co jest znacząco częściej pozytywnie oceniane: recenzowany, organizatorski, kontakt(y), autoprezentacja, staż(e).

Gdy sprawdzono najczęstsze konteksty wyróżnionych słów wartościujących, to zauważono, że można wskazać zarówno wyrażenia służące wprost wartościowaniu, takie jak to, że osoba *umiejętnie* prowadzi analizę, polemikę, czy *syntetyzuje* wiedzę itp., czy też *wnikliwość* i *zaawansowanie* badaczy i badaczek. Można też wskazać słowa określające recenzowane osiągnięcia (*wielopłaszczyznowy*), określające osoby jako: *ukształtowanych* badaczy, którzy często podejmują *niełatwe* problemy i tematy, poziom *zawansowania* badaczy, ich *nieustanny* rozwój czy *niezaprzeczalne* wartości i walory recenzowanych prac. Oceniane są też osiągnięcia organizatorskie, *rozwijanie* kontaktów, *wielowymiarowa* aktywność i działalność – a więc to, co jest wskazane jako warunki dla habilitacji. Doceniane są także: *wielowymiarowe* ujęcie tematów badawczych. Są też wyrażenia związane z tematyką prac czy książek, takie jak *autoprezentacja*, oraz używane na wiele sposobów, jak kontakt.

Lista oparta na wysokiej wartości współczynników (Pn i Pp) może być dobrą podstawą do tworzenia „słownika wartościowania” w recenzjach naukowych. Jednak oparcie się tylko na współczynniku może wprowadzić pewne nieścisłości, takie jak wyrazy będące wysoko na liście rangowej nie ze względu na ich funkcję oceniającą, ale np. ze względu na rzadkie pojawienie się w tekstach tylko jednej grupy recenzji.

Tab. 6. Zestawienie wyrazów o największym współczynniku Pp (powyżej 5,8*). Źródło: badania własne.

Nr	Słowo (Iem)	N razem	N negatywne	N pozytywne	Neg/poz znormaliz.	Poz/neg znormaliz.
1	umiejętnie	218	1	217	0,04	26,79
2	zróżnicowany	101	1	100	0,08	12,35
3	wnikliwość	94	1	93	0,09	11,48
4	umniejszać	82	1	81	0,10	10,00
5	wielopłaszczyznowy	79	1	78	0,10	9,63
6	recenzowany	78	1	77	0,10	9,51
7	ukształtowany	76	1	75	0,11	9,26
8	niełatwy	76	1	75	0,11	9,26
9	organizatorski	75	1	74	0,11	9,14
10	kontakt	848	13	835	0,13	7,93
11	zaawansowanie	61	1	60	0,13	7,41
12	zrozumiałość	60	1	59	0,14	7,28
13	nieustannie	60	1	59	0,14	7,28
14	niezaprzeczalny	58	0	58	0,00	7,16
15	autoprezentacja	57	0	57	0,00	7,04
16	wielowymiarowy	55	1	54	0,15	6,67
17	wariancja	53	0	53	0,00	6,54
18	uhonorować	53	1	52	0,16	6,42
19	staż	53	1	52	0,16	6,42

* powyżej średniej +3 x odchylenie standardowe

Podsumowując wyniki, można wskazać kilka prawidłowości:

1. W ocenianych osiągnięciach docenia się przede wszystkim ich *wielkość, zróżnicowanie, kompleksowość, wielopłaszczyznowość i wielowymiarowość*, a także podejmowanie *niełatwych problemów i przedstawianie ich w zrozumiały sposób*.
2. Inną formą pozytywnej oceny, która jest stosowana zarówno wobec osób, jak i pracy, jest wskazywanie różnych obszarów *zasługujących na docenienie, wyróżnienie, uznanie czy wysoką ocenę*. Są to zarówno *aktywność habilitantów (naukowa, organizacyjna, dydaktyczna)*, jak i *organizacja tekstu książki, dobór materiału i literatury, działalność itp.*
3. Cechy, którymi określone są osoby, wskazują przede wszystkim na: *umiejętności w różnych obszarach takich jak dobór literatury, analiza, wybór przykładów, korzystanie z dorobku dyscypliny itp.* I związane z nimi *docieklivość i wnikliwość w badaniach i uzupełniającą erudycja habilitantów*.

4. Ceniona jest także *samoświadomość* badawcza i *nieustanny rozwój*.
5. Natomiast słowa takie jak: *biegłość*, *prestiz* i (tylko częściowo) *aprobata* dotyczą obszaru konkretnych recenzowanych badań i książek.
6. Językowym zabiegiem grzecznościowym, który pozwala recenzentom na wtrącenie uwag i sugestii, jest użycie wyrażen składających się z wyrażenia „nie umniejszając” i dodanie np. *wartości pracy* czy *znaczenia osiągnięć*.

Podsumowując tę część analizy, można stwierdzić, że w przypadku recenzji z konkluzją pozytywną mamy do czynienia z częstym użyciem słów mocno wartościujących pozytywnie przy jednocześnie niewielkiej liczbie określeń pozytywnego wydźwięku emocjonalnego.

Wnioski

W niniejszym tekście przedstawiono wyniki analiz wartościowań w korpusie wszystkich recenzji awansowych w dyscyplinie językoznawstwo. Jak pokazano, zaproponowana i zastosowana metoda daje oczekiwane wyniki – pozwala na wyróżnienie zarówno słownictwa wspólnego dla obu podkorpusów, jak i charakterystycznego. Wskazano obszary podlegające wartościowaniu, które są zgodne z ustaleniami z wcześniejszych badań jakościowych (Szkudlarek-Śmiechowicz 2015). Dotyczą one głównie pracy, szczególnie tekstów naukowych, działalności: naukowej, dydaktycznej i organizacyjnej, a także erudycji i biegłości w różnych obszarach związanych z pracą naukową.

Zaproponowano współczynniki: pozytywny i negatywny (Pp i Pn), które pozwalają także na względnie proste wyróżnienie słów używanych wprost do wartościowania, których użycie przynosi satysfakcjonujące rezultaty.

Udało się także wyróżnić charakterystyczne słowa-operatory, które używane są do wprowadzania wartościowania, takie jak: *zasługujący*, *nie umniejszając*.

W przypadku recenzji zakończonych negatywną konkluzją stanowczo mniej mówi się o osobach, a więcej o niejasnościach, błędach, pomyłkach, rozbieżności w oczekiwaniach recenzentów, a tym, co i jak napisane znajduje w przedstawionych dokumentach habilitacyjnych. Znacząco częściej używa się też wyrazów o nacechowaniu wartościującym i emocjonalnym. Szczegółowo wskazuje się miejsca, szczególnie akapity, w których dochodzi do rozbieżności, co w znacznej mierze wyjaśnia to, że recenzje negatywne są średnio dwukrotnie dłuższe.

W recenzjach zakończonych pozytywną konkluzją pisze się ogólnie mniej, mniej wartościuje i stosunkowo rzadko używa słów odnoszących się do emocji (*radość*, *przyjemność*).

Wszystkie wskazane różnice między badanymi recenzjami pozwoliły na osiągnięcie głównych celów pracy: wyróżnienie zarówno tego, co jest oceniane i w jaki sposób określane. Ustalenia te będą pomocne do dalszej pracy nad przygotowaniem analiz wartościowania i wydźwięku emocjonalnego dla recenzji prac awansowych.

Ciekawe wyniki może dać analiza przy pomocy algorytmów samouczących (*machine learning*), która jest planowana w przyszłości.

Na koniec konieczne jest jeszcze podkreślenie ograniczeń przedstawianego badania, do których zaliczyć trzeba przede wszystkim ograniczony zakres dziedzinowy – wyniki analizy innych dyscyplin nie muszą się pokrywać, szczególnie w obszarze zróżnicowania językowego. Do ograniczeń metody należy

zaliczyć przede wszystkim niedoskonałe (jeszcze) algorytmy rozpoznawania tekstu, jednak w miarę poprawy tychże, a także wdrożenia metod korekty, wyniki będą znacząco lepsze. Niezbędne jest także przynajmniej częściowe sprawdzanie kontekstów słów, które są wartościujące i spełniają formalne wymogi do uznania za takie, ponieważ czasem analiza kolokacji wskazuje, że ich użycia związane są wprost z tematyką prac habilitacyjnych i nie służą do wartościowania.

Bibliografia

- Bahrainian, Seyed-Ali, Andreas Dengel (2013) „Sentiment Analysis and Summarization of Twitter Data”. [W:] *2013 IEEE 16th International Conference on Computational Science and Engineering*. Sydney: NSW; 227–234.
- Binxuan, Huang, Ou Yanglan, Kathleen M. Carley (2018) „Aspect Level Sentiment Classification with Attention-over-Attention Neural Networks”. [W:] *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*. Cham: Springer; 197–206.
- Chakraborty, Souvic, Goyal Pawan, Mukherjee Animesh (2020) „Aspect-Based Sentiment Analysis of Scientific Reviews”. [W:] *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2020*; 207–216. <https://doi.org/10.1145/3383583.3398541> (dostęp: 26.04.2024).
- Chen, Rong, Hyon Sunny (2005) „Faculty Evaluation as a Genre System: Negotiating intertextuality and interpersonality”. [W:] *Journal of Applied Linguistics*. Vol. 2 (2); 153–184, <https://doi.org/10.1558/jal.v2i2> (26.04.2024).
- Demšar, Janez, Tomaž Curk, Aleš Erjavec, Črt Gorup, Tomaž Hočevar, Mitar Milutinović, Martin Možina, Matija Polajnar, Marko Toplak, Anže Starič, Miha Štajdohar, Lan Umek, Lan Žagar, Jure Žbontar, Marinka Žitnik, Blaž Zupan (2013) „Orange: Data Mining Toolbox in Python”. [W:] *Journal of Machine Learning Research*. Vol. 14; 2349–2353.
- Fleischer, Michael (2010) *Wartości w wymiarze komunikacyjnym*. Łódź: Primum Verbum.
- Grech, Michał (2024) „Przydatność analizy korpusowej do badania wartościowania w recenzjach awansowych”. [W:] *Acta Universitatis Lodzensis. Folia Linguistica*. T. 58/2; 323–338.
- Hyon, Sunny (2008) „Convention and Inventiveness in an Occluded Academic Genre: A Case Study of Retention–Promotion–Tenurereports”. [In:] *English for Specific Purposes*. Vol. 27 (2); 175–192. <https://doi.org/10.1016/j.esp.2007.07.003> (dostęp: 20.04.2024).
- Hyon, Sunny (2011) „Evaluation in Tenure and Promotion Letters: Constructing Faculty as Communicators, Stars, and Workers”. [W:] *Applied Linguistics*. Vol. 32/4; 389–407, <https://doi.org/10.1093/applin/amr003> (dostęp: 16.03.2024).
- Kiklewicz, Aleksander (2017) „Lingwistyczne aspekty kategorii wartości”. [W:] Aleksander Kiklewicz, Jolanta Piwowar (red.) *Wartości i wartościowanie we współczesnej humanistyce. III: perspektywa językoznawcza*. Olsztyn: Centrum Badań Europy Wschodniej UWM; 33–72.
- Laskowska, Elżbieta (1992) *Wartościowanie w języku potocznym*. Bydgoszcz: Wydawnictwo Uczelniane WSP w Bydgoszczy.
- Munir, Ahmad, Shabib Aftab, Iftikhar Ali, Noureen Hameed (2017) „Hybrid Tools and Techniques for Sentiment Analysis: A Review”. [W:] *International Journal of Multidisciplinary Sciences and Engineering*. Vol. 8/4; 28–33.
- Remisiewicz, Łukasz (2019) „Niespójność recenzji a wynik postępowania habilitacyjnego. Próba analizy w dyscyplinie socjologia”. [W:] *Nauka*. Vol. 4; 109–127.

- Remisiewicz, Łukasz (2023) „Selekcja osób w postępowaniach habilitacyjnych w polskiej socjologii”. [W:] *Nauka*. Vol. 3; 127–144.
- Rozkosz, Ewa (2018) „Ewaluacja osiągnięć naukowych w postępowaniach habilitacyjnych”. Rozprawa doktorska. Toruń: Uniwersytet Mikołaja Kopernika w Toruniu.
- Saif, Mohammad, Peter D. Turney (2013) „NRC Emotion Lexicon”. [W:] *Journal National Research Council*. Ottawa: National Academies Press.
- Song, Youwei, Jiahai Wang, Tao Jiang, Zhiyue Liu, Yanghui Rao (2019) „Targeted Sentiment Classification with Attentional Encoder Network”. [W:] *Lecture Notes in Computer Science*. Vol. 11730; 1073–1080.
- Szkudlarek-Śmiechowicz, Ewa (2015) „Wartościowanie w recenzjach naukowych (na podstawie recenzji w postępowaniu o nadanie stopnia doktora habilitowanego)”. [W:] Małgorzata Karwatowska, Robert Litwiński, Adam Siwiec (red.) *Człowiek, zjawiska i teksty kultury w komunikacji społecznej*. Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej; 251-268.
- Szpunar, Magdalena (2020) „O (nie)rzetelnych recenzjach w nauce”. [W:] *Przegląd Pedagogiczny*. Vol. 1; 353–366.

Źródła internetowe

- Centralna Komisja ds. Stopni i Tytułów, <https://www.ck.gov.pl/promotion/p/2/type/1%2Ck/dziedzina/1/dyscyplina/8.html> (dostęp: 1.06-1.07.2023).
- Hamdan, Hussam, Patrice Bellot, Frederic Bechet (2016) *Sentiment Analysis in Scholarly Book Reviews*. arXiv, <http://arxiv.org/abs/1603.01595> (dostęp: 26.04.2024).
- Machado, Samuel, Ribeiro, Ana Carolina, Jorge Vaz Oliveira e Sá (2019) „Machine Learning Algorithms and Techniques for Sentiment Analysis in Scientific Paper Reviews: A Systematic Literature Review”. [W:] *CAPSI 2019 Proceedings*. Vol. 6, <https://aisel.aisnet.org/capsi2019/6> (dostęp: 16.03.2024).
- spaCy, <https://spacy.io> (dostęp: 20.02.2024).