



Państwowa Wyższa
Szkoła Zawodowa
w Koninie

Professor Ordinarius Dr habil. Barbara Lewandowska-Tomaszczyk
Head, Department of Research into Language, Literature and Translation
State University of Applied Sciences in Konin
Przyjaźni 1, 62 510 Konin



tel.

(63) 249-72-00

Email: blt@konin.edu.pl

**Recenzja osiągnięć naukowych dla oceny dorobku dr Jana Rybickiego
w postępowaniu habilitacyjnym w dziedzinie *nauk humanistycznych*,
dyscyplinie *językoznawstwo***

(1) Sylwetka Habilitanta

Dr Jan Rybicki jest absolwentem anglistyki Akademii Pedagogicznej w Krakowie, gdzie uzyskał także stopień doktora nauk humanistycznych na podstawie rozprawy *A Computer-Assisted Comparison of Two English Translations of Henryk Sienkiewicz's Trilogy*, w 2000 roku. Przez kilka lat zatrudniony tam na stanowisku adiunkta, w roku 2011 rozpoczął zatrudnienie w Uniwersytecie Jagiellońskim w Instytucie Filologii Angielskiej, także na stanowisku adiunkta. W tych latach uzyskał dwukrotnie stypendium Fundacji Kościuszkowskiej, które odbywał w prestiżowej amerykańskiej uczelni Rice University.

(2) Osiągnięcia naukowe

Głównym obszarem zainteresowań naukowych dr Jana Rybickiego jest stylometria, szczególnie w porównaniu tekstów oryginałów i ich przekładów, oraz atrybucja tekstów. Jest to dział językoznawstwa coraz energiczniej rozwijający się wraz z rozwojem technologii informacyjnych, używany w lingwistyce kryminalistycznej, mający swoje miejsce w analizach autorstwa dzieł literackich. Dr Rybicki zajmuje się tym analizą tekstów literackich. Jest osobą aktywną w kręgach humanistyki cyfrowej, członkiem komitetu wykonawczego *Association for Literary and Linguistic Computing*, obecnie - *European Association for Digital Humanities* przez dwie

kadencje oraz członkiem Rady Naukowej prestiżowego czasopisma naukowego *Literary and Linguistic Computing/Digital Scholarship in the Humanities*.

Był dr Jan Rybicki uczestnikiem COST Action IS0901 *Women Writers in History: Toward a New Understanding of European Literary Culture* (2009–2013), w ramach której powstały niektóre z jego publikacji (publikacja dotycząca porównania języka autorek i autorów tekstów). Kierował także lub był głównym wykonawcą projektów Narodowego Centrum Nauki (dwa projekty serii Opus).

Oprócz zainteresowań naukowych w dziedzinie humanistyki cyfrowej, szczególnie stylometrii, w których prezentuje swoje główne osiągnięcia badawcze dla celów obecnej recenzji, oraz innych publikacji o tej tematyce (25 publikacji innych zgłoszonych w spisie publikacji, 626 cytowań łącznych, *Google scholar citation list*, 16.1.2019 r.), będących analizą poszczególnych przypadków użycia rozwijanej metodologii, jest także dr Rybicki aktywnym tłumaczem literatury anglosaskiej. Przetłumaczył 28 powieści, co jest sygnałem powiązań zainteresowań badawczych Habilitanta z praktyką językową oraz refleksją tłumaczeniową i porównawczą. Dr Rybicki prezentuje także swoje zainteresowania analizą literatury, szczególnie w ujęciu przekładowym i komparatystycznym, mające swój wyraz w publikacjach (także współautorskich), szczególnie o wątku recepcji prozy Sienkiewicza w krajach anglosaskich.

(3) Główne osiągnięcia naukowe

Cykl szesnastu powiązanych publikacji na temat:

SYGNALY STYLOMETRYCZNE W JĘZYKU LITERACKIM W ORYGINALE I PRZEKŁADZIE

Publikacje dr Jana Rybickiego zgłoszone do recenzji lokują się w przedziale badań ilościowych nad językiem na dużych lub bardzo dużych zbiorach tekstów. Są to w dużej mierze artykuły współautorskie, szczególnie w odniesieniu do prac o charakterze teoretyczno-metodologicznym. Prace wykonane przez dr Rybickiego w tej grupie publikacji dotyczą przygotowania tekstów elektronicznych, udziału w analizie komputerowej, analizę wyników, oraz redakcji tekstu. W pracach o charakterze

weryfikacyjno-interpretacyjnym, dr Rybicki prezentuje do oceny większość swoich publikacji jedno-autorskich.

Cykl recenzowanych publikacji zawiera dwie grupy zagadnień tematycznych. Pierwsza grupa to publikacje związane z zagadnieniami metodologicznymi (sześć artykułów, w tym jeden autorski), pozostałe zaś publikacje (dziewięć autorskich), to artykuły analityczno-interpretacyjne. Udział autorski dr Rybickiego w opublikowanych pracach mieści się w przedziale od 20 % do 60 %, przy czym prace w drugiej grupie tematycznej są w przeważającej większości autorskie, lub wkład dr Rybickiego przekracza 50 %. Odniosę się do każdej z publikacji bardziej szczegółowo.

Fakt, o którym Autor pisze w autoreferacie, iż ilościowa analiza tekstów potwierdza intuicje dotyczące analiz jakościowych tekstów, jest dziś weryfikowany na obszernych materiałach korpusów językowych i popierany wynikami badań statystycznych. W dużej mierze także następna cecha badań kwantytatywnych jest także widoczna w pracach publikowanych. Jest ona związana z analizami typu *corpus-driven*, w odróżnieniu od *corpus-based* i wskazuje na fakt, że dane ilościowe i analiza statystyczna takich elementów językowych jak frekwencje słów czy n-gramów, i to nie tylko wyrazów autosemantycznych, znaczących, lecz także synsemantycznych, tzw. strukturalnych, czy funkcyjnych, mogą być źródłem nowych, nie zakładanych przez badacza tez. Ponadto – zarówno jedno, jak również drugie – mogą stać się wskaźnikiem kryterialnym („stylometrycznym odciskiem palca”) w procesie atrybucji testów.

Prace dr Rybickiego koncentrują się na identyfikacji metod i narzędzi analizy tekstu, które byłyby przydatne do możliwie dokładnej charakterystyki stylometrycznej autorów tekstów i ich przekładów z jednej strony, zaś z drugiej - na weryfikacji tych metod i narzędzi na wybranych materiałach tekstów literackich oraz ich przekładów. Proces ten ma prowadzić do ustalenia tzw. stylometrycznego sygnału autora czy tłumacza, w ich - jak można ją określić – tożsamości językowej. Ten sygnał stylometryczny wyliczany jest w terminach Delt Burrowsa i jej modyfikowanych wersji, czyli poprzez porównywanie ciągów frekwencyjnych najczęstszych słów występujących w danym zestawie tekstów. Porównanie takie przeprowadzane jest we

wszystkich parach tekstów dla określenia miary odległości międzytekstowej, czyli stopnia podobieństwa oraz rozbieżności między nimi.

W grupie prac metodologicznych występują cztery artykuły, trzy podstawowe, czwarty – bardziej techniczny, skupiający się na analizie programu komputerowego samego narzędzia statystycznego używanego w prezentowanych analizach.

Dwie prace współautorskie (J. Rybicki, M. Eder, *Deeper Delta Across Genres and Languages: Do We Really Need the Most Frequent Words?*, 2011 M. Eder, J. Rybicki, *Do birds of a feather really flock together, or how to choose training samples for authorship attribution*, 2013) składają się na ambitny projekt weryfikujący metodologię obliczeniową atrybucji autorskiej w prozie, poezji oraz tekstach dramatycznych dla siedmiu, typologicznie różnych języków. W pierwszej pracy wzięto pod uwagę dwa parametry – liczbę i frekwencję wyrazów - oraz zastosowano pakiet statystyczny R. Uzyskano interesujące wyniki, wrażliwe jednak na typ systemu językowego, zaś praca, opublikowana w czasopiśmie z listy A MNiSW, cieszy się sporą liczbą cytowań (73 cytowania w dn. 16.1.2019 w *google citation index*). W drugiej publikacji zastosowano ten sam korpus, zaś liczba tekstów i autorów była zmienna. Zastosowano też metodę sprawdzianu krzyżowego przy imponującej liczbie 500 prób. Najwyższe wyniki uzyskało badanie dla korpusu angielskiego, najśłabsze – dla danych polskich, co Autorzy przypisywali m.in. językowym różnicom typologicznym, ale także stopniowi reprezentatywności doboru testów. Skłoniło to dr Rybickiego do następnej – autorskiej - próby (J. Rybicki, *Success Rates in Most-frequent-word-based Authorship Attribution. A Case Study of 1000 Polish Novels from Ignacy Krasicki to Jerzy Pilch*, 2015), przeprowadzonej na dość dużym korpusie reprezentatywnych próbek literatury języka polskiego w dużym wymiarze historycznym, starannie dobranych autorów oraz podobnych parametrów ilościowych. Bardzo interesującym wynikiem badania jest stwierdzenie, iż popularne w językoznawstwie korpusowym n-gramy są gorszym miernikiem sygnału autorskiego niż pojedyncze słowa. Nie bez znaczenia jednak wydaje się tu fakt, iż dr Rybicki używał ulepszonych algorytmów statystycznych (dyskusja dotyczy n-gramów anotowanych częścią mowy o zakresie od 2 do 6). Wynikiem badań recenzowanych na podstawie tych publikacji może być jednak stwierdzenie, iż nie da się wyłonić maksymalnie skutecznych warunków atrybucji tekstów. W każdym z badanych

przypadków stwierdzano silną jej zależność od typu języka, od gatunku tekstu, okresu, w którym powstał, itd.

W pewnej mierze niezależnie od badań ilościowo-analitycznych, przedstawionych do oceny powstał artykuł z roku 2016 (M. Eder, J. Rybicki, M. Kestemont, *Stylometry with R: A Package for Computational Text Analysis*, "R Journal" 2016), który prezentował pakiet komputerowy *Stylo* (rozszerzony program R z nakładką graficzną), używany w recenzowanym materiale. Dr Rybicki miał w nim mniejszy udział, głównie przy sprawdzaniu skuteczności działania programu (publikacja w czasopiśmie o 5-letnim IF 2.114).

Recenzowane w pierwszej – metodologicznej – grupie publikacje tworzą zwarty, logiczny cykl, w których każda z kolejnych prac podejmuje wyjaśnienie wątpliwych obszarów w poprzednich badaniach.

Druga część cyklu zawiera prace weryfikacyjne oraz interpretujące w odniesieniu do poszczególnych utworów literackich w tekstach oryginalnych oraz w przekładach literatury anglosaskiej na język polski oraz polskiej na język angielski. Interesującą linią przewodnią prac jest pytanie badawcze o widzialność i znak rozpoznawczy autora oraz tłumacza, przy użyciu opracowanych metodologii stylometrycznych. Publikacja sprawdzająca atrybucje tekstów w niejasnych przypadkach autorstwa i przekładu (J. Rybicki, *Ślady żony tłumacza. Alma Cardell Curtin i Jeremiah Curtin*, 2011), poza niewątpliwym tłumaczeniowym, atrybutywnym walorem, prezentuje także swój walor metodologiczny. Jest nim użycie wersji pakietu *Stylo* w procesie uczenia nienadzorowanego przy zastosowaniu bardziej lokalnego narzędzia tzw. pseudo-bootstrapu konsensusowego (pozyskiwanie uśrednionych dendrogramów w analizie skupień). Narzędzie to charakteryzuje się iteratywnym podejściem do poszukiwanych rozwiązań i działa na obszernych korpusach tekstów.

Bardzo interesującą w odniesieniu do wyników publikacją autorską dr Rybickiego, oraz najczęściej cytowaną z autorskich prac (49 cytowań 16.1.2019) jest J. Rybicki, *The Great Mystery of the (Almost) Invisible Translator: Stylometry in Translation*. (2012), która odkrywa silniejsze sygnały leksykalne językowej tożsamości autora tekstów oryginalnych wobec językowej tożsamości tłumacza. Praca jest

metodologicznie bez zarzutu, wnioski są logiczne, popierające znane w translatologii tezy o niewidzialności tłumacza. Podobną tematykę podejmuje kolejna praca, w języku polskim (J. Rybicki, *Stylometryczna niewidzialność tłumacza*, 2013). Oprócz podobnych wniosków, dr Rybicki identyfikuje tu jednak także sygnały tożsamości językowej niektórych z bardziej znanych tłumaczy literatury na język polskich (T. Boy-Żeleński).

Dr Rybicki kontynuuje tematykę relacji między tożsamością językową autora a jego tłumacza w następnej pracy (J. Rybicki, *Stylometric translator attribution: Do translators leave lexical traces?*, 2013). Potwierdza w niej uzyskane poprzednio wnioski na takich tekstach, w których jeden oryginał odpowiada jednemu przekładowi. Ponadto, rozszerza swoją analizę na porównanie przekładów własnych z przekładami innych tłumaczy. Wnioski, które prezentuje można, tu zgadzam się z dr Rybickim, rozpatrywać w terminach znanych w naukach o przekładzie teorii polisystemów Itamara Even-Zohara. Rozrzut między silniejszymi a słabszymi wyznacznikami stylometrycznej tożsamości tłumacza sygnalizowany jest w J. Rybicki, *Tabakowska kontra Pawelec? Stylometria polskich przekładów twórczości N. Daviesa*. 2012, opartym na obszernym materiale prozy eseistyczno-historycznej.

Współautorska publikacja J. Rybicki, M. Heydel, *The Stylistics and Stylometry of Collaborative Translation: Woolf's 'Night and Day' in Polish*, 2013 (czasopismo listy A), analizuje tekst kolaboratywnie tłumaczony. Ważnym wynikiem analizy jest potwierdzenie pytania badawczego pracy o autorstwo odpowiednich części oryginału. Jednocześnie badanie pozytywnie weryfikuje sprawdzaną w poprzednich pracach konsensusową metodologię pseudo-bootstrapową. Następną, współautorską, publikacją (J. Rybicki, D. Hoover, M. Kestemont. *Collaborative authorship: Conrad, Ford and rolling delta*, 2014), pogłębia podejście do tematyki atrybucji tekstów, przy autorstwie kolaboratywnym, poprzez dodanie do pakietu *Stylo* funkcji *rolling.delta* dla analizy autorstwa powieści współautorskiej Josepha Conrada i Forda Madox Forda. Jako materiał tekstowy dr Rybicki używa analizę cech stylometrycznych samodzielnie tworzonych tekstów tych autorów. Oprócz dużej wagi dla językoznawstwa, bardzo interesujące wyniki prezentowanej analizy mają także duże znaczenie dla badaczy literatury, w szczególności w dziedzinie konradologii.

Wynikiem zarówno badań stylometrycznych dr Rybickiego z jednej strony, jak również udziału w Akcji COST IS0901 na temat pisarstwa kobiet w ujęciu historyczno-kulturowym z drugiej, jest publikacja J. Rybicki, *Vive la différence: Tracing the (Authorial) Gender Signal by Multivariate Analysis of Word Frequencies*, 2016, także z czasopisma listy A. Praca ta raportuje wyniki badań nad sygnałami płci kulturowej w korpusie angielskich tekstów z Chawton House przy pomocy wskaźnika Zeta i jego użyciu w pakiecie *Stylo* (generowanie list słów charakterystycznych; wielowymiarowa analiza skupień preferowanych przez daną płć). Interesujące wyniki omawianych badań, szczególnie w odniesieniu do wykazu pojęć, jak Autor je nazywa preferowanych, byłyby ważnym punktem odniesienia do porównań w tej mierze z tekstami współczesnymi, w których – jeśli chodzi o ówczesną tematykę i słowa charakterystyczne (pojęcia związane z rodziną i uczuciami przez pisarki oraz z honorem, rycerskością itd. przez pisarzy), oczekiwałabym zaskakujących zmian, zapewne nie ze względu na wspomniane kobiece i męskie preferencje, lecz na ogólny kontekst społeczno-kulturowy. Autor odnosi się już do części tej kwestii (str. 755), biorąc pod uwagę teksty pisane później, jednak wyniki, choć interesujące (np. w literaturze autorstwa mężczyzn preferowane formy to *wojna* i *pieniądze*), nie są w pełni jednoznaczne. Jest również ważne, jak – bardziej dokładnie w odniesieniu do kwestii weryfikacji - plasuje się zastosowana metodologia (Autor wskazuje w autoreferacie na „2-3 błędy na kilkadziesiąt tekstów”), która według Autora potwierdzałaby wyniki o dominacji w pisarstwie płci kulturowej autora.

Ponadto, nowy sposób wizualizacji wyników w postaci analizy sieciowej, co Autor potwierdza w dalszych złożonych do oceny publikacjach, jest rzeczywiście czytelniejszy interpretacyjnie i można z nim wiązać pewne nadzieje dla identyfikacji dalszych i silniejszych podobieństw tekstowych. To właśnie w publikacji J. Rybicki, *Pierwszy rzut oka na stylometryczną mapę literatury polskiej*, 2014, Autor potwierdza swoje wnioski generując mapy stylometryczne ‘czytanych na dystans’ tekstów literackich, układające się zarówno w porządku chronologicznym, jak również autorskim oraz typologicznym-literackim. Dalsze uszczegółowienie tej metodologii znajdujemy w pracy J. Rybicki, *Drugi rzut oka na stylometryczną mapę literatury polskiej*, 2017, w której Autor prezentuje modyfikację miary odległości międzytekstowej, przez wprowadzenie Delty kosinusowej, przez co oczekuje poprawy skuteczności wyników atrybucji autorskiej oraz przekładowej, co w dużej mierze

potwierdza się w analizie. W porównaniu z wynikami poprzedniej omawianej tu publikacji tym razem jeszcze silniej dostrzec można zróżnicowanie stylu i gatunków literackich. W odniesieniu zaś do tekstów przekładowych Autor dostrzega udaną próbę skupiania się tekstów o zbliżonej specyfice kulturowo-językowej (np. literatura polska i rosyjska). Traktuje także te wyniki, w sposób moim zdaniem zasadny, jako ilościowe dowody na istnienie stylometrycznego *translationese*, systemu pośredniego między językiem źródłowym a docelowym w przekładzie.

Dwie ostatnie publikacje z prezentowanego cyklu są przykładem dalszych zastosowań metodologii stylometrycznej do badania innych typów tekstów. W swojej publikacji J. Rybicki, *Partners in Life, Partners in Crime?* Autor analizuje ponownie atrybucję tekstów tłumaczeniowych, tym razem z jednak literatury niemieckiej, w odniesieniu do przypuszczalnej współpracy translatorskiej, W pracy współautorskiej M. Choiński, J. Rybicki, *Is God Really Angry at Sinners? A Stylometric Study of Jonathan Edwards's Representations of God.*, 2017, Autor bada język amerykańskich homilii purytańskich, analizując nacechowanie pojęcia leksykalnego *God*. Posługuje się tu głównie analizą frazeologiczną i interpretacją częstotliwości występowania kolokatów.

Podsumowując ogólnie ocenę, chcę zauważyć, iż stosowanie metodologii ilościowej nie prowadzi Autora do przypadkowej rotacji metodologicznej i każdorazowej pełnej zmiany podejścia analitycznego w kolejnych pracach. Dr Rybicki zresztą wyraża to dobitnie w artykule z 2016 roku (Vive...): "it is more important to use stable methodology, even if this means losing a percent of attribution success rate here and there, rather than to multiply unknown variables by testing ever-new algorithms while trying to solve a literary (or linguistic) question. " (str. 749/50). Moim zdaniem jest to cenna zasada przy analizie porównawczej tekstów, której wyniki zmieniają się pod wpływem modyfikacji używanych metod, lecz przecież rozkładają się na teksty różnych odmian i języków, różnych autorów, powstałych w różnym czasie w historii.

Recenzowany cykl publikacji dr Jana Rybickiego jest przemyślanym, usystematyzowanym zbiorem prac. Wnoszą one do nauki zarówno perspektywę humanistyczną filologa, jak również nowe spojrzenie na metodologie ilościowe badań języka oraz rzucają światło na analizy ilościowe atrybucji tekstów autorskich i tłumaczeniowych. Metodologia stylometryczna nie rozwiązuje naturalnie wszystkich

problemów identyfikacji atrybutów idiolektycznego stylu autorskiego czy translatorskiego. Po pierwsze, nie zajmuje się wszystkimi subtelnymi właściwościami tworzonego tekstu, po drugie zaś – nie zawsze jest możliwe ustalenie tekstu referencyjnego. Pomimo tych problemów, metody stylometryczne używane przez Autora są, co wynika z prezentowanych prac, coraz bardziej pomocne zarówno w analizie zróżnicowania stylistycznego tekstów z jednej strony, jak również, z drugiej, w ustalaniu autorstwa tekstów oraz w przeprowadzaniu – weryfikowalnego i stabilnego - porównania siły bliskości i rozbieżności między nimi oraz siły powiązań między tekstami odmiennymi.

(4) Działalność dydaktyczna i popularyzatorska

Oprócz kursowych zajęć dydaktycznych, dr Rybicki bierze czynny udział w seminariach i warsztatach prowadzonych we współpracy między Instytutem Filologii Angielskiej UJ a Instytutem Języka Polskiego PAN w Krakowie dla doktorantów i pracowników akademickich instytucji zagranicznych.

Dr Rybicki uczestniczy także aktywnie w spotkaniach i dyskusjach popularyzujących naukę w audycjach radiowych, TV oraz w mediach internetowych.

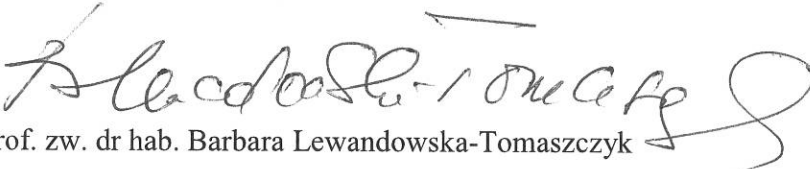
(5) Działalność organizacyjna

Dr Jan Rybicki jest aktywnym działaczem na niwie humanistyki cyfrowej. Działa w towarzystwach temu poświęconych, był też współorganizatorem bardzo udanej edycji kongresu *Digital Humanities* w 2016 roku w Krakowie. Organizował poprzednio konferencje i seminaria stylometryczne: EADH (2010), seminarium Stylometry@Kraków (2012) oraz Stylometry Workshop (2016), zaś w latach 2014 i 2017 r - panele stylometryczne w ramach innych konferencji. W konsorcjum DARIAH-PL ds. humanistyki cyfrowej jest członkiem Rady Konsorcjum, zaś w jego strukturach europejskich DARIAH-EU, pełni funkcję członka Rady Naukowej. W ramach COST Action CA16204 *Distant Reading for European Literary History* (2017-2021), jest członkiem Komitetu Zarządzającego oraz przedstawicielem Polski.

(6) Wniosek końcowy

Stwierdzam, po analizie wagi poszczególnych elementów dorobku naukowego dr Jana Rybickiego i skonfrontowaniem ich z art. 18 ust. 5, 2003 r. Centralnej Komisji do Spraw Stopni i Tytułów oraz Dz. Ustaw 2018 poz. 1669, w sprawie oceny osiągnięć osoby ubiegającej się o nadanie stopnia doktora habilitowanego w dziedzinie nauk humanistycznych, dyscyplinie językoznawstwo, iż przedstawione osiągnięcia naukowe oraz recenzowany dorobek naukowy dr Jana Rybickiego z nawiązką spełniają warunki do postawienia wniosku o przyznanie Autorowi stopnia doktora habilitowanego. Dorobek naukowy, dydaktyczny oraz organizacyjny w pełni uzasadniają wniosek o nadanie dr Janowi Rybickiemu stopnia doktora habilitowanego w dziedzinie nauk humanistycznych, dyscyplinie językoznawstwo.

16.01.2019 r.


Prof. zw. dr hab. Barbara Lewandowska-Tomaszczyk