

Łódź, 12 stycznia 2020 r.

dr hab. Piotr Pęzik, prof. UŁ
Instytut Anglistyki Uniwersytetu Łódzkiego

Recenzja dorobku naukowego, dydaktycznego i organizacyjnego dr. Yury'a Fedorushkova w związku z postępowaniem o nadanie stopnia naukowego doktora habilitowanego nauk humanistycznych w dyscyplinie językoznawstwo.

Informacje ogólne o habilitancie	1
Ocena dorobku naukowego	1
2.1 Ocena ogółu publikacji	1
2.2 Ocena monografii stanowiącej główne osiągnięcie naukowe	3
2.3 Inne osiągnięcia naukowe	7
Ocena dorobku dydaktycznego i organizacyjnego	8
4. Ocena innych informacji dot. kariery zawodowej	8
5. Wnioski końcowe	9

1. Informacje ogólne o habilitancie

Pan dr Yury Fedorushkov ukończył w 2003 r. studia magisterskie na Wydziale Neofilologii Uniwersytetu im. Adama Mickiewicza. W 2008 r. uzyskał stopień doktora nauk humanistycznych w tej samej jednostce, broniąc rozprawę pt. "Metody ekscpecji neologizmów języka rosyjskiego (na materiale współczesnej prasy rosyjskiej)." Temat rozprawy doktorskiej wpisuje się w obszar korpusowych metod leksykograficznych. W historii zatrudnienia habilitant wykazuje dwie jednostki naukowe: Katedrę Historii Języka Rosyjskiego i Dialektologii Uniwersytetu im. M. Łomonosowa (2008-2009 r.), a następnie stanowisko adiunkta w Instytucie Filologii rosyjskiej UAM.

2. Ocena dorobku naukowego

2.1 Ocena ogółu publikacji

Poza zgłaszaną jako główne osiągnięcie monografią, dr Yury Fedorushkov jest autorem krótkiej monografii pt. "Ekscerpeja wyrazów cudzysłowowych z tekstów prasy rosyjskiej" (2009),

będącej rozszerzeniem rozprawy doktorskiej, “Kieszonkowego słownika piłki nożnej rosyjsko-polskiego i polsko-rosyjskiego” (2018)” oraz współautorem dwóch innych publikacji słownikowych wydanych jeszcze przed doktoratem tj. w 2006 i 2007 r.

Na załączonej przez habilitanta liście “artykułów w czasopismach naukowych” widnieje 13 pozycji. Warto uściślić, iż poza artykułami opublikowanymi rzeczywiście w czasopismach, takich jak *Studia Rossica Posnaniensia*, *Scripta Neophilologica Posnaniensia* (wydawane kolejno przez Instytut i Wydział, w którym zatrudniony jest habilitant), czy *Studia Rossica Gedanensia* pojawiają się tam publikacje z serii monograficznych (Fedorushkov 2010) i tomów konferencyjnych (np. Fedorushkov 2007). Większość z w/w publikacji to prace z obszaru korpusowych analiz leksykograficznych oraz metodologii leksykografii i frazeografii. Pojedyncze publikacje poświęcone są kognitywnym analizom “lingwokoncepcyjnym” pojęć sakralnych (Fedorushkov 2012, 2015). Dwie publikacje podejmują wątek wykorzystania narzędzi do wizualizacji usieciowienia semantycznego i frazeologicznego jednostek leksykalnych na potrzeby glottodydaktyki (Fedorushkov 2014).

Jako rozdział w monografii habilitant wymienia publikację z 2018 r., w której wspólnie z dwoma innymi autorami (Jurkiewicz, Graliński) opisuje wczesną wersję warsztatu ekstrakcji frazemów. Jego pełna wersja opisana została w zgłaszanej jako główne osiągnięcie monografii z r. 2019.

Na liście publikacji, które ukazały się po uzyskaniu przez habilitanta stopnia doktora widnieje niewiele pozycji w czasopismach czy seriach wydawniczych o znacznym zasięgu i renomie. Prace poświęcone leksykograficznemu i frazeograficznemu badaniu korpusowym co prawda opisują oryginalne projekty badawcze habilitanta, często przeprowadzane na dużych zbiorach danych, ale równocześnie wydają się być niewystarczająco solidnie osadzone w głównym nurcie badawczym tego obszaru językoznawstwa. W szczególności, brakuje w nich wyraźnych odwołań i porównań do stanu badań międzynarodowych z tego zakresu. Te dwa czynniki mogły zadecydować o znikomej cytowalności dotychczasowego dorobku habilitanta. Z pobieżnej kwerendy wynika, że zliczający zwykle nadmiarowo cytowania indeks Google Scholar zawiera pojedyncze zewnętrzne odwołania do w/w prac. Możliwe, że takowe cytowania zostały zarejestrowane przez inne systemy bibliograficzne, ale w dostarczonych dokumentach nie znajdujemy ich wykazu.

Pomimo niskiej cytowalności, publikacyjne osiągnięcia habilitanta z obszaru metodologii leksykografii i frazeografii oraz leksykografii praktycznej pozwalają stwierdzić, iż dorobek ten wypełnia w minimalnym zakresie wymogi w postępowaniu o stopień doktora habilitowanego nauk humanistycznych.

2.2 Ocena monografii stanowiącej główne osiągnięcie naukowe

Monografia pt. *Adaptacja warsztatu leksykograficznego do automatyzacji rozpoznawania frazemów w systemie uczącym się (na przykładzie konstrukcji werbo-nominalnych języka rosyjskiego)*. Wydawnictwo Naukowe UAM, Poznań 2019, s. 292, ISBN 978-83-232-3548-4.

Ta licząca 292 strony praca wpisuje się w obszar badań nad prefabrykacją frazeologiczną jako fundamentalnym aspektem produkcji i recepcji każdego języka naturalnego. Większą część opisanych w niej prac stanowią badania i rozważania metodologiczne nad możliwością przygotowania, implementacji a następnie ewaluacji półautomatycznego systemu wydobywania jednostek frazeologicznych (frazemów) z dużych korpusów językowych. Mimo postępów badań w tym obszarze, ekstrakcja frazeologii z dużych zbiorów referencyjnych stanowi bardzo ważną i wciąż niedostatecznie rozwiniętą gałąź korpusologii.

Monografia składa się ze wstępu, czterech rozdziałów zasadniczych i konkluzji.

Rozdział I składa się z trzech odrębnych sekcji opisujących podstawowe aspekty “bazy lingwistycznej” prezentowanego warsztatu leksykograficznego. Pierwsza z nich wprowadza wzorzec leksykalno-gramatyczny, do którego w dalszych etapach pracy ma zostać ograniczony proces identyfikacji frazemów. Wzorcem tym są połączenia otwartej grupy czasowników z wprowadzanymi przez przyimek ‘на’ rzeczownikami w bierniku. Jest to dość arbitralne ograniczenie, które komplikuje ocenę uniwersalności proponowanych procedur dostosowania warsztatu leksykograficznego do zadania ekstrakcji frazeologii. Z drugiej strony, taką decyzję można zinterpretować jako próbę mikrobadań charakterystyki jednego z podstawowych typów predykatywnych związków wyrazowych. Przy ogromie materiału ekstrahowanego z dużych korpusów użytych w pracy takie ograniczenie można uznać za zasadne, pod warunkiem, że nie wyciąga się zbyt daleko idących prognoz w kwestii ekstrapolacji proponowanych rozwiązań na innego typu konstrukcje.

Druga część rozdziału stanowi obszernie omówienie kryteriów identyfikacji jednostek frazematycznych. Autor wysnuwa z dyskusji kryteriów identyfikacji frazemów łącznie 13 wniosków. Wynikiem tych rozważań jest propozycja klucza frazematycznego, czyli zbioru kwalifikatorów decyzyjnych używanych do oznaczania frazemów. Jeżeli te wnioski oraz klucz frazematyczny potraktować jako wytyczne do anotacji frazemów w danych uczących i testowych, to należałoby się zastanowić nad dość fundamentalną kwestią pomiaru tzw. zgodności anotatorów (ang. inter-rater agreement). Innymi słowy, powstaje pytanie, na ile zgodnie i konsekwentnie dwoje lub więcej biegłych w w/w kluczu frazematycznym anotatorów zidentyfikuje reprezentatywną próbę instancji omawianej konstrukcji werbo-nominalnej jako instancje frazematyczne albo swobodne. Można przypuszczać, iż duża złożoność wytycznych mogłaby taką zgodność zaniżyć. Nawet mierzona w różnym czasie zgodność anotatora z samym sobą (ang. ‘intra-rater agreement’) może ulegać wahaniom przy bardzo rozbudowanych wytycznych. Określenie wskaźnika zgodności anotatorów jest o tyle istotne, że przy jego skrajnie niskich wartościach (np. 0,2 -- 0-3 wg. wskaźnika Kappy Cohena) “uczenie”

klasyfikatora w oparciu o takie dane nie jest zasadne. Z pewnością dodanie pomiarów zgodności anotacji przybliżyłoby tę część pracy do standardów ewaluacyjnych językoznawstwa komputerowego.

Ostatnia część pierwszego rozdziału opisuje zasoby korpusowe użyte do opracowania zbiorów uczących i testowych w procesie automatycznej ekstrakcji frazemów. Trudno ocenić ich całkowitą wielkość ze względu na niejednorodną charakterystykę ilościową (jednostki wielkości są teksty zdania/ słowa), ale jest oczywiste, że zdecydowana ich większość (poza Narodowym Korpusem Języka Rosyjskiego) jest o kilka rzędów wielkości mniejsza od danych ze zbioru Common Crawl, czyli dużego zrzutu rosyjskojęzycznych witryn internetowych, co oznacza, że ten ostatni zasób można uznać za dominujący korpus źródłowy. Zaletą CommonCrawl jest niewątpliwie wielkość zawartych w nim zasobów tekstowych. Wśród wad należy wymienić wysokie zaszumienie wynikające z automatycznych procedur zbierania danych, niską reprezentację pewnych gatunków i typów funkcjonalnych tekstów (np. mowa, źródła zamknięte itp.) oraz w zasadzie nieznaną strukturę bibliograficzną CommonCrawl jako korpusu.

Rozdział 2 opisuje procedurę selekcji konstrukcji werbalno-nominalnych o potencjalnym statusie prefabrykatu z pozyskanych korpusów j. rosyjskiego. Ekscerpcja konstrukcji opiera się na identyfikacji pozycyjnych struktur morfosyntaktycznych, które stanowią przybliżenie relacyjnych konstrukcji składniowych, a następnie na dalszej ich selekcji m.in. wg kryteriów leksykalnych. Techniki używane do preselekcji zbitek słownych autor określa mianem “filtrów ankodowo-regeksowych”. W zasadzie chodzi tu o połączenie prostej analizy morfoskładniowej (opartej na bezkontekstowej metodzie słownikowej) z wyrażeniami regularnymi. Informatyczny aspekt wydobywania takich wzorców jest więc dość banalny, ale ich językoznawcza interpretacja stanowi duże wyzwanie metodologiczne.

Zorientowanego w tematyce korpusowej czytelnika może dziwić brak odniesień do jakichkolwiek prac nt. znanych technik identyfikacji jednostek frazeologicznych z danych korpusowych w tym i kolejnych rozdziałach. Wiele tego typu metod począwszy od analizy konkordancji i prostych kolokacji dwuwyrzowych po bardziej zaawansowane techniki pozycyjne i relacyjne ekstrakcji frazemów weszło do podstawowego zestawu narzędzi leksykografów epoki ‘korpusowej’. Wyszukiwarki korpusowe z podstawową obsługą standardu CQP (Corpus Query Language) umożliwiają wydobycie i agregację tysięcy instancji wzorca konstrukcji złożonych z dowolnych sekwencji części mowy, pasujących do zdefiniowanych w zapytaniach wyrażen regularnych¹. Wiele z takich systemów posiada wbudowane moduły do identyfikacji potencjalnych kolokacji i innego typu jednostek frazeologicznych. Jedynym pozornym odniesieniem do stanu prac nad ekstrakcją frazeologii w Rozdziale 2 wydaje się być

¹ Na przykład zapytanie o sekwencje czasownika, przyimka “na” i rzeczownika w bierniku w zrównoważonej części NKJP zwraca ponad 319 tys. wystąpień tego typu struktur. Zob. zapytanie <pos=v.*> na <pos=n.*:acc.*> w <http://pelcra.clarin-pl.eu/NKJP>. Podobne możliwości uzyskania wstępnego “surowca” i to z ujednoznaczonych morfologicznie danych (m.in. rosyjskich) oferuje silnik SketchEngine.

krótka wzmianka o modelowaniu tematycznym, które jest techniką identyfikacji i etykietowania grup tematycznych w korpusach, a nie metodą ekstrakcji kolokacji.

Rozdział 3 omawia w bardzo ogólnym zarysie ewaluowane przez autora podejście do wykrywania frazemów (SARWKF). Z krótkiego opisu wynika, że w procesie klasyfikacji wykorzystywane są wektorowe reprezentacje słów (uzyskane przez trening BOW/ Skip-gram płytkiej sieci neuronowej Word2vec z jedną warstwą ukrytą). Reprezentacje te wyliczane są dla słów występujących w oknie kontekstowym frazemu a następnie, jak można zgadywać, uśredniane lub w inny sposób łączone z wektorowymi reprezentacjami słów wchodzących w skład samego frazemu. Powszechnie znane jest wykorzystanie tego typu reprezentacji do klasyfikacji krótkich fragmentów tekstów, wykrywania wydźwięku, czy też dopasowania odpowiedzi z bazy wiedzy (zob. Joulin et al. 2016). Podstawową ich zaletą w uczeniu nadzorowanym jest transfer reprezentacji znaczenia i cech gramatycznych słów pozyskanych z korpusu referencyjnego. W takim podejściu reprezentacje słów przestają być dyskretnymi symbolami, a stają się porównywalnymi tensorami, dzięki czemu możliwe jest obliczenie np. podobieństwa dystrybucyjnego między całkowicie różnymi ortograficznie synonimami. Z punktu widzenia głównego tematu osiągnięcia, tzn. “dostosowania warsztatu leksykograficznego” do systemu uczenia maszynowego mającego wykrywać frazemy, w Rozdziale 3 odczuwalny jest brak uzasadnienia operacjonalizacji reprodukcji frazematycznej w oparciu o wektorowe reprezentacje słów. Autor nie przedstawia intuicji stojącej za doбором zestawu tego typu zmiennych objaśniających. Nie wyjaśniono jak wektorowe reprezentacje słowoform mogą mieć przełożenie na wykrywanie statusu frazematycznego danej kombinacji wyrazowej. Jeżeli przyjąć, że mamy tu do czynienia z klasyfikacją binarną (frazem/ nie-frazem), w której zmiennymi objaśniającymi są uśrednione wektory słów z kontekstu frazemu, a zmienną celu jest klasa czy też regresyjny wskaźnik frazematyczności, to trudno z dostępnego opisu wywnioskować, czego taki klasyfikator miałby się nauczyć z samych wektorowych reprezentacji słów jako zmiennych objaśniających. Na str. 213 znajduje się jedynie wzmianka o długości kontekstów jako ‘przykładowej’ cesze w procesie trenowania klasyfikatora SGD, ale jest oczywiste, że długość okna kontekstowego nie może być kluczowym aspektem identyfikacji frazemów. Ponieważ działanie algorytmu klasyfikacji nie zostało intuicyjnie wyjaśnione, trudno jest oceniać, jaki wpływ na jego wyniki ma proponowana procedura “adaptacji warsztatu leksykograficznego”, a to niestety podważa w pewnym stopniu główną konstrukcję wyводу. Innymi słowy, nie można w sposób kontrolowany dostosowywać danych treningowych do algorytmu, którego działanie nie jest choćby intuicyjnie znane. Dla porównania, znane próby operacjonalizacji statusu frazeologicznego połączeń wyrazowych jako zależności między wektorowymi reprezentacjami frazemu a reprezentacjami jego składowych opisuje m.in. Cordeiro (2017). Kryterium idiomatyczności opisywane w tej ostatniej pracy jest intuicyjne i udaje się je zweryfikować na frazematkach zawierających wyrazy synsemantyczne.

Z punktu widzenia ewaluacji systemu przydatne byłoby wyjaśnienie, czy opisywany klasyfikator nie ‘przeucza się’ (ang. over-fits) konkretnych kombinacji słów lub ich synonimów,

odrzucając jednocześnie frazemy o charakterystyce niezaświadczonej bezpośrednio lub pośrednio, np. synonimicznie w zbiorze uczącym. Brakuje tu jasnej informacji o proporcjach liczebności klas frazemów do nie-frazemów w danych uczących i testowych, tzn. o średniej proporcji frazemów do nie-frazemów w korpusach źródłowych, zbiorze treningowym i zbiorze testowym. W przypadku klasyfikacji binarnej, z którą mamy tu do czynienia istnieje ryzyko wyuczenia tzw. prawdopodobieństw jednej, bardziej prawdopodobnej klasy, tzn. przyjęcia przez klasyfikator strategii oznaczania stałej proporcji instancji testowych jako przypadki z danej klasy. Ze względu na brak szczegółowych informacji statystycznych o zbiorze uczącym i testowym trudno jest spekulować na temat istotności podawanej przez autora wartości metryki dokładności na poziomie 72%². Można zakładać że chodzi tutaj o miarę "accuracy" (stosunek klasyfikacji poprawnych do sumy klasyfikacji poprawnych i niepoprawnych). Miara ta nie uwzględnia różnic w dystrybucji i zbilansowaniu rozpatrywanych klas. Autor pisze co prawda, że system klasyfikacji osiągnął dokładność 'zdecydowanie statystycznie istotną', ale tego twierdzenia nie uzasadniono szczegółowo. Nie jest jasne, czy jest to jedynie odczucie autora, czy może zastosowano tu jakikolwiek test istotności uwzględniający w/w aspekty ewaluacji.

Rozdział 4 opisuje zbiór manualnych procedur ewaluacyjnych, ukierunkowanych na wykrywanie skonwencjonalizowanych konstrukcji werbalno-nominalnych, które nie wystąpiły w danych uczących. Taka dodatkowa weryfikacja danych wyjściowych jest nieodzownym aspektem procesu ekscerpacji frazemów, pozwalającym na ocenę przydatności całego systemu w zastosowaniach leksykograficznych. Należy zauważyć, że procedura weryfikacji "jednostek bingo" wydaje się uwzględniać jedynie jednostki prawdziwie i fałszywie pozytywne (ang. true/false negatives) z pominięciem jednostek fałszywie negatywnych (ang. false negatives). Jej wyniki można więc interpretować głównie jako analizę precyzji, a nie pokrycia ekstrakcji.

W konkluzji monografii autor wyraża przekonanie, że przedstawione procedury adaptacją warsztatu leksykalnego do zadania ekscerpacji frazemów można uogólnić do konstrukcji innego typu. Trudno jest określić na ile skalowalne są zaproponowane rozwiązania ze względu na ograniczenie badań do jednego schematu tj. występujących po sobie bezpośrednio trzech segmentów wyrazowych, tworzących sekwencję czasowników połączonych przyimkiem 'na' z rzeczownikami rodzaju żeńskiego w liczbie poj., zakończone dwoma morfemami fleksyjnymi.

Pomimo dość jasnego sformułowania ogólnych celów i efektów badań opisanych w monografii, jej lekturę utrudnia tendencja autora do tworzenia nowych, często złożonych lub enigmatycznych terminów na określenie znanych już pod innymi, bardziej skonwencjonalizowanymi nazwami pojęć, technik i elementów warsztatu korpusowego i frazeograficznego. Na przykład 'Surowiec R' to chyba nic innego jak 'zbiór uczący' (ang. training set), 'Odtwarzalność Kontekstowa' to prosta częstość zbitki w korpusie, a 'Megalista Faktu Rejestracji' to wystąpienie potencjalnego frazemu w słownikach. Z kolei 'System Automatyzacji Rozpoznawania Wielowyrazowców w

² Na str. 208 podawana jest dokładność 0,665% -- zapewne chodziło o 66,5%.

Kluczu Frazematycznym' (SARWKF) to po prostu system wykrywania frazemów. Wydaje się, że pełną charakterystykę proponowanego klasyfikatora można było wprowadzić jednorazowo, a następnie następnie posługiwać się prostszym jego określeniem. Skomplikowane innowacje terminologiczne zmuszają czytelnika do częstego zaglądania do glosariusza pojęć.

Niezależnie od wymienionych powyżej zastrzeżeń metodologicznych, za istotne elementy znaczącego wkładu habilitanta w rozwój dyscypliny językoznawstwa można uznać następujące aspekty głównego osiągnięcia:

- a) Opracowanie opisanego propozycji klucza frazematycznego. Co prawda jego praktyczne wykorzystanie w innych pracach z obszaru frazeografii korpusowej, wymaga walidacji zgodnej ze standardami NLP, ale propozycję autora można traktować jako cenne źródło zweryfikowanej jakościowo i ilościowo koncepcji operacjonalizacji reprodukcji frazeologicznej.
- b) Opracowanie i przeprowadzenie eksperymentów na dużych zbiorach danych. Chodzi tu zarówno o użyte źródłowe dane korpusowe, jak też o bardzo dużą liczbę wyekstrahowanych z nich i przeanalizowanych instancji potencjalnych frazemów w fazie przygotowania danych uczących i testowych.
- c) Opracowanie i przeprowadzenie opisaną w rozdziale czwartym metodyki weryfikacji wyników klasyfikacji frazemów wydobywanych z danych korpusowych.

2.3 Inne osiągnięcia naukowe

Lista staży i wizyt dr Fedorushkova w instytucjach naukowych nie jest obszerna. Habilitant wymienia pobyt w dwóch instytucjach w ciągu jedenastu lat od momentu uzyskania stopnia doktora:

- 1) konsultacje i kwerendy biblioteczne w Uniwersytecie Opolskim, których charakter trudno do końca ocenić na podstawie dostarczonego opisu oraz
- 2) dwutygodniowy pobyt w Uniwersytecie Amsterdamskim, podczas którego m.in. wygłosił wykład dla studentów tamtejszej slawistyki.

Od r. 2007 dr Yury Fedorushkov wziął udział w 15 konferencjach organizowanych w większości przez ośrodki w Polsce. Wygłoszone na nich referaty dotyczyły głównie zastosowania metod korpusowych do ekscerpcji jednostek wielowyrazowych z korpusów.

Z autoreferatu wynika, iż jedynym uzyskanym przez habilitanta grantem był realizowany przez 7 miesięcy projekt finansowany przez Centrum Polsko-Rosyjskiego Dialogu i Porozumienia.

Projekt ten zakończył się bardzo praktycznym wynikiem w postaci kieszonkowego słownika dla osób podróżujących na Mistrzostwa Świata w Piłce Nożnej.

Z drugiej strony, nie jest jasna intencja podawania nieudanej próby uzyskania grantu NCN w tej sekcji autoreferatu. Informacja ta może sugerować, że habilitant podjął tylko jedną próbę uzyskania tego typu grantu od r. 2009. Być może chodzi o program tzw. 'grantów na granty', ale nie jest to jasno sprecyzowane.

3. Ocena dorobku dydaktycznego i organizacyjnego

W latach 2008-2009 habilitant pracował na stanowisku starszego wykładowcy w Katedrze Historii Języka Rosyjskiego i Dialektologii Uniwersytetu im. M. Łomonosowa. Od 2009 r. pracuje jako adiunkt w Instytucie Filologii rosyjskiej UAM. Wypromował dość liczną grupę 28 licencjatów i dwóch magistrantów, co dobrze świadczy o jego pracy i doświadczeniach dydaktycznych.

Działalność organizacyjną i popularyzatorską dr. Fedorushkova należy uznać za bardzo umiarkowaną. Nie jest on członkiem krajowych lub międzynarodowych towarzystw naukowych, redakcji czasopism naukowych, z autoreferatu nie wynika by recenzował prace czy projekty badawcze. Z informacji zawartych w autoreferacie nie wynika, by organizował lub współorganizował konferencje o charakterze ściśle naukowym. Zorganizował za to warsztaty dla nauczycieli oraz dwukrotnie konkurs recytatorski. Wziął także udział w konkursie na najlepsze tłumaczenie poezji Wisławy Szymborskiej.

4. Ocena innych informacji dot. kariery zawodowej

W sekcji "Leksykografia NLP" autoreferatu habilitant rozwija skrót "NLP" jako "programowanie neurolingwistyczne", czyli jedną z pseudonaukowych technik psychoterapeutycznych. Zapewne zamiarem habilitanta było powołanie się na ang. termin "Natural Language Processing" (przetwarzanie języka naturalnego), tj. popularną w ostatnich latach dziedzinę z pogranicza informatyki, językoznawstwa komputerowego i sztucznej inteligencji. Być może doszło tu do jakiejś niefortunnej amalgamacji pojęciowej tzw. sieci neuronowych jako rodziny algorytmów uczenia maszynowego, programowania i przetwarzania języka naturalnego.

5. Wnioski końcowe

Habilitant w istotny sposób poszerzył swój dorobek publikacyjny od czasu uzyskania stopnia doktora w 2009 r. Jego działalność dydaktyczną w tym okresie można uznać za w pełni zadowalającą. Zgłaszana jako główne osiągnięcie w przewodzie habilitacyjnym monografia posiada wymienione szczegółowo powyżej walory decydujące o znacznym wkładzie w rozwój frazeografii korpusowej jako ważnego obszaru językoznawstwa.

Biorąc pod uwagę powyższe należy stwierdzić, że dr Yury Fedorushov spełnia określone w Ustawie z dnia 3 lipca 2018 r. kryteria uzyskania stopnia doktora habilitowanego nauk humanistycznych w dyscyplinie językoznawstwa. W związku z tym wnioskuję do Komisji Habilitacyjnej o dopuszczenie pana dr. Yury'ego Fedorushov do dalszych etapów przewodu habilitacyjnego.

Piotr Pgzik