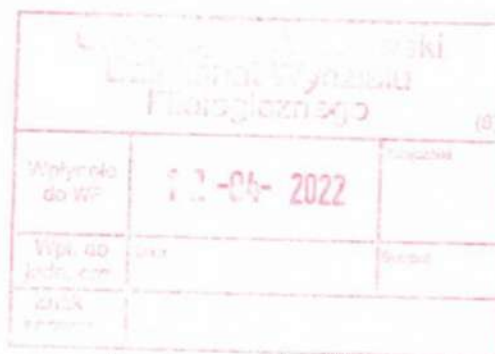


Prof. dr hab. Grażyna Habrajska  
Katedra Dziennikarstwa i Komunikacji Społecznej  
Uniwersytet Łódzki



Recenzja rozprawy doktorskiej mgr Wojciecha Jastrzębskiego pt.

*Komentarze internetowe i ich dyskursywny charakter.*

*Budowa i analiza korpusu na podstawie materiałów z portalu YouTube.*

Rozprawa doktorska mgr Wojciecha Jastrzębskiego poświęcona jest w głównej mierze konstruowaniu sub-korpusów komentarzy internetowych o zróżnicowanej charakterystyce. Stanowi to jego autorski wkład do, prowadzonych od dawna, korpusowych badań naukowych, w których do tej pory nie analizowano komentarzy umieszczanych w Internecie.

Autor zdawał sobie sprawę z trudności terminologicznych, związanych ze stosunkowo „młodymi” badaniami naukowymi dyskursu w ogóle, a komentarzy internetowych w szczególności. Ze względu na bardzo obszerną literaturę dotyczącą dyskursu, mgr Jastrzebski słusznie przyjmuje jedno wybrane ujęcie zaproponowane przez Michaela Fleischera, nie referując szczegółowo różnych koncepcji. Natomiast komentarze ujmuje wyłącznie technicznie, co także ma uzasadnienie w przypadku badań korpusowych. W skrótowy, powierzchowny sposób przedstawia badania jakościowe nad komunikacją językową w Internecie, ze szczególnym uwzględnieniem analiz komentarzy, zaznaczając, że nie są one celem jego rozważań.

Zgodnie z założeniami rozprawy, najwięcej miejsca poświęca autor rozważaniom dotyczącym budowy korpusów i analizy korpusowej. Za wcześniejszymi opracowaniami wyróżnia korpusy ogólne, np. Narodowy Korpus Języka Polskiego i specjalistyczne, np. dotyczące określonego wydarzenia czy określonej tematyki. Zwraca uwagę na to, co można badać wykorzystując korpusy specjalistyczne, wymienia zatem między innymi określenia emocji, niestandardową pisownię,

częstotliwość słów, hashtagów, n-gramów itd. i takie właśnie badania w swojej rozprawie prowadzi.

Szczegółowo opisane zostały przygotowania do budowy korpusu komentarzy internetowych o różnorodnej tematyce. Autor określił kryteria doboru źródeł, takie jak: zasięg mediów, zróżnicowanie materiałów, struktura portalu, personalizacja treści, dostępność narzędzi do zbierania danych czy sposób działania systemu komentarzy. Do swoich badań wybrał portal YouTube o zasięgu ogólnopolskim, który cieszy się dużą popularnością, zatem można było spodziewać się tam licznych różnorodnych komentarzy. Wybór tego portalu wydaje się słuszny do realizacji założonego celu, jaki postawił przed sobą doktorant, także ze względu na jego strukturę i możliwość pominięcia personalizacji treści, poprzez wykorzystanie zakładki „Na Czasie”, gdzie znajdują się różnorodne komentarze, pozwalające na umieszczenie w nich zarówno tekstu jak i emoji. Dla celów rozprawy wybrane zostały komentarze do filmów umieszczonych w tej zakładce na przestrzeni roku (X 2018 – IX 2019), by uniknąć monotematyczności.

Autor w rozdziale trzecim dokładnie prezentuje poszczególne etapy budowania korpusu oraz wykorzystywane w tym celu narzędzia. Opis narzędzi, trudny dla niewtajemniczonego w szczegóły techniczne czytelnika, został przedstawiony w sposób przejrzysty, zrozumiały dla każdego, często w formie tabel, opatrzonych szczegółowymi, podsumowującymi komentarzami.

W rozdziale czwartym doktorant scharakteryzował w sposób ogólny utworzony przez siebie korpus. Najpierw określone zostały osoby, które umieszczały analizowane komentarze pod filmami, bez ujawniania szczegółowych informacji o nich. Takie podejście zgodne było ze wstępnym założeniem autora przejętym od M. Fleischera, że „komunikacje (...) są niezależne od jednostki”. Jest to jedynie możliwe podejście w przypadku ponad dziewięciu milionów komentarzy stanowiących korpus. Następnie wyodrębnione zostały słowa kluczowe, stanowiące podstawę analiz korpusowych, dla całego korpusu i dla sub-korpusów. W tabeli 8 przedstawiona została lista stu słów kluczowych, wyodrębnionych z całego korpusu. Tu w komentarzu oczekiwałabym



jednak większej precyzji. Z materiału wynika wyraźnie, że mamy do czynienia z kilkoma rodzajami słów: wulgaryzmami, wykrzyknikami, słowami polskimi zapisanymi bez znaków diakrytycznych oraz „słowami” nieczytelnymi dla przeciętnego użytkownika języka polskiego (np. *friz*, *suba*, *zebo*, *autotune* i liczne inne). W przypadku tych ostatnich oczekiwałabym wyjaśnienia, co oznaczają, a przynajmniej w jakich kontekstach były używane w analizowanym korpusie. Bez takich wyjaśnień dane liczbowe umieszczone w tabeli są dla czytelnika bezużyteczne. Bardzo interesująco przedstawiają się słowa kluczowe w sub-dyskursach. Z tabeli 11 wynika, że ich częstotliwość jest nieprzewidywalna i zaskakująca (np. w sub-dyskursie edukacja dominuje *czołg*, a w nauce i technice *zgnieść*). O czym to może świadczyć? Warto byłoby tu też dodać komentarz, jakie rodzaje słów dominują w poszczególnych sub-dyskursach, czy te zgodne z ogólną polszczyzną, czy bez znaków diakrytycznych, czy wulgaryzmy itd.

Kolejny rozdział poświęcił autor szczegółowym analizom w obrębie sub-korpusów. Rozpoczął od wyodrębnienia w każdym z nich najpopularniejszych komentarzy, na podstawie przede wszystkim „polubień”. Następnie zwrócił uwagę na sposoby przyciągania uwagi przez komentujących na swoje komentarze, głównie bezpośrednio prosząc o lajki. W tym świetle łatwo zauważyć wątpliwą wartość „polubień” w ocenianiu komentarzy. Doktorant zwraca też uwagę na kompozycję wizualną komentarzy, przede wszystkim emotikony, zwielokrotnienie liter, pisanie wielkimi literami, ale także inne nietypowe kompozycje, takie jak gry czy rebusy. Podsumowanie tych zabiegów graficznych w sposób przejrzysty przedstawione zostało w tabeli 17, a statystyki symboli Emoji w korpusie w tabeli 18. Obie tabele opatrzone zostały przez doktoranta szczegółowymi komentarzami. Najciekawszą w tym rozdziale jest analiza sub-korpusu kontrowersyjnych dyskusji, a zwłaszcza ciągi wyrazowe w dyskusjach wydobywane z korpusu i nie koniecznie ważna byłaby tu ich frekwencja, ale przede wszystkim możliwość skutecznego wykorzystania. Uzyskane wyniki można byłoby wykorzystać do utworzenia bazy dla słownika struktur języka polskiego służącego do tłumaczeń. Bardzo szczegółowo zostały też przedstawione

części mowy w dyskusjach. W tabeli 20 pokazana została ich frekwencyjność. Brakuje mi tutaj wyjaśnienia skrótów kategorii, które nie odpowiadają ogólnie znanym z gramatyk. Niektóre z tych kategorii opisuje autor w komentarzu, ale dla porządku powinno się podać ich standardowe odpowiedniki. Ciekawe wnioski wynikają także z analizy wulgaryzmów w komentarzach. Przy bardzo dużej częstotliwości występowania uderza ich bardzo ubogi repertuar. Autor słusznie zauważa, że użycie słów uznawanych za wulgarne nie zawsze ma charakter agresywny, co jest faktem od dawna znanym, natomiast wyraża zdziwienie, że nie ma zbyt wielu wulgaryzmów w sub-korpusie *Wiadomości i polityka* w przeciwieństwie do tego, co zdiagnozowała wcześniej Elżbieta Laskowska. Myślę, że łatwo to wyjaśnić odwołując się do młodego wieku internautów komentujących filmy na YouTube, których polityka rzadko, a zwykle w ogóle nie interesuje. Weryfikacja takiej tezy w rozprawie nie była możliwa ze względu na przyjętą zasadę nieuwzględniania informacji o autorach komentarzy w analizie materiału.

Ostatni, szósty rozdział rozprawy, stanowią analizy zapożyczeń, neologizmów i zjawisk słowotwórczych, występujących w komentarzach do filmów. Tabela 26 zawiera listę „nowych” słów, które, po odrzuceniu nazw własnych, nie zawsze są nowe, np.: *łapka, bomba, kontent, sezon, epicki, kciuk, dzwon, zacny, pelikan, ustawka, dram, autobus, suchar, słodziak, bąbelek, chmura, dzida* itd., być może, a nawet prawdopodobnie, zyskały one nowe znaczenie w komentarzach internetowych, ale, moim zdaniem, nie można ich nazywać neologizmami bez opisanie znaczeń, w jakich funkcjonują w analizowanym materiale. Nie wydaje się też trafne określenie „wyrazy niestownikowe” (podrozdział 6.2) dla słów błędnie zapisanych (głównie bez znaków diakrytycznych), nazw własnych, zapożyczeń czy różnego rodzaju skrótów/symboli, które nigdy nie znalazły się w słowniku języka polskiego, z założenia promującego poprawną polszczyznę i pomijającego nazwy własne. Moim zdaniem, nie ma ogólnej nazwy polskiej, która obejmowałaby te wszystkie tak różne kategorie słów, które nie tworzą żadnego spójnego układu. Może lepsze byłoby określenie wyrazy charakterystyczne dla komentarzy internetowych?



Albo wyrazy wyróżniające komentarze? Nie jest też trafny tytuł podrozdziału 6.3 „wyrazy z sufiksami”, w którym autor wyróżnia tylko dwa sufiksy (-owicz i -er nie -ber) najczęściej wykorzystywane do włączania zapożyczeń do polskiego systemu językowego. Może lepiej byłoby nazwać tę część rozważań „Sufiksy najczęściej występujące w komentarzach”, albo „Sufiksy -owicz i -er w komentarzach na YouTube”? Tu należy zwrócić uwagę, że w licznie występujących odmianach słowa *youtuber* mamy to do czynienia z sufiksem -er, a nie -ber: YouTube+er. Natomiast słusznie pominięte w analizie słowa *cyber*, *uber*, *bimber* w ogóle nie mają przyrostka, a *bober* ma przyrostek -er tylko w gwarowej nazwie odmiany grochu *bób* / *bober*, nie sądzę jednak, by w komentarzach do filmów wyraz ten miał takie znaczenie. Rozdział ten kończy opis wyrazów z reduplikacją liter, które autor zebrał w tabeli 32. Słusznie tu zauważa, że w niektórych przypadkach jest to wynik niestaranności komentującego.

Rozprawę kończy starannie opracowane podsumowanie. Cechy analizowanych komentarzy zestawione zostały w tabeli 34 ze szczegółowym objaśnieniem jej zawartości. Uwagę w podsumowaniu zwraca syntetyczna charakterystyka czterech typów komentarzy: ekspresywnych, specjalistycznych, interdyskursywnych i cyfrowonatywnych, które w rozprawie nie zostały tak przejrzysto wyodrębnione. Wątpliwości moje, w kontekście recenzowanej rozprawy, budzi określenie „komentarze interdyskursywne” w odniesieniu do, jak pisze doktorant: „komentarzy o „neutralnej” charakterystyce, najbliższej komentarzowi rozumianemu jako taki, znanego również z innych mediów” (s. 139). Trafniejsza byłaby chyba nazwa „komentarze intermedialne”. Ja doskonale rozumiem, że autor stara się ściśle trzymać przyjętej przez siebie definicji dyskursu zaproponowanej przez Michaela Fleischera, którą także pozwolę sobie tu przytoczyć (cytat na który powołuje się w rozprawie doktorskiej mgr Jastrzębski):

Dyskurs jest kulturowo uwarunkowanym sposobem tego, jak i za pomocą jakich interpretantów dana subkultura się wypowiada, zabiera głos w świecie znakowym, a to znaczy w kulturze, oraz zapewnia swoją spójność (koherencję)

W rozprawie jednak autor nie odwołuje się ani do pojęcia subkultury i nie stara się jej określić, ani do pojęcia interpretantów, posługuje się tradycyjnymi terminami lingwistycznymi, nie wskazuje też koherencji w obrębie dyskursu. Uważam, że bardziej stosowny do treści rozprawy byłby tytuł: *Komentarze internetowe i ich charakterystyka językowa* (bowiem szeroko scharakteryzowane zostały części mowy; słownictwo: wulgaryzmy, zapożyczenia; reduplikacje), z pominięciem, modnego wprowadzie, ale niebezpiecznego pojęcia dyskursu, które jednak w większości ujęć wiąże się z analizą jakościową i uwzględnia treści pozajęzykowe. Do rozprawy dołączono 5 załączników oraz na płycie pełen analizowany materiał.

Niewątpliwą zaletą rozprawy doktorskiej mgr Wojciecha Jastrzębskiego jest utworzenie korpusu tekstów komentarzy umieszczanych pod filmami w portalu YouTube oraz sub-korpusów, które mogą być wykorzystywane w dalszych badaniach porównawczych.

**Przedstawiona do recenzji rozprawa mgr Wojciecha Jastrzębskiego spełnia wymogi niezbędne do uzyskania stopnia naukowego doktora. Realizuje merytoryczne, metodologiczne i badawcze standardy naukowe, stawiane wobec rozpraw doktorskich. Wnoszę zatem o dopuszczenie mgr Wojciecha Jastrzębskiego do dalszych etapów przewodu doktorskiego.**

Prof. dr hab. Grażyna Habrajska

