

Dr hab. Rafał L. Górski
Wydział Filologiczny
Uniwersytetu Jagiellońskiego

Kraków, 14.02.2019

Recenzja osiągnięcia naukowego „*Sygnały stylometryczne w języku literackim w oryginale i przekładzie*” oraz całokształtu dorobku naukowego, dydaktycznego, współpracy naukowej i popularyzacji nauki pana dr Jana Rybickiego w związku z postępowaniem w sprawie nadania stopnia naukowego doktora habilitowanego

Dr Jan Rybicki rozpoczął karierę naukową w roku 1991 jako asystent w Instytucie Neofilologii ówczesnej Wyższej Szkoły Pedagogicznej w Krakowie. W roku 2000 uzyskał tamże doktorat na podstawie dysertacji *A Computer-Assisted Comparison of Two English Translations of Henryk Sienkiewicz's Trilogy*, napisanej pod kierunkiem dra hab. Jana Prokopa. Pracował na tej uczelni do roku 2011, najpierw na stanowisku asystenta, później adiunkta, z dwiema przerwami, gdy wykładał jako Stypendysta Fundacji Kościuszkowskiej na Rice University w Teksasie. W roku 2011 wygrał konkurs na stanowisko adiunkta w Instytucie Filologii Angielskiej Uniwersytetu Jagiellońskiego, gdzie dotąd pracuje.

1. Ocena głównego osiągnięcia

Lingwistyka i literaturoznawstwo to bez wątpienia odrębne dyscypliny naukowe z niewielkim punktem styczności, połączone ze sobą na zasadzie pewnej tradycji w naukę zwaną filologią. Do tych dwu dyscyplin dołączyło ostatnio przekładoznawstwo, któremu znów odmawia się odrębności i traktuje najczęściej jako subdyscyplinę językoznawstwa stosowanego bądź, rzadziej, komparatystyki literackiej. Jednym z takich punktów stycznych tych trzech dyscyplin jest stylometria. Stylometria interesuje się policzalnymi cechami tekstu, są to z reguły cechy językowe, które z praktycznych względów ograniczają się niemal wyłącznie do leksyki. Pytania które stawia stylometra mogą być istotne już to dla literaturoznawcy, już to językoznawczy czy przekładoznawcy. Piszę o tym ponieważ kariera Habilitanta przebiega nieco nietypowo, doktorat uzyskał on bowiem w zakresie literaturoznawstwa – choć trzeba od razu dodać, że to czym się zajmował to przekład literacki, który badał metodami humanistyki cyfrowej. Jego dorobek mieści się więc pośrodku tego co tradycyjnie zwiemy filologią, a więc w punkcie

styczności trzech wymienionych wyżej dziedzin; z pewnością nie jest to jądro lingwistyki, ale mieści się w szeroko pojętym językoznawstwie, jako badanie cech językowych (a ściślej leksykalnych) tekstu. Obszar, w którym się dr Rybicki porusza dobrze charakteryzuje tytuł czasopisma naukowego, w którym opublikował On kilka prac: „*Literary and linguistic computing*”.

Habilitant jako główne osiągnięcie wskazał cykl prac, który zatytułował „*Sygnały stylometryczne w języku literackim w oryginale i przekładzie*”. Na cykl składa się 16 prac. 6 z nich zostało ogłoszonych w czasopismach z tzw. listy A (5 *Literary and Linguistic Computing* – obecnie zatytułowane *Digital Scholarship in the Humanities*, 1 w *R Journal*), 3 w czasopismach listy ERIH, zaś 6 w to rozdziały w książkach zbiorowych.

Tym co łączy cykl to stylometria, obszar badawczy, który zajmuje się badaniem stylu tekstu, stylu rozumianego bardzo specyficznie, mianowicie jako ilościowa charakterystyka wybranych elementów językowych, które w wypadku Habilitanta ograniczają się do leksyki. Z punktu widzenia informatyka, stylometria to po prostu kategoryzacja tekstu, z punktu zaś widzenia filologa istotne jest to, co podobieństwo jednego tekstu do drugiego mówi na jego temat. W cyklu znajdziemy oba podejścia.

1.1. W pracach [1] i [2] Habilitant wraz ze współautorem Maciejem Ederem badają skuteczność atrybucji autorskiej, a przede wszystkim to, na ile owa skuteczność zmienia się w odniesieniu do różnych języków, typów tekstów i liczby słów branych pod uwagę. Z kolei w [2] ten sam zespół autorów zbadał na ile na atrybucję autorską wpływa dobór zestawu uczącego.

Te dwie prace są o tyle istotne, że zastosowania metod ilościowych w humanistyce są nowe, stąd stosunek do dowolnego uzyskanego wyniku był dosyć bezkrytyczny. To między innymi tym badaniom zawdzięczamy to, że wymagania, jakie stawia się eksperymentom w zakresie humanistyki cyfrowej znacząco wzrosły w ostatniej dekadzie. Obie prace doczekały się niemal 100 cytowań. Rozwinięciem tych badań jest praca [3].

Inny charakter ma praca [4] opublikowana w czasopiśmie informatycznym poświęconym językowi programowania i środowisku do obliczeń statystycznych R. W tym wypadku nie chodzi o referowanie wyników badań, ale opis pakietu *Stylo*. Udział Habilitanta w jego stworzeniu polegał na dodaniu interfejsu graficznego, co z pewnością przyczyniło się do zwiększenia i tak dużej popularności tego narzędzia do badań stylometrycznych. W tym więc wypadku należy raczej oceniać oprogramowanie niż sam tekst. O przydatności pakietu *Stylo* świadczy 51 cytowań, przy czym należy zauważyć, że tekst ten jest dużo młodszy niż sam pakiet, co w oczywisty sposób zmniejsza liczbę potencjalnych odwołań bibliograficznych.

1.2. Prace [6], [7], [8], [10] to próba zastosowania stylometrii do tego by uzyskać odpowiedź na bardzo istotne z punktu widzenia przekładoznawstwa pytanie. Chodzi mianowicie o to, kto jest „autorem tłumaczenia” – tłumacz, czy pisarz, który stworzył tekst oryginalny. Ściślej: jest to pytanie z jednej strony o zakres swobody twórczej tłumacza, a z drugiej o to, na ile tekst oryginału przenika do tłumaczenia. Powtarzane eksperymenty wskazują na to, że stylometrycznie bliskie są sobie teksty, które wyszły spod pióra jednego pisarza, a nie te, które są dziełem jednego tłumacza. Przy czym chodzi o podobieństwo w rozkładzie ilościowym wyrazów częstych, czyli wyrazów synsemantycznych, albo też autosemantycznych, ale wieloznacznych, tak więc tematyka utworów nie ma tu nic do rzeczy.

W pracy [6] Habilitant przeprowadza eksperymenty na korpusach tłumaczeń z angielskiego na francuski, francuskiego na angielski oraz z angielskiego na polski, wreszcie polskiego na angielski. Generalnie wykresy potwierdzają, że tłumaczenia wykazują raczej podobieństwa pomiędzy tekstami jednego pisarza niż tekstami jednego tłumacza.

Podobne eksperymenty są powtórzone w pracach [7] i [8]. Pierwsza z nich jest o tyle wyjątkowa w twórczości Habilitanta, że zajmuje się w niej przekładem tekstów nieliterackich. W pracy tej Autor opisuje ciekawy eksperyment, mianowicie bada sygnał autorski w pojedynczym tekście przetłumaczonym przez trzech różnych tłumaczy – tekst jest podzielony na szereg fragmentów, z których na potrzeby badania każdy jest traktowany tak, jak by był odrębnym utworem. W takiej sytuacji widać separację fragmentów, które wyszły spod różnych piór tłumaczy. Wynik jest co prawda w jakimś zakresie raczej przewidywalny, ale z pewnością warto było tę intuicję poprzeć twardymi danymi. Praca [8] jest metodologicznie bliska [6], operuje jednak częściowo na innych danych. [10] z kolei pokazuje, że w twórczości pojedynczego tłumacza widać separację podług autora tekstu źródłowego. W tym miejscu warto zwrócić uwagę, że materiałem badawczym jest twórczość samego Habilitanta, który – zanim zajął się przekładoznawstwem – zajmował się przekładem literatury anglojęzycznej.

Do tej części twórczości naukowej Habilitanta mam trzy uwagi. Po pierwsze szkoda, że nie pokusił się on o ogłoszenie wyników swoich badań w publikacjach o większym zasięgu, przede wszystkim w którymś z czołowych czasopism przekładoznawczych. Po drugie badacz trochę traci z oczu ewolucję teorii i praktyki przekładu, umieszczając w tym samym korpusie przekłady współczesne i dawne. Zaznaczam wszakże, że jest to zarzut niewielkiego kalibru, gdyż w tych pionierskich pracach istotne jest sformułowanie pytania badawczego i pierwsze wyniki, które zachęcają do replikacji i prowadzenia badań w bardziej kontrolowanych próbach (sam Habilitant przekonuje w [1], [2] i [3] jak ważna jest kontrola wyników eksperymentu!), odpowiadających na szczegółowsze pytania, choć trzeba przyznać, że ten postulat częściowo

spełniają prace [7] i [10]. Po trzecie wreszcie szkoda, że Habilitant poprzestaje na prezentacji wykresów. Przy wszystkich swoich zaletach bootstap consensus tree wymaga pewnej interpretacji, która mimo wszystko nie jest do końca jednoznaczna i co więcej zmusza czytelnika do jej samodzielnego przeprowadzenia. Jakkolwiek jest to znakomite narzędzie wizualizacji, ściśle dane liczbowe wzmocniłyby tezy tych prac.

Swoistym wstępem do tego podcyklu jest [5], trzech „detektywistycznych” prac autora, tj. ustalanie wątpliwej atrybucji autorskiej. W pracy tej dr Rybicki dowodzi, że pewien udział w redakcji tłumaczeń Curtina, który przyswoił czytelnikom anglojęzycznym Sienkiewicza, miała jego żona. Warto podkreślić, że w tej pracy Autor wychodzi od hipotezy postawionej na podstawie tradycyjnej (niecyfrowej) filologii (Mikoś 1990). Podobnie praca [9] identyfikuje te części tekstu Virginii Wolf, które zostały przetłumaczone przez dwu tłumaczy.

1.3. [12] i [14] to prace operujące na danych (przynajmniej z punktu widzenia literaturoznawcy) wielkoskalowych. Nie stawiają one żadnej hipotezy, raczej w efektowny sposób pokazują możliwości jakie niesie za sobą taka właśnie skala. Bootstarp Consensus Network grupuje teksty wg gatunków, epoki i stylu. Jak twierdzi Autor – teksty wyłamujące się z tego schematu zwykle wcześniej były identyfikowane przez literaturoznawców, jako niepasujące do schematu epoki czy gatunku; teraz tezy te znalazły potwierdzenie czysto językowe.

1.4. [11] to również praca dotycząca identyfikacji wkładu pojedynczego autora w tekst zbiorowy, w tym wypadku chodzi o utwory spółki autorskiej Josepha Conrada i Forda Madoxa Forda. Badacze używają do tego metody zwanej rolling delta. Polega ona na cięciu tekstu na plasterki, które następnie są porównywane z tekstami pisanymi przez współautorów samodzielnie. Pozwala to na przypisanie (głównego) autora poszczególnym partiom tekstu. Jest to jedno z pierwszych zastosowań tej metody.

1.5. [13] z kolei podejmuje temat śladu płci autora w angielskich tekstach osiemnasto- i wczesnodziewiętnastowiecznych. W tym wypadku to, co daje najsilniejszy i w atrybucji autorskiej najpewniejszy sygnał, czyli najczęstsze słowa zawodzą – identyfikują autorów, ale nie ich płeć. Lepiej działa wypracowana przez Burrowsa metoda polegająca na wychwytywaniu w tekście swoście definiowanych słów kluczowych. Przy okazji Habilitant konstatuje, że leksemy zdradzające płeć autorek najslabiej ujawniają się w tych utworach, które weszły do kanonu literatury.

1.6. Odrębne miejsce w cyklu zajmuje [15]. Praca ta jest na swój sposób podobna jest do badań z nurtu „językowego obrazu świata” – jednym z jej celów jest zbadanie „językowego obrazu” Boga w pismach Edwardsa. W pierwszej części Autorzy omawiają kolokacje wyrazu

God. Następnie kolokacje posłużyły za dyskryminatory do analizy skupień. Wykazują, że te właśnie kolokacje dość dobrze odróżniają teksty homiletyczne od jego pism osobistych (wspomnień, listów itp.) a także pozostałej twórczości.

1.7. [16] jest trzecią pracą z zakresu „filologii detektywistycznej”. Jest ona próbą ustalenia pisarza kryjącego się pod pseudonimem Elena Ferrante. Zagadka ta przyciąga uwagę stylometrów, dość powiedzieć, że artykuł ten jest publikowany w tomie poświęconym temu pytaniu. Habilitant wskazuje na pisarza Domenico Starnonego, jako tego, który się ukrył pod pseudonimem Ferrante, raczej niż jego żonę Anitę Raję (oboje są o to „podejrzewani”).

2. Pozostałe osiągnięcia Habilitanta

Na pozostały dorobek dra Rybickiego składają się 3 publikacje z listy A (w dwu wszakże Habilitant deklaruje marginalny wkład), 1 z listy ERIH, 6 z czasopism spoza tych dwu list oraz 15 rozdziałów w książkach.

Na szczególną uwagę zasługuje praca *Burrowing into Translation: Character Idiolects in Henryk Sienkiewicz's Trilogy and its Two English Translations*, cytowana 36 razy. Mimo pewnego błędu metodologicznego, który wykrył później sam Autor, stała się ona inspiracją dla wielu badaczy. Habilitant zajmuje się w niej „idiolektami” (w istocie stworzonymi przez pisarza, ale przecież zróżnicowanymi) poszczególnych bohaterów Trylogii i ich odpowiednikami w dwu angielskich tłumaczeniach.

Jeśli chodzi o wskaźniki bibliometryczne, to indeks Hirscha wg Google Scholar wyraża się liczbą dwucyfrową, zaś liczba cytowań 615. Są to – jak na humanistykę – wyniki znakomite.

3. Ocena osiągnięć dydaktycznych, współpracy naukowej i popularyzacji nauki

Dr Rybicki jest postacią znaną w środowisku humanistyki cyfrowej. Był członkiem komitetu wykonawczego Association for Literary and Linguistic Computing, a także Association of Digital Humanities Organizations. Wchodzi w skład rady naukowej czasopisma *Digital Scholarship in the Humanities* (dawniej *Literary and Linguistic Computing*). Reprezentuje on Uniwersytet Jagielloński w radzie konsorcjum DARIAH-PL, zaś w DARIAH-EU uczestniczy w pracach jej Rady Naukowej, wreszcie zasiada w komitecie sterującym akcji COST „Distant Reading for European Literary History”. Habilitant jest często zapraszany na rozmaite szkoły humanistyki cyfrowej, co świadczy o tym, że jest cenionym wykładowcą.

4. Konkluzja

W ocenie recenzenta cykl ten spełnia warunki stawiane przez ustawę z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki z późniejszymi zmianami. Warto tu zwrócić uwagę na fakt, że w skład głównego osiągnięcia wchodzi aż 6 artykułów ogłoszonych w czasopismach z listy A. Powszechnie wiadomo, że jest ona dla humanistyki dosyć niereprezentatywna i nie uwzględnia wielu prestiżowych periodyków, niemniej jeśli już czasopismo humanistyczne się na niej znajduje, to jest to niezaprzeczalny dowód jakości. Również liczba cytowań dobrze świadczy o pracy Habilitanta.

Podsumowując stwierdzam, że dorobek naukowy, wyodrębniony cykl publikacji stanowiący osiągnięcie naukowe, osiągnięcia dydaktyczne, współpraca naukowa i popularyzacja nauki dra Jana Rybickiego oceniona zgodnie z Rozporządzeniem Ministra Nauki i Szkolnictwa Wyższego z dnia 1 września 2011 r. (Dz. U nr 196 poz. 1165) *W sprawie kryteriów oceny osiągnięć osoby ubiegającej się o nadanie stopnia doktora habilitowanego* **spełnia w wysokim stopniu wymogi** prawne dotyczące nadania stopnia naukowego doktora habilitowanego wynikające z ustawy z dnia 14 marca 2003 „O stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki” z późniejszymi zmianami.

Rafał L. Górski