

Autoreferat

1. Imię i nazwisko.

Monika Teresa Czerepowicka

2. Posiadane dyplomy, stopnie naukowe lub artystyczne – z podaniem podmiotu nadającego stopień, roku ich uzyskania oraz tytułu rozprawy doktorskiej:

- a) magister filologii polskiej – Uniwersytet Warszawski, Wydział Polonistyki, Warszawa 2000.

Praca magisterska „Weryfikacja empiryczna składniowej bazy danych czasowników i quasi-czasowników polskich: nie posiadać się – śledzić”, napisana pod kierunkiem prof. dr. hab. Marka Świdzińskiego, zrecenzowana przez dr Dorotę Kopcińską.

- b) doktor nauk humanistycznych w zakresie językoznawstwa – Uniwersytet Warmińsko-Mazurski w Olsztynie, Wydział Humanistyczny, Olsztyn 2006.

Rozprawa doktorska napisana pod kierunkiem prof. dr. hab. Zygmunta Saloniego pt. „Opis lingwistyczny wyrażen niestandardowych składniowo typu *na lewo, do dziś, po trochu, na zawsze* we współczesnym języku polskim”, recenzowana przez prof. dr. hab. Włodzimierza Gruszczyńskiego oraz prof. dr. hab. Aleksandra Kiklewicza.

3. Informacja o dotychczasowym zatrudnieniu w jednostkach naukowych lub artystycznych.

Od 1.04.2006 r. do dziś – Uniwersytet Warmińsko-Mazurski w Olsztynie, Wydział Humanistyczny, stanowisko: adiunkt.

4. Omówienie osiągnięć, o których mowa w art. 219 ust. 1 pkt. 2 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2021 r. poz. 478 z późn. zm.). Omówienie to winno dotyczyć merytorycznego ujęcia przedmiotowych osiągnięć, jak i w sposób precyzyjny określać indywidualny wkład w ich powstanie, w przypadku, gdy dane osiągnięcie jest dziełem współautorskim, z uwzględnieniem możliwości wskazywania dorobku z okresu całej kariery zawodowej.

Moje zainteresowania naukowe koncentrują się wokół trzech obszarów: zdyscyplinowanego opisu gramatycznego polszczyzny, leksykografii elektronicznej oraz wykorzystania korpusów i innych narzędzi elektronicznych w praktyce językoznawczej i pozajęzykoznawczej. Najwcześniejsze doświadczenia badawcze zdobywałam podczas anotacji morfoskładniowej korpusów tekstów oraz prac leksykograficznych, stąd dostrzegam wartość, ale i trudności prowadzenia językoznawczych badań empirycznych. Mój dorobek naukowy obejmuje

artykuły dotyczące kwestii ogólnych, podstawowych, związanych z rejestracją jednostek leksykalnych, opisem gramatycznym, warsztatem leksykograficznym, a także szczegółowych rozwiązań słownikowych. Byłam współpracownikiem zespołu pracującego nad „Słownikiem gramatycznym języka polskiego” (sgjp.pl). Jestem współautorką dwóch słowników fleksyjnych wielowyrazowych jednostek leksykalnych: „Słownika elektronicznego jednostek frazeologicznych” (SEJF) oraz „Słownika paradygmatów frazeologizmów werbalnych” (VERBEL).

Zarysowane powyżej doświadczenia naukowe i leksykograficzne przyczyniły się do rozwinięcia badań nad problemami opisu jednostek wielowyrazowych, czego efektem jest moje główne osiągnięcie naukowe, tj. monografia:

„Wielowyrazowe jednostki leksykalne w polszczyźnie. Studium leksykograficzne”, Olsztyn 2023, ss. 266, Wyd. UWM, ISBN 978-83-8100-380-3.

Niemal do końca XX w. praktyka leksykograficzna podążała za znanym sobie i utrwalonym przez strukturalizm postrzeganiem jednostek języka jako obiektów ciągłych, mających postać pojedynczych wyrazów. Konsekwencją marginalizowania jednostek wielowyrazowych jest nikłe zainteresowanie nimi w badaniach językoznawczych, przekładające się na brak opracowań poświęconych ich szczegółowemu opisowi gramatycznemu. Dopiero śmiałe tezy Igora Mielczuka oraz Andrzeja Bogusławskiego zapoczątkowały zmiany w postrzeganiu jednostek leksykalnych. Dlatego tak konieczne jest usystematyzowanie dostępnej wiedzy na temat jednostek nieciągłych. Cele mojej monografii są więc dwojakiego rodzaju. Po pierwsze, jest nim analiza opisu leksykograficznego wielowyrazowych jednostek leksykalnych, zwłaszcza w zakresie ich cech gramatycznych. Adnotacja o fleksji jednostek nieciągłych w tradycyjnych słownikach pojawia się rzadko, zazwyczaj jako uzupełnienie informacji zasadniczej, czyli objaśnienia znaczenia. Pośrednio zakłada się, że związki frazeologiczne powielają odmianę członu głównego po uwzględnieniu regularnych uzgodnień systemowych, jak dziedziczenie wartości fleksyjnych wyrazu określającego po określanym (*biały kruk, białego kruka*). Eksplicytna informacja fleksyjna zarówno jednostek ciągłych, jak i wielosegmentowych jest warunkiem koniecznym kompletnego opisu leksykograficznego. Dlatego tak istotne jest przeprowadzenie systematycznych badań nad polską leksykografią jednojęzyczną. Zbadanie, w jaki sposób postrzega się i definiuje jednostki wielowyrazowe oraz które rozwiązania słownikowe, wynikające z refleksji językoznawczej, sprzyjają jasności, konsekwencji opisu, stanowią o rozwoju warsztatu leksykograficznego, wychodzą naprzeciw potrzebom użytkowników, mogą przyczynić się do spopularyzowania słowników i wiedzy językoznawczej.

Dzięki analizie słowników uwidaczniają się sposoby notacji, opisu i pośrednio interpretacji jednostek wielowyrazowych. Natomiast dzięki włączeniu w obszar badań dzieł dawnych

i współczesnych rysują się zarówno różnice, jak i podobieństwa w traktowaniu połączeń wyrazowych w perspektywie synchronicznej i diachronicznej. Płynące z badań wnioski wskazują na aktualny stan leksykografii, uzupełniają i systematyzują teorię leksykograficzną dotyczącą wielowyrzowych jednostek leksykalnych. Wyznaczają także zadania współczesnej leksykografii, a formułowane postulaty stwarzają możliwości zastosowania w praktyce słownikarskiej.

Drugim celem monografii jest wskazanie jednego z możliwych sposobów opisu fleksji jednostek wielowyrzowych, stosowanego w wybranych słownikach elektronicznych – SEJF i VERBEL. Zastosowana w nich koncepcja opisu leksykograficznego bazuje na pojęciu *fleksemu*, zaproponowanym przez Janusza S. Bienia i Zygmunta Saloniego. Wieloetapowy opis warstwy morfologicznej języka i wynikające z niego pojęcia zostały dostosowane do opisu jednostek nieciągłych, z czego szczegółowo zdaję sprawę w monografii.

Zmiana technologiczna, dzięki której powstają słowniki elektroniczne, stwarza oczywistą konieczność włączenia w obszar badań leksykografii elektronicznej, co z kolei powoduje uwzględnienie w nich perspektywy bardziej technicznej – maszynowego przetwarzania polszczyzny. Monografia ma więc charakter interdyscyplinarny, sytuuje się na przecięciu językoznawstwa, leksykografii i inżynierii lingwistycznej.

Książka składa się z siedmiu rozdziałów, wstępu, zakończenia i bibliografii. W rozdziale drugim omawiam pojęcie leksykografii, jej rolę, znaczenie, a także związaną z nią współcześnie terminologię i perspektywy badawcze. W tej pracy opowiadam się za jej szeroką perspektywą badawczą i praktyczną. Do zadań leksykografii zaliczam zarówno tworzenie baz danych, jak i dzieł powstających na ich podstawie, traktujących o języku (słowniki) i świecie (encyklopedie). Rozwój technik gromadzenia i publikacji informacji każe postrzegać ją mniej jako subdyscyplinę językoznawstwa, a bardziej jako niezależną gałąź wiedzy, czerpiącą z innych dziedzin – językoznawstwa (opracowanie danych językowych), informatyki i teorii informacji (sposoby gromadzenia, publikacji wiedzy w bazach danych, słownikach i encyklopediach), jak i szczegółowych dyscyplin (koniecznych w wypadku encyklopedii). W swoich badaniach koncentruję się jednak na słownikach języka polskiego i związkach językoznawstwa oraz informatyki, realizujących się w postaci leksykografii elektronicznej.

Kluczowym pojęciem w monografii jest wielowyrzowa jednostka języka (ang. *multi-word expressions, multi-word unit*), które stosuję wymiennie z następującymi: *wielowyrzowa / wielosegmentowa jednostka leksykalna, wielowyrzowa / wielosegmentowa jednostka języka, jednostka frazeologiczna i frazeologizm*. Sposoby rozumienia wielowyrzowej jednostki leksykalnej w literaturze przedmiotu szczegółowo dyskutuję w rozdziale trzecim. Przedmiotem opisu są modele teoretyczne, które znajdują zastosowanie w przygotowywaniu

słowników. Wśród lingwistycznych znalazły się koncepcje następujących badaczy: Stanisława Skorupki, Andrzeja Marii Lewickiego, Andrzeja Bogusławskiego, Macieja Grochowskiego, Wojciecha Chlebdy i Piotra Müldnera-Nieckowskiego. Za sprawą leksykografii elektronicznej oprócz koncepcji czysto językoznawczych uwzględniam także propozycje bardziej techniczne – analizujące zagadnienie z punktu widzenia automatycznego przetwarzania języka, tj. model Sens ↔ Tekst Igora Mielczuka, Jurija Apresjana, Aleksandra Żółkowskiego, poziomy frazeologiczności Timoty’ego Baldwina i Su Nam Kima oraz skalę frazeologiczności Michała Woźniaka. Prezentowane teorie lingwistyczne ukazują, że frazeologiczność jest zjawiskiem złożonym, z trudnością dającym się zamknąć w ramach dychotomicznych podziałów (jest – nie jest jednostką wielosegmentową). Trzeba podkreślić, że nie ma jednej uniwersalnej definicji wielowyrzawowej jednostki języka. Proponowane przez badaczy kryteria frazeologiczności zależą od celu prowadzonych badań, co, przekładając na leksykografię, oznacza rodzaj opracowywanego słownika i jego przeznaczenie.

Analiza różnych modeli teoretycznych doprowadziła do zaproponowania własnej definicji jednostki wielowyrzawowej. Do jej cech definicyjnych zaliczam:

- 1) podzielność formalną – jednostka wielowyrzawowa składa się z co najmniej dwóch słów tekstowych;
- 2) nieregularność znaczeniową – znaczenie jednostki nie jest bezpośrednią sumą znaczeń jej składników;
- 3) nieciągłość – komponenty nie muszą bezpośrednio po sobie występować, mogą być rozdzielone innymi słowami, nienależącymi do jednostki;
- 4) stałość dystrybucyjną – przynajmniej jeden ze składników występuje w ściśle określonym sąsiedztwie leksykalnym, poza połączeniem nie występuje;
- 5) nieregularność składniową – funkcja składniowa jednostki nie wynika wprost z relacji syntaktycznych zachodzących pomiędzy członami.

Dwie pierwsze cechy uznać należy za podstawowe, dające się wskazać we wszystkich jednostkach wielowyrzawowych. Natomiast dalsze cechy (3–5) – jako wymienne: wystarczy spełnić jedną z nich, by dane połączenie odróżnić od konstrukcji i zaliczyć do wielowyrzawowych jednostek języka.

W rozdziale czwartym przedstawiam zakres podjętych badań leksykograficznych. Dyskutuję w nim sposób prezentacji i opisu wielosegmentowych jednostek języka w wybranych polskich słownikach – ogólnych i frazeologicznych. Materiał sięga XIX w., uwzględnia dwa wieki sztuki słownikarskiej, które są szczególnie istotne tak pod względem jej rozwoju ilościowego, jak i jakościowego. W doborze kierowałam się zasięgiem słownika, jego rozpoznawalnością, znaczeniem w rozwoju polskiej leksykografii oraz stosowanymi rozwiązaniami

warsztatowymi dotyczącymi opisu jednostek wielowyrazowych. Badania rozpoczynam od analizy pierwszego jednojęzycznego słownika polszczyzny – autorstwa Samuela Bogumiła Lindego¹, kończę na współcześnie powstającym *Wielkim słowniku języka polskiego PAN*. Celem przeglądu jest ukazanie sposobów opisu wielowyrazowych jednostek języka w wybranych polskich słownikach, tak ogólnych, jak i specjalistycznych – frazeologicznych. Przegląd ten pośrednio pokazuje także sposób kształtowania się refleksji językoznawczej na temat jednostek nieciągłych oraz zajmującego się nimi działu językoznawstwa. Początkowo jednostki te interpretowane są jako nietypowe przykłady łączliwości poszczególnych leksemów, z czasem zaczynają funkcjonować na specjalnych prawach – nie są opisywane jako zwykłe egzemplifikacje, lecz mniej lub bardziej stałe połączenia, by w końcu zyskać status niezależnych jednostek języka i równie wnikliwie jako one opisywanych. Ewolucja poglądów dotyczących jednostek nieciągłych idzie w parze z rozwojem warsztatu leksykograficznego. Frazeologizmy zaczyna się opatrywać informacją słownikową (np. w postaci kwalifikatorów), która z czasem staje się wyspecjalizowanym opisem gramatycznym i szerzej – leksykograficznym.

W rozdziale piątym podejmuję problematykę opisu leksykograficznego, uwzględniającego szczegółowy zakres informacji fleksyjnej na podstawie wybranej koncepcji opisu morfologicznego polszczyzny. W związku z równoczesnym zastosowaniem perspektywy funkcjonalnej i formalnej w opisie jednostek wielosegmentowych rozszerzam pojęcie *fleksemu* wprowadzone do polskiego językoznawstwa przez Janusza S. Bienia i Zygmunta Saloniego. Dyskusja oraz motywacja takiego rozwiązania znajduje się w rozdziale piątym. W rozdziale szóstym analizuję sposób opisu fleksji jednostek wielosegmentowych w dwóch słownikach fleksyjnych – SEJF i VERBEL.

Słownik elektroniczny jednostek frazeologicznych (SEJF) jest słownikiem fleksyjnym zawierającym około 3 tys. jednostek wielowyrazowych, których człon główny należy do klasy leksemów imiennych. Trzon zbioru stanowią wyrażenia sklasyfikowane jako rzeczownikowe, a także przymiotnikowe i przysłówkowe oraz sporadycznie inne nieodmienne. Słownik VERBEL powstał w wyniku kontynuacji prac nad opisem morfoskładniowym jednostek wielowyrazowych, jego zawartość stanowi ponad 5 tys. werbalnych jednostek wielowyrazowych, zróżnicowanych zarówno pod względem składni wewnętrznej, jak i zewnętrznej. Zannotowane jednostki należą do jednego z trzech typów składniowych: zwrotów grupy werbalnej, zwrotów związku głównego (Lewicki 1986) oraz zwrotów z paradygmatem analogicznym do czasowników niewłaściwych lub wtórnie niewłaściwych.

¹ Wskazując na jednojęzyczny charakter dzieła Lindego, mam świadomość, że słownik ten jest także źródłem licznych ekwiwalentów w dialektach słowiańskich oraz językach obcych polszczyźnie, co z pewnej perspektywy czyni z niego słownik wielojęzyczny.

Oba słowniki – SEJF i VERBEL – metodologicznie oparte są na tych samych podstawach, dążą do jak najbardziej konsekwentnego, rygorystycznego opisu gramatycznego jednostek leksykalnych, wszystkie ich formy podając *in extenso*.

Szczegółowe informacje dotyczące budowy i odmiany jednostek wielowyrzowych w przedstawionych powyżej słownikach fleksyjnych umożliwiają podjęcie próby systematyzacji materiału leksykalnego, która znajduje się w rozdziale siódmym. Podstawę systematyzacji wyrażen stanowią następujące cechy jednostek:

- przynależność do określonej klasy gramatycznej,
- budowa (liczba składników i ich właściwości),
- kompletność paradygmatu.

Zbiór wielowyrzowych jednostek imiennych tworzą głównie hasła rzeczownikowe, składające się z dwóch lub trzech komponentów leksykalnych. Ich paradygmaty fleksyjne zazwyczaj mają regularny charakter, powielają odmianę głównego składnika. Ewentualne ograniczenia dotyczą odmiany przez liczbę. Hasła przysłówkowe i przymiotnikowe, stanowiące margines materiału leksykalnego, w głównej mierze także mają budowę dwu- lub trójdzielną, choć w tych klasach jednostek częściej występują także połączenie składające się z większej liczby komponentów.

Systematyzacja formalno-fleksyjna zwrotów bierze pod uwagę kilka kryteriów:

- klasę morfologiczną komponentu głównego,
- wartość aspektu,
- ilość i porządek linearny składników,
- typ wypełnienia pozycji podmiotu,
- kompletność paradygmatu.

Oprócz wymienionych powyżej w systematyzacji uwzględnia się również inne cechy zwrotów, jak obecność cząstki *się* oraz negacji.

W osobne zestawy pogrupowano jednostki reprezentujące odmienne cechy morfologiczne i składniowe:

1. zwroty ufundowane na czasowniku właściwym przyłączające podmiot osobowy,
2. zwroty ufundowane na czasowniku właściwym przyłączające podmiot nieosobowy,
3. zwroty ufundowane na czasowniku niewłaściwym,
4. zwroty ufundowane na czasowniku wtórnie niewłaściwym,
5. zwroty grupy głównej.

Zbiór wielowyrazowych jednostek czasownikowych w głównej mierze tworzą zwroty grupy werbalnej dysponujące pełnym paradygmatem fleksyjnym, dziedziczonym po komponencie czasownikowym. Ograniczenia fleksyjne, choć stosunkowo nieczęste, mają zróżnicowany charakter. Mogą wynikać z cech składniowych, morfologicznych, pragmatycznych lub semantycznych jednostek. Pierwsze wiążą się z typem wypełnienia pozycji podmiotu – rzeczowniki nieosobowe ograniczają możliwości odmiany przez podstawowe kategorie fleksyjne czasowników: osobę i liczbę. Ograniczenia morfologiczne, jak np. brak bezosobnika, są z reguły dziedziczone po komponencie czasownikowym. Z kolei braki w paradygmatach wynikające ze znaczenia lub sposobu użycia zwrotów wydają się najmniej przewidywalne, choć część z nich wyznacza pragmatyka trybu rozkazującego. Zwroty grupy głównej oraz oparte na czasowniku niewłaściwym lub wtórnie niewłaściwym tworzą nieliczny podzbiór, który w całości (z definicji) charakteryzuję jako defektywny pod względem fleksyjnym.

Systematyzacja wzorów strukturalno-paradygmatycznych frazeologizmów uwidacznia pewną prawidłowość. Zdecydowana większość jednostek reprezentuje zaledwie kilka wzorów strukturalno-fleksyjnych, jednocześnie stosunkowo duża liczba wzorów jest przypisana do pojedynczych jednostek leksykalnych. Innymi słowy, mimo że ograniczenia fleksji należą do relatywnie małego zbioru jednostek (wśród zwrotów 1234 z 5076), to przekładają się na dużą liczbę wzorów (336 z 504). Obserwacje te skłaniają do weryfikacji hipotezy o możliwości zastosowania prawa Zipfa w stosunku do nieciągłych jednostek języka, co stanowi treść rozdziału ósmego.

Z charakterystyki wyróżnionych klas frekwencyjnych wyraźnie wynika, że defektywność paradygmatyczna narasta wraz ze spadkiem rangi wzorów. Wzory szczególnie częste i bardzo częste obejmują ponad 60% wszystkich jednostek wielowyrazowych zanotowanych w słownikach fleksyjnych SEJF i VERBEL. Tworzą grupę najliczniejszą i jednocześnie najbardziej regularną pod względem fleksyjnym, a także nieskomplikowaną pod względem formalnym – są to zazwyczaj bi- lub trigramy: połączenia czasownika z rzeczownikiem lub wyrażeniem przyimkowym, rzeczownika i przymiotnika lub adverbium o postaci wyrażenia przyimkowych. Najczęstsze wzory odpowiadają jednak za stosunkowo niewielki procent wszystkich zwrotów. Dlatego, aby objąć wszystkie jednostki werbalne, trzeba uwzględnić inne, bardzo liczne wzory, z których większość odpowiada nielicznym zwrotom.

Spadek liczebności klas frekwencyjnych idzie w parze z defektywnością paradygmatu. Początkowo są to ograniczenia systemowe, wynikające zazwyczaj ze sposobu wypełnienia pozycji podmiotu w jednostkach werbalnych. Wraz ze wzrostem rangi wzorów widać wyraźną tendencję do pojawiania się ograniczeń pozaskładniowych, często nieprzewidywalnych

na podstawie budowy formalnej lub morfologicznej jednostek. Porównanie liczebności poszczególnych klas frekwencyjnych stanowi potwierdzenie praw Zipfa w przypadku jednostek wielowyrazowych, analizowanych z perspektywy ich budowy formalnej i kompletności paradigmatycznej.

Pozostałe osiągnięcia naukowe omówię zgodnie z wyróżnionymi obszarami.

a) **Zdyscyplinowany opis gramatyczny jednostek leksykalnych**

Doświadczenia zdobyte podczas prac anotacyjnych dość szybko wskazały mi obiekt dociekań językoznawczych. Stałe związki wyrazowe, połączenia zleksykalizowane wymykają się analizie maszynowej. Konieczne jest ustalenie ich postaci, wskazanie właściwości, tak by stworzyć opis gramatyczny porównywalny z opisem leksemów, czyli jednosegmentowych jednostek języka. Szczegółowe rozstrzygnięcia interpretacyjne, dotyczące jednostek systemu językowego obecne są m.in. w artykułach: „**Opis dystrybucji niektórych form tzw. zaimków trzecioosobowych**”, oraz „**Po pierwsze, z bliska, po polsku, czyli o wybranych wyrażeniach dwusegmentowych**”. Przedstawione w tekstach badania mają charakter empiryczny. Ich celem jest analiza wyrażen wybranych typów i nowatorska propozycja interpretacyjna. Konkluzje poprzedzone są analizami kwerend korpusowych.

Podsumowaniem tych zainteresowań jest monografia, będąca zmodyfikowaną wersją książkową pracy doktorskiej obronionej na Wydziale Humanistycznym UWM, tj. „**Opis powierzchniowoskładniowy wyrażen niestandardowych typu *na lewo, do dziś, po trochu, na zawsze* we współczesnym języku polskim**”. W wyniku analiz kwerend korpusowych i leksykograficznych proponuję w niej opis kilku typów grup połączeń przyimków – z rzeczownikami, przymiotnikami i leksemami nieodmiennymi.

Wynikiem zainteresowań polską fleksją są m.in. artykuły współautorskie, w których analizowałam wybrane leksemy czasownikowe, wyjątkowe ze względu na swoje właściwości morfologiczne: „**O odmianie czasowników o podstawach *mleć* i *pleć***” (współautor: Zygmunt Saloni) oraz „**Status czasownikowej formy *kole* we współczesnej polszczyźnie** (współautorzy: Sebastian Przybyszewski, Iwona Kosek).

Czasowniki o podstawach *mleć* i *pleć* przysparzają kłopotów użytkownikom polszczyzny, mają bowiem odmianę indywidualną i nietypową. W gramatykach omawia się je jako osobną grupkę. W klasyfikacji koniugacyjnej Jana Tokarskiego należą do grupy XI. Od długiego czasu znajdują się jednak pod silnym wpływem regularnej koniugacji według grupy VI (o podstawach *mieść*, *pieść*). Analiza kwerend korpusowych wyraźnie pokazuje, że formy tworzone według tego wzorca wchodziły do polskiego systemu stopniowo. Początkowo zmia-

nie uległy formy imiesłowu przymiotnikowego i odsłownika (z *melty*, *melcie* na *mielony*, *miele*). Oddziaływanie było na tyle silne, że formy nowe uznano za jedynie poprawne i wprowadzono do paradygmatów podawanych w gramatykach i słownikach – obok tradycyjnych form nieregularnych typu *miele*, *piele* oraz bezokolicznika *mleć*, *pleć*. Otrzymano w konsekwencji paradygmat hybrydalny. Wyjątkowość paradygmatu powinna mieć odzwierciedlenie w praktyce leksykograficznej. Autorzy artykułu proponują, by dla wszystkich czasowników o podstawach *mleć* i *pleć* podawać paralelne leksemy o podstawach *mieścić* i *pieścić*, opatrując je jednak odpowiednimi kwalifikatorami, np. *przest.* (*mleć* i *pleć*) lub *daw.* (*melto* i *melcie*), wskazującymi, że użycie ich form jest dziś wyraźnie ograniczone, nacechowane; neutralne, szeroko używane są formy od *mieścić* i *pieścić*.

Forma *kole* bywa opisywana jako człon izolowany, komponent frazeologizmów {coś} *kluje* / *kole* {kogoś} w *oczy*, {coś} *kluje* / *kole* {kogoś} w *zęby*. Tymczasem kwerendy korpusowe oraz analiza słowników polszczyzny pokazują, że można interpretować ją jako fragment paradygmatu czasownika *kluć* – formę wariantywną, opartą na temacie czasu teraźniejszego (*kole/kol*), zarówno w wystąpieniach samodzielnych, jak i frazeologicznie uwikłanych.

Dwie kolejne prace dotyczą kwestii fundamentalnych związanych z granicami i odmianą wielowyrazowych jednostek werbalnych. Artykuły te powstały w zespole autorów słownika VERBEL. Pierwszy z nich: „**Realizacja pozycji podmiotu otwieranej przez zwrot frazeologiczny a jego paradygmat (na wybranych przykładach)**” (współautorzy: Iwona Kosek, Sebastian Przybyszewski) dotyczy jednego z czynników mających wpływ na zasób form frazeologizmu, mianowicie sposobu wypełniania pozycji podmiotu-mianownika otwieranej przez związek frazeologiczny. W analizach uwzględniono podgrupę frazeologizmów werbalnych, rozumianych funkcjonalnie. Rozważane w artykule kwestie i podawane przykłady dotyczą różnych poziomów ograniczeń. Dyskutowane przykłady ujawniają pewne semantyczno-referencyjne ograniczenia selekcyjne argumentu mianownikowego. Z przeprowadzonej analizy wynika, że pozycja podmiotu markowana w formie podstawowej jednostki w postaci zaimka *coś* może oznaczać kilka różnych typów semantycznych realizacji pozycji podmiotu, mających wpływ na zasób form w paradygmacie. Są to: *coś_A* odnoszące się do abstraktów (nazw cech i stanów), *coś_B* odnoszące się tylko do konkretów lub do konkretów i abstraktów oraz *coś_C* odnoszące się do zwierząt. Wyróżnione trzy typy podmiotu nieosobowego wpływają na zasób form osobowych i rodzajowych zwrotów. Mój wkład w powstanie artykułu polegał na sformułowaniu hipotezy badawczej, zgromadzeniu materiału badawczego oraz przeprowadzeniu analizy danych językowych.

Z kolei praca: „**Wstępne uwagi o negacji w wielowyrazowych jednostkach werbalnych**” (współautor: Sebastian Przybyszewski) podejmuje kwestie granicy zwrotów grupy głównej. W szerokim ujęciu negacja jest obecna w ciągach wielowyrazowych najczęściej poprzez użycie partykuły *nie*. Jednak może się też przejawiać poprzez inne leksemy, np. przyimek *bez* (por. {coś} *mija bez echa*; {coś} *jest bez ceny*), spójnik *ani* (por. {ktoś} *ani się obejrzy*; {coś} *{kogoś} ani ziębi, ani grzeje*) lub zarówno poprzez partykułę *nie*, jak i spójnikiem *ani*, np. {ktoś} *nie zna dnia ani godziny*, {ktoś} *nie może ruszyć ręką ani nogą*. W wyniku analiz przykładów korpusowych oraz informacji słownikowej autorzy dochodzą do wniosku, że w kontekście negacji można wyróżnić kilka grup fraz czasownikowych:

1. jednostki neutralne pod względem kategorii negacji, to znaczy takie, które równie dobrze tworzą formy twierdzące, jak i przeczące, np. {ktoś} *zbija bąki*, {ktoś} *daje komuś kosza*, {coś} *idzie na marne*;
2. jednostki werbalne, których komponentem jest partykuła przecząca. W takiej sytuacji nie jest wykładnikiem negacji, a składnikiem jednostki wielowyrazowej, np. {ktoś} *za kołnierz nie wylewa*, {ktoś} *nie urodził się wczoraj*. Jednostki tego typu mogą występować w zdaniach bez *nie*, ale nie są to zdania asertoryczne (por. w pytaniach retorycznych);
3. jednostki, które zwykle występują w obu typach kontekstów – negacyjnym i afirmatywnym, ale wystąpienia z negacją zdecydowanie dominują i stanowią ich standardowy sposób użycia, co znajduje odzwierciedlenie w informacji słownikowej, gdzie są kwalifikowane jako ciągi zawierające *nie* (tak są zazwyczaj interpretowane w WSJP PAN), bądź też pojawiają się z komentarzem gramatycznym ZWYKLE NEG (szczególnie w ISJP PWN), np. {ktoś} *owija w bawełnę* lub {ktoś} *jest z cukru*;
4. jednostki, w których negacja jest ograniczona morfologicznie – zjawisko dotyczy zwrotów, których formy rozkaźnikowe w funkcji podstawowej są zanegowane, np. {ktoś} *odwraca kota ogonem*, {ktoś} *chowa głowę w piasek*;
5. jednostki występujące wyłącznie w kontekście pozytywnym (polskie *Positive Polarity Items*). Oznacza to, że nie można bezpośrednio przyłączyć partykuły *nie* do komponentu werbalnego danego zwrotu (brak negacji wewnątrzzdaniowej), co nie wyklucza możliwości pojawienia się *nie* w negacji przyzdaniowej, np. {ktoś} *zbiera grosz do grosza*. Do tej grupy można prawdopodobnie zaliczyć również np.: {coś} *ma się do czegoś jak pięść do oka*, {ktoś} *ma z kimś do pogadania*, {ktoś} *zna kogoś jak zły szeląg*.

Mój wkład w przygotowanie artykułu polegał na sprecyzowaniu problemu badawczego, wykonaniu kwerendy korpusowej oraz częściowo słownikowej, a także sporządzeniu analizy zgromadzonego materiału. Pełniłam także funkcję autora korespondencyjnego.

b) Leksykografia elektroniczna

W swoim dorobku mam zarówno słowniki, jak i prace poświęcone zagadnieniom teoretycznym, związanym z opisem leksykograficznym. Zajmowałam się także pracami redakcyjnymi wydawnictw naukowych. Pełniłam funkcję redaktora słownika gramatycznego polskich czasowników, tj. **Zygmunt Saloni, „Czasownik polski. Odmiana, słownik 12 000 czasowników”, Warszawa, wyd. I 2001, wyd. II 2002, wyd. III 2006.** Choć publikacja została tradycyjnie opublikowana, to powstała na podstawie elektronicznej bazy danych.

- Słowniki elektroniczne

Jestem współautorką dwóch elektronicznych słowników fleksyjnych wielowyrazowych jednostek leksykalnych: **„SEJF. Słownik elektroniczny jednostek frazeologicznych”, <http://zil.ipipan.waw.pl/SEJF>.** Warszawa 2012 (współautor Agata Savary), oraz **„VERBEL. Elektroniczny słownik paradygmatów polskich frazeologizmów czasownikowych”:** online: <http://uwm.edu.pl/verbel/> (współautorzy: Iwona Kosek, Sebastian Przybyśzewski).

Słownik SEJF zawiera tabele odmiany ok. 3 tys. jednostek wielowyrazowych, funkcjonalnie rzeczowników, przymiotników oraz przysłówków, a także marginalnie innych jednostek nieodmiennych. Mój wkład w powstanie słownika polegał na opracowaniu listy haseł na podstawie dostępnych źródeł leksykograficznych i naukowych, przygotowaniu opisu w formalizmie Multiflex. Agata Savary, współautorka, dostosowała aplikację leksykograficzną Toposław do potrzeb słownika oraz wykonała jego automatyczną walidację.

Słownik VERBEL zawiera ponad 5 tys. jednostek werbalnych, fragmentarycznych, tj. wymagających uzupełnienia – podmiotu (zwroty związku głównego) lub innego podrzędnika (zwroty grupy werbalnej). Mój polegał na współtworzeniu koncepcji ogólnej słownika, współprzygotowaniu listy haseł, analizie paradygmatów wybranych haseł oraz całościowym przygotowaniu opisu formalnego jednostek.

- Teoria leksykograficzna

Kwestie interpretacji i lematyzacji wybranej grupy jednostek wielowyrazowych omawiam w artykule **„O szczególnych nie-istniejących rzeczownikach we współczesnej polszczyźnie”.** Analizuję w nim wyrażenia, których przynajmniej jeden z komponentów poza wskazanym połączeniem nie występuje, jak np. *hokus-pokus, klituś-bajduś, fiksum dyr dum.* Na podstawie obserwacji materiału badawczego, na który składają fragmenty tzw. kartoteki

poznajskiej, korpusy tekstów i wybrane słowniki języka polskiego, proponuję szczegółowe rozwiązania leksykograficzne. Część z nich została przyjęta w *Słowniku gramatycznym języka polskiego*.

Prace nad słownikami SEJF o VERBEL były poprzedzone analizami teoretycznymi, związanymi ze specyfiką leksykografii elektronicznej oraz możliwościami objęcia opisem maszynowym jednostek wielowyrazowych języka polskiego. Tym zagadnieniom poświęcony jest artykuł: **„Computational Lexicography of Multi-Word Units: How Efficient Can It Be?”** (współautorzy: Filip Graliński, Filip Makowiecki, Agata Savary). Celem artykułu jest opis eksperymentów kodowania wielowyrazowych jednostek leksykalnych w dwu formalizmach: Multiflex i POLENG. Obydwa narzędzia służą do opisu morfologicznego jednostek wielowyrazowych w różnych językach, zwłaszcza słowiańskich, charakteryzujących się bogatym systemem fleksyjnym. Opisywane języki formalne były stosowane w różnych obszarach przetwarzania maszynowego języków naturalnych (NLP), jak: leksykografia obliczeniowa, tłumaczenie maszynowe, ekstrakcja terminów, identyfikacja nazwanych jednostek i normalizacja tekstu. Multiflex oferuje precyzyjny i jednoznaczny opis, a także dostosowuje się do różnych języków i modeli morfologicznych. Pozwala na łączenie wielu typów wariantów jednostek wielowyrazowych, takich jak akronimy, inwersje itp. Jednak jego użycie jest stosunkowo powolne i nie można opisać jednostek nieciągłych, tj. takich których komponenty nie występują w bezpośrednim sąsiedztwie. POLENG z kolei oferuje podejście uzupełniające: obejmuje szybszy proces półkontrolowanego kodowania, pozwala na przetwarzanie jednostek z nieciągłościami. Skuteczność formalizmu została potwierdzona w profesjonalnym tłumaczeniu maszynowym obszernych danych językowych. Jego działanie jest jednak bardziej ukryte, a łączenie wariantów, zwłaszcza fleksyjnych, można przeprowadzić jedynie w ograniczonym stopniu. Prace z udziałem leksykografów wykazały, że oba narzędzia można efektywnie wykorzystać do tworzenia zasobów morfologicznych jednostek wielowyrazowych. Pozwoliły także na zaproponowanie dalszych udoskonaleń formalizmów, takich jak weryfikacja kompletności opisu morfologicznego, sprawdzanie kompletności paradygmatu i kodowanie nowych haseł. Mój wkład w powstanie artykułu polegał na wykonaniu eksperymentu leksykograficznego z udziałem formalizmu Multiflex, dyskusji kwestii związanych z fleksją wyrażen oraz dostarczeniu właściwego materiału egzemplifikacyjnego.

Problematyka zawartości słownika, problemów leksykograficznych związanych z opisem fleksyjnym i formalnym wielowyrazowych jednostek nominalnych obecna jest w dwóch artykułach: **„SEJF – Słownik elektroniczny jednostek frazeologicznych”** oraz **„SEJF – A Grammatical Lexicon of Polish Multiword Expressions”** (współautorka: Agata Savary).

W pierwszym artykule koncentruję się na problematyce istotnej z polonistycznego punktu widzenia. Przedstawiam w nim etapy i sposób opisu fleksyjnego jednostek w formalizmie Multiflex, omawiam kwestie związane z uzyskaniem kompletnego opisu fleksyjnego, a także ich reprezentacji w słowniku (lematyzacja zgodna z danymi korpusowymi, wahania rodzaju jednostek rzeczownikowych, szyk wariantywny).

Drugi artykuł ma charakter bardziej techniczny. Zwraca uwagę na kwestie szczegółowe opisu formalnego jednostek wielowyrazowych, automatycznej walidacji słownika, a także omawia zastosowanie słownika w automatycznej anotacji jednostek wielowyrazowych w bankach drzew składniowych. Wyniki te pokazują, że SEJF można z powodzeniem zastosować do automatycznych adnotacji, ze względu na szczegółowe opisy gramatyczne zawarte w tym zasobie. Automatyczne ujednoznacznienie, czyli odróżnienie połączeń idiomatycznych od kompozycyjnych, pozostaje wyzwaniem w przetwarzaniu języków naturalnych. Walidacja słownika jednoznacznie wskazała, że koncepcja opisu fleksyjnego oparta na pojęciu fleksemu jest zasadniczo skuteczna w przypadku jednostek nominalnych. Próby wykorzystania jej do opisu jednostek werbalnych pokazały jednak konieczność modyfikacji. Mój wkład w przygotowanie artykułu polegał na opisie zawartości słownika, etapów opisu fleksyjnego i formalnego, wskazaniu problematyki związanej z wykorzystaniem pojęcia fleksemu. Pełniłam także funkcję autora korespondencyjnego.

Kwestiom związanym z obecnością jednostek obcego pochodzenia w słownikach języka polskiego poświęcony jest artykuł: „**Jednostki obce w słowniku języka polskiego na przykładzie Słownika elektronicznego jednostek frazeologicznych (SEJF)**”. Na podstawie analiz słowników ogólnych i frazeologicznych polszczyzny oraz kwerend korpusowych proponuję listę tego typu jednostek oraz ich opis gramatyczny. Artykuł kończy lista wyrażen zanotowanych w słowniku SEJF, zawierająca przypisaną każdemu połączeniu kwalifikację klasy gramatyczną, a w przypadku rzeczowników – także wartość rodzaju.

Koncepcja słownika VERBEL została omówiona w artykule: „**O projekcie elektronicznego słownika odmiany frazeologizmów czasownikowych**” (współautorzy: Iwona Kossek, Sebastian Przybyszewski). Udział współautorów w przygotowanie tego artykułu oraz innych towarzyszących powstaniu słownika jest równy. Każdy z autorów brał udział w dyskusjach nad kształtem, zawartością, materiałem leksykalnym i podstawami metodologicznymi opracowania. Realizacji słownika towarzyszyły wystąpienia, a także artykuły współautorskie.

Problemy lematyzacji zwrotów zostały podjęte w: „**The problem of lemmatisation in the Polish inflectional dictionary of verbal MWES**”. Wybór odpowiedniej formy hasłowej powinien być nie tylko wygodnym rozwiązaniem technicznym, ale także dostarczać użytkow-

nikowi jak najszerszych informacji. W polskiej tradycji leksykograficznej nazwą hasła czasownikowego jest bezokolicznik. Wybór tej formy wiąże się z pewnymi problemami: forma bezokolicznika nie dostarcza informacji na temat odmiany głównego komponentu, hasło wielowyrazowe podane w bezokoliczniku nie zdaje sprawy z różnic znaczeniowych oraz składni połączenia (*{ktoś} idzie pod nóż* ‘o człowieku: iść na operację; *{coś} idzie pod nóż* ‘o zwierzęciu: być przeznaczonym do uboju). W artykule wykazujemy, że w wypadku polskich zwrotów celowe jest stosowanie 3.os. lp. jako formy podstawowej. W ten sposób naturalnie wskazuje się pozycje otwarte w jednostce wielowyrazowej i możliwości ich wypełnienia. Dzięki zastosowaniu formy finitywnej możliwe jest wskazanie najbardziej typowej formy, np. *{ktoś} robi {komuś} wodę z mózgu* (zazwyczaj w czasie teraźniejszym).

W artykule: „**Źródła w tworzeniu elektronicznego słownika paradygmatów frazeologizmów werbalnych**” poddano analizie typy źródeł i sposobów ich wykorzystania na poszczególnych etapach opracowywania słownika VERBEL. W konkluzjach wskazano, że metaźródła słownikowe mają zastosowanie jako istotny element w ustalaniu i weryfikowaniu listy jednostek. Źródła korpusowe natomiast – poza funkcją egzemplifikacyjną, znajdują zastosowanie w weryfikowaniu hipotez, choć w ograniczony sposób. Ograniczenia te (także przy preferowanym podejściu *corpus-based approach*) wynikają z charakteru słownika fleksyjnego, w którym wszystkie formy z założenia należy podać *in extenso*. Przyjmując prymat systemu językowego nad uzusem, o włączeniu lub niewłączeniu danej formy do paradygmatu rozstrzyga analiza właściwości znaczeniowo-strukturalnych zwrotu. Korpus natomiast może posłużyć do podania pewnych elementów charakterystyki pragmatycznej i znalezienia kontekstów nieprzewidzianych przez badacza. W tworzeniu słownika fleksyjnego zwrotów większość decyzji dotyczących statusu danych form jest oparta na kompetencji językowej leksykografów, bowiem frekwencja poszczególnych form zwrotów jest nierówna. Wiele form nie znajduje potwierdzenia w korpusach, co nie przesądza o pełności paradygmatu jednostki wielowyrazowej. Wydaje się, że nawet tam, gdzie źródła słownikowe czy opracowania naukowe wskazują status danej formy, każdorazowo powinien on być weryfikowany w oparciu o inne źródła (w tym korpusowe) i kompetencję językową słownikarzy.

Słownik wraz z opisem podstaw teoretycznych, zawartości i budowy znajduje się w monografii: „**VERBEL. Elektroniczny słownik paradygmatów polskich frazeologizmów czasownikowych. Teoria, problemy, prezentacja, wersja 2.0**”. W rozdziale 4. monografii – *Budowa słownika* omawiam założenia zastosowanego opisu formalnego, strukturę artykułu hasłowego oraz poszczególne etapy opisu zgromadzonych w słowniku jednostek.

Prezentacja słownika znajduje się w artykule: „*Not Only Meaning... Verbel: The Electronic Dictionary of Paradigms of Polish Verbal Multiword Expressions*” (współautorzy: Sebastian Przybyszewski, Iwona Kosek). W pierwszej części artykułu omówiono podstawy teoretyczne i metodologiczne stosowane przy tworzeniu słownika, a także problemy związane z jego tworzeniem, charakterystyczne dla języków złożonych fleksyjnie, takich jak polski. W drugiej części omówiono koncepcję i treść słownika. Na koniec opisano narzędzia przetwarzania języka naturalnego, które leżą u podstaw pracy nad słownikiem.

Problematyce tworzenia słowników elektronicznych poświęciłam artykuł: „*Dictionaries and lexical databases of Polish language*”. Przedstawiam w nim wpływ technik komputerowych na organizację, rozmiar, przeznaczenie i zawartość słowników. W badaniach uwzględniłam trzy polskie słowniki elektroniczne: *Uniwersalny słownik języka polskiego PWN*, *Wielki słownik języka polskiego PAN* oraz *Słownik gramatyczny języka polskiego*. W wyniku analiz dowodzę, że na sposób organizacji i prezentacji wiedzy we wspomnianych dziełach wpływ mają bazy danych, na podstawie których te słowniki powstały. Dzięki bazom oraz odpowiednim sposobom prezentacji zawartych w nich danych użytkownicy mogą dotrzeć do szczegółowych informacji o jednostkach leksykalnych, grupować je, a także kompilować podsłowniki, spełniające określone warunki (m.in. tematyczne, frekwencyjne, gramatyczne).

Z kolei w artykule: „*The structure of a dictionary entry and grammatical properties of multi-word units*” omawiam wpływ właściwości gramatycznych jednostek językowych na opis leksykograficzny, w szczególności na strukturę hasła słownikowego. Zagadnienie to zostało przedstawione na przykładzie słownika VERBEL. Choć jest to słownik specjalistyczny, zawierający formalny opis jednostek wielowyrazowych, to dzięki odpowiedniej strukturze haseł możliwe jest korzystanie z niego także przez osoby niebędące specjalistami. Możliwość wyboru zakresu informacji leksykograficznej w słowniku gwarantuje dwustopniowy schemat hasła: od ogółu do szczegółu, czyli od ogólnych informacji opisowych o jednostce do zestawienia wszystkich jej form (słów morfologicznych) z przypisanymi im cechami fleksyjnymi. Wydaje się, że dane podane na etapie początkowym (ogólne informacje o klasie gramatycznej, wartość aspektu, obecność odpowiednika aspektowego, znaczenie i przykłady) oraz na etapie końcowym (wszystkie formy fleksyjne jednostki wraz z przypisaną im charakterystyką morfoskładniową) stanowią wystarczający opis leksykograficzny. Umieszczenie poszczególnych rodzajów informacji w odrębnych zakładkach słownika daje użytkownikowi swobodę ich stosowania. Jest bardzo prawdopodobne, że przeciętny użytkownik będzie zadowolony z ogólnego opisu i być może zajrzy także do listy form. Z kolei spe-

cialistę – językoznawcę, leksykografa, informatyka – zaciekawiają różne etapy opisu i techniczne sposoby ich prezentacji. Choć słownik zawiera szczegółowe informacje opisane według specjalistycznego modelu językowego, jego podstawowe użycie nie wymaga rozległej wiedzy specjalistycznej. Zakładki zawierające szczegóły opisu formalnego można pominąć bez utraty funkcjonalności słownika.

c) **Korpusy tekstów oraz inne narzędzia elektroniczne w praktyce językoznawczej i pozajęzykoznawczej**

Problematykę korpusów, ich wykorzystania w zaawansowanych badaniach językoznawczych podjęłam w swoich najwcześniejszych pracach: „**Polskie korpusy tekstów (wybrane zagadnienia)**” oraz „**Co skreślano i co dopisywano w korpusie *Słownika frekwencyjnego polszczyzny współczesnej***” (współautor Zygmunt Saloni). Artykułom przyświeca przekonanie o konieczności jednoznacznego i zdyscyplinowanego opisu gramatycznego zleksykalizowanych połączeń wyrazowych. Na początku XXI wieku problematyka dostępności korpusów tekstów języka polskiego, uwzględnienie ich w badaniach językoznawczych nie był szczególnie często podejmowany. W swoich tekstach analizuję kwestie metodologiczne związane z kompilacją korpusów, jak decyzje segmentacyjne i interpretacyjne danych typów form wyrazowych. Przedstawiam także możliwości zastosowania obszernych zbiorów tekstowych w badaniach gramatycznych i stylistycznych. Nakreślona w artykułach perspektywa badań empirycznych, konieczność objęcia opisem językoznawczym możliwie najbardziej obszernego materiału językowego stała się dla mnie wyznacznikiem w dalszych poszukiwaniach naukowych.

W innych pracach z tego obszaru analizuję możliwości zastosowania specjalistycznych narzędzi elektronicznych. Zakres wykorzystania specjalistycznej aplikacji w pracach leksykograficznych wielowyrzowych jednostek leksykalnych jest treścią artykułu: „***Toposław* jako narzędzie znakowania jednostek wieloczłonowych**”. Kwestii szczegółowej, mianowicie tym wyrażeniom nominalnym, których rodzaj różni się od wartości tego parametru ich głównego komponentu (np. *koziół ofiarny, mól książkowy, dziadek do orzechów*) poświęcony jest artykuł „**Problemy opisu związków frazeologicznych w formalizmie *Multiflex* (na przykładzie rodzaju wyrażen frazeologicznych)**” (współautorka: Iwona Kosek). W pracy przedstawiono problemy rygorystycznego opisu tego typu wyrażen oraz potencjalne rozwiązania w aplikacji *Toposław*. Mój wkład w powstanie artykułu polega na sformułowaniu problemu badawczego, zgromadzeniu materiału potrzebnego do jego weryfikacji, przygotowaniu opisu formalizmu oraz przeprowadzeniu analizy danych.

Z kolei użyteczność aplikacji *Jasnopis*, badającej stopień trudności polskich tekstów użytkowych, analizowałam w dwóch artykułach: „**Możliwości zastosowania narzędzia do pomiaru zrozumiałości tekstu *Jasnopis* w praktyce szkolnej**” oraz: „***Jasnopis* a wsparcie osób z zaburzeniami mowy – na przykładzie materiałów ze środowiskowej debaty konsultacyjnej osób słabosłyszących**” (drugi we współautorstwie z Magdaleną Osowicką-Kondratowicz). W artykułach tych dowodzę, że właściwie zaprojektowane programy komputerowe, jak np. *Jasnopis*, mogą znaleźć zastosowanie w różnorodnych dziedzinach życia społecznego, jak edukacja czy równy dostęp do informacji prawnej i języka w szczególności dla osób niepełnosprawnych (np. słabosłyszących oraz mniejszości językowej Głuchych). Mój wkład w powstanie artykułu współautorskiego polega na sformułowaniu hipotezy badawczej oraz zgromadzeniu i analizie materiału badawczego

Problematyka ogólnodostępnych, nie specjalistycznych narzędzi elektronicznych w warsztacie logopedy jest treścią: artykułu: „**Wykorzystanie programów komputerowych w praktyce logopedycznej**”. Tekst ten stanowi przewodnik dla logopedów. Jego tezy stały się szczególnie aktualne w okresie pandemii covid-19.

5. Informacja o wykazywaniu się istotną aktywnością naukową albo artystyczną realizowaną w więcej niż jednej uczelni, instytucji naukowej lub instytucji kultury, w szczególności zagranicznej.
 - a. Członek Management Committee w sieci naukowej PARSEME (PARSInG and Multiword Expressions). Towards linguistics precision and computational efficiency in natural language processing (akcja COST nr IC1207), 2013–2017, kierownik projektu: prof. Agata Savary, <https://www.cost.eu/actions/IC1207/#tabs+Name:Management%20Structure>;
 - b. Koordynator zespołu konsorcjum badawczego z ramienia UWM, przygotowującego wniosek grantowy w konkursie MAESTRO6 (2014) – „Koncepcja słownika kombinatorycznego i jego wielopoziomowa weryfikacja na przykładzie języka polskiego”, kierownik projektu: prof. dr hab. Adam Przepiórkowski, instytucja prowadząca: Instytut Podstaw Informatyki Polskiej Akademii Nauk (wniosek pomyślnie zrecenzowany, nieskierowany do finansowania);
 - c. Staż naukowy (stypendium zewnętrzne) w Université François Rabelais de Tours, Blois, Francja, czerwiec – lipiec 2014; opiekun stażu: prof. Agata Savary;
 - d. Członek Komitetu Naukowego międzynarodowych konferencji PARSEME General Meeting: Ateny, Grecja, 10–11 marca 2014; Valetta, Malta, 19–20 marca 2015; Ohrid, Macedonia, 7–8 kwietnia 2016; Dubrownik, Chorwacja, 27–28 września 2017.

- e. Członek Management Committee w sieci naukowej UniDive – Universality, diversity and idiosyncrasy in language technology (akcja COST nr CA21167), 2022–2026, kierownik projektu: prof. Agata Savary, <https://www.cost.eu/actions/CA21167/#tabs+Name:Management%20Committee>;
- f. Udział w kursach tematycznych i szkołach lingwistycznych:
- Wiosenne Kursy Lingwistyczne im. Vilema Mathesiusa, Uniwersytet Karola, Praga, Czechy, 15–28 marca 2000, 7–20 marca 2004, 6–13 marca 2005;
 - ELSNET Summer School of Text and Speech Corpora, Uniwersytet Karola, Praga, Czechy, 18–30 lipca 2001;
 - PARSEME 1st Training School, Uniwersytet Karola, Praga, Czechy, 19–23 stycznia 2015.
- g. Udział w konferencjach naukowych:
1. Konferencja Młodych Sławistów POLYSLAV, Warszawa 26–28.09.2002, referat „Jego – go, czyli uwagi o zaimkach osobowych w języku polskim”;
 2. Przyszłość słowników w Polsce, Kamień Śląski 1–2.04.2004, referat „Nietypowe jednostki dwusegmentowe w Słowniku frekwencyjnym polszczyzny współczesnej”;
 3. Gramatyka i semantyka współczesnego języka polskiego, Toruń 2–4.09.2004, referat „*Po pierwsze, z bliska, po polsku*, czyli o wybranych wyrażeniach dwusegmentowych”;
 4. Konferencja Młodych Sławistów POLYSLAV, Bergamo 23–25.09.2004, referat „Jak hasłować niektóre niestandardowe wyrażenia przyimkowe”;
 5. Glosa do leksykografii polskiej, Warszawa 8–9.09.2009, referat „O szczególnych nie-istniejących rzeczownikach we współczesnej polszczyźnie”;
 6. Nowe zjawiska w języku, tekście, komunikacji III, Olsztyn, 13–15.05.2010, referat „Multiflex jako narzędzie znakowania jednostek wielocłonowych”;
 7. III Glosa do leksykografii polskiej, Warszawa 15–16.09.2012, referat „Jednostki obce w słowniku języka polskiego na przykładzie elektronicznego słownika wyrażzeń *Toposław*”;
 8. Poznań Linguistic Meetings, 8–10.09.2012, referat “The grammatical dictionary of Polish multiword units”;
 9. Human Language Technology Days, Warszawa, 26–28.09.2012, plakat “Electronic Dictionary of Polish Multiword Units”;

10. III Ogólnopolskie Sympozjum Naukowe z cyklu Perspektywy współczesnej frazeologii polskiej. Między teorią a praktyką frazeologiczną, Poznań, 23.10.2012, referat: „Frazeologia w satyrze politycznej na przykładzie twórczości Janusza Szpołańskiego”;
11. Bogactwo współczesnej polszczyzny. w stulecie *Języka Polskiego*, Kraków, 12–13.04.2013, „SEJF – Słownik elektroniczny jednostek frazeologicznych”;
12. Przyszłość językoznawstwa – językoznawstwo przyszłości. Konferencja jubileuszowa Instytutu Języka Polskiego PAN, Kraków, 10–12.06.2013, referat „Frazeologia w kontekście lingwistyki komputerowej – o nowym projekcie leksykograficznym” (współautor: Iwona Kosek);
13. General Meeting of PARSEME, Warszawa 16–18.09.2013, referat “An electronic dictionary of Polish multiword units as an example of using text and speech corpus in lexicographical studies”;
14. General Meeting of PARSEME, Ateny, 10–11.03.2014, poster “Verbal multiword expressions in Polish”;
15. Nowe zjawiska w języku tekście, komunikacji V. Frazeologia w stylach i gatunkach, Olsztyn, 21–22.05.2014, referat „Skąd się biorą dziury w serze, czyli o typach nieciągłości wewnątrz jednostek frazeologicznych”;
16. Europhras 2015. Computerised and Corpus-based Approaches to Phraseology: Monolingual and Multilingual Perspectives, Malaga 29.06–1.07.2015, referat: “The problem of lemmatization in the Polish inflectional dictionary of verbal MWEs” (współautorzy: Iwona Kosek, Sebastian Przybyszewski);
17. 7th language & Technology Conference, Poznań 27–29.11.2015, referat: “SEJF – The Grammatical Lexicon of Polish Mutliword Expressions” (współautorka: Agata Savary);
18. Nowe zjawiska w języku, tekście, komunikacji VI. Język edukacji, język w edukacji, Olsztyn 21–22.04.2016, referat: „*Jasnopis* w praktyce szkolnej”;
19. Język jest labiryntem ścieżek. Nowe kierunki i nowe zadania w badaniach nad językiem polskim, Warszawa 18–19.06.2016, referat: „Negacja w wielowyrzowych jednostkach werbalnych” (współautor: Sebastian Przybyszewski);
20. Intra- and Interlinguale Zugänge zur deutschen Phraseologie und Paromiologie, Wrocław 16–18.03.2017, referat: „Paradygmat a znaczenie związku frazeologicznego (na wybranych przykładach polskich i niemieckich)”, współautorzy: Iwona Kosek, Sebastian Przybyszewski;

21. Źródła w badaniach językoznawczych. 75. Zjazd Polskiego Towarzystwa Językoznawczego, Łódź 18–19.09.2017, referat: „Źródła w tworzeniu elektronicznego słownika paradygmatów frazeologizmów werbalnych” (współautorzy: Iwona Kosek, Sebastian Przybyszewski);
22. XXVII Międzynarodowe Kolokwium Slawistyczne, Lwów 16–17.05.2019, referaty: „Reprezentatywność korpusu specjalistycznego na przykładzie korpusu tekstów naukowych IPHRAS” oraz „Fleksja frazeologizmów czasownikowych języka polskiego – stan badań, problemy opisu (na podstawie prac nad *Słownikiem paradygmatów polskich frazeologizmów czasownikowych*)” we współautorstwie z Iwoną Kosek i Sebastianem Przybyszewskim;
23. Polonistyka w świetle tradycji i wyzwań współczesności, Lwów 4–6.10.2019, referat: „Po co (nam) słownikowe bazy danych?”;
24. Electronic Lexicography in the 21st Century (eLex), Brno 5–7.07.2021, referat: “The structure of a dictionary entry and grammatical properties of multi-word units”.

6. Informacja o osiągnięciach dydaktycznych, organizacyjnych oraz popularyzujących naukę lub sztukę.

Działalność dydaktyczna

Podczas zatrudnienia w UWM na stanowisku adiunkta realizuję rocznie średnio 280 godzin dydaktycznych. Wśród nich są wykłady autorskie, ćwiczenia kursowe, konwersatoria specjalizacyjne oraz seminaria dyplomowe prowadzone na kierunkach: filologia polska, logopedia, lingwistyka w biznesie.

Wykłady autorskie:

- Zagadnienia językoznawstwa komputerowego, filologia polska, od 2012 r.
- Wybrane zagadnienia muzykoterapii, wykład monograficzny dla studentów Wydziału Humanistycznego, od 2019 r.

Ćwiczenia kursowe:

- Gramatyka opisowa współczesnego języka polskiego, filologia polska, od 2006 r.
- Funkcjonalny opis języka polskiego, logopedia, od 2011 r.
- Kultura języka polskiego, filologia polska, filologia polska z językiem ukraińskim, lingwistyka w biznesie, od 2006 r.
- Logorytmika i umuzykalnienie, logopedia, od 2014 r.

- Muzyka w terapii logopedycznej, logopedia, 2015–2020 r.
- Warsztaty pracy z tekstem, lingwistyka w biznesie, od 2022 r.

Konwersatoria specjalizacyjne:

- Komputerowe opracowanie tekstu, specjalność wydziałowa: Edytorstwo tekstu, 2009–2015 r.
- Technologie informacyjne w humanistyce, filologia polska, 2012–2018 r.
- Metodyka pracy z komputerem, logopedia, 2017–2021 r.
- Logopedia komputerowa i internetowa, logopedia, 2019–2020 r.
- Lingwistyka komputerowa, lingwistyka w biznesie, od 2021 r.

W wyniku prowadzonych seminariów dyplomowych wypromowałam dwanaścioro licencjuszy (r. akadem. 2014/2015) na kierunku filologia polska oraz dwanaście licencjuszek na kierunku logopedia (2021/2022). Zrecenzowałam ponad 80 prac licencjackich i magisterskich na kierunkach: filologia polska, logopedia, lingwistyka w biznesie.

Byłam koordynatorem blisko 30 przedmiotów akademickich, w tym także praktyk zawodowych na kierunku filologia polska.

Od 2019 r. pełnię obowiązki promotora pomocniczego w przewodzie doktorskim mgr Barbary Szwedy w dyscyplinie językoznawstwo (temat pracy: „Analiza lingwistyczna stron internetowych firm farmaceutycznych”; promotor: prof. dr hab. Alina Naruszewicz-Duchlińska).

Działalność organizacyjna

- zastępca dyrektora Instytutu Filologii Polskiej UWM w latach 2010–2014. Do moich obowiązków należało m.in. przygotowanie dokumentacji i zmiany struktury studiów zgodnie z Krajową Ramą Kwalifikacji. Podczas swojej kadencji kierowałam także zespołem opracowującym nowy program studiów – logopedia, studia stacjonarne i stopnia (2010–2012), oraz byłam członkiem zespołu opracowującego kierunek logopedia, st. stacjonarne II stopnia (2012–2014);
- członek Towarzystwa Miłośników Języka Polskiego, od 2010 r.
- członek wydziałowej komisji ds. ocen nauczycieli akademickich, 2012–2014;
- członek wydziałowej komisji ds. zapewniania jakości kształcenia, 2012–2015;
- członek Rady Wydziału Humanistycznego UWM, 2013–2016;
- członek zespołu ds. jakości kształcenia na kierunku logopedia, od 2015 r.;
- członek Polskiego Towarzystwa Logopedycznego, od 2015 r.

- członek Wydziałowej Komisji Rekrutacyjnej, 2016–2018;
- członek komisji konkursowej na Wydziale Humanistycznym UWM, 2017–2023;
- opiekun roku, 2016–2020;
- redaktor, członek kolegium redakcyjnego czasopisma naukowego „Prace Językoznawcze”, od 2014 r.;
- członek komisji ekspertów oceniających projekty w ramach konkursu: „Studencki Grant Rektora” (nauki humanistyczne), od 2023 r.

Działalność popularyzatorska

- prowadzenie warsztatów pisarsko-edytorskich w wybranych placówkach edukacyjnych w Olsztynie (np. Szkoła Podstawowa nr 2, Szkoła Podstawowa nr 15, Przedszkole Niepubliczne TEMPOO, Playschool), 2012–2016;
- organizacja i prowadzenie warsztatów podczas XIV Olsztyńskich Dni Nauki i Sztuki (Zabawy z językiem), 2014, 2016 r.,
- organizator, członek jury w Ogólnopolskim Konkursie na Najtrudniejszy Skrętacz Języka Polskiego i Lingwołamkę, od 2015 r.
- członek jury w miejskim dyktandzie z okazji Dnia Języka Ojczystego, 2016 r.;
- współpraca z Uniwersytetem Dzieci – warsztaty edytorskie, mentor ścieżki językoznawczej, 2016–2017;
- członek jury Ogólnopolskiego Seminarium Kół Naukowych w Olsztynie, 2017, 2022 r.
- systematyczny udział w audycjach Polskiego Radia Olsztyn – „Wieczór z kulturą”, cykl „Potyczki z językiem” w latach 2017–2022.

7. Oprócz kwestii wymienionych w pkt. 1-6, wnioskodawca może podać inne informacje, ważne z jego punktu widzenia, dotyczące jego kariery zawodowej.

a) Projekty naukowe, realizowane w UWM

- 2002–2005 – wykonawca w grantie KBN „Słownik gramatyczny języka polskiego”, kierownik projektu: prof. dr hab. Zygmunt Saloni;
- 2004–2005 – główny wykonawca w grantie KBN „Opis lingwistyczny wybranych wyrażen niestandardowych składniowo we współczesnym języku polskim”, kierownik projektu: prof. dr hab. Zygmunt Saloni;

- 2014–2018 – główny wykonawca w projekcie NCN „Opis paradygmatyczny polskich frazeologizmów czasownikowych. Słownik elektroniczny”, kierownik projektu: dr hab. Iwona Kosek, prof. UWM;
- 2017–2019 – wykonawca w projekcie Polsko-Niemieckiej Fundacji Nauki „D/P-IPHRAS: Interphraseologie als Element der Wissenschaftssprache (D/P-IPHRAS: Interfrazologia jako element języka nauki)”, kierownik po stronie niemieckiej – prof. Bettina Bock, kierownik po stronie polskiej – dr hab. Joanna Targońska.

Współpraca zewnętrzna

- 2001–2004 – Instytut Podstaw Informatyki PAN, grant KBN nr 7T11C04320 „Anotowany korpus pisanego języka polskiego z dostępem przez Internet (z uwzględnieniem zastosowań w inżynierii lingwistycznej)”, kierownik projektu: prof. dr hab. Adam Przepiórkowski, zakres prac: ujednoznacznianie wyników analiz morfologicznych słów w tekstach;
- 2009–2014 – współpraca z IPI PAN w ramach projektu „NEKST – Adaptacyjny system wspomagający rozwiązywanie problemów w oparciu o analizę treści dostępnych źródeł elektronicznych” (<https://nekst.ipipan.waw.pl/index.html>), zakres prac: opis jednostek frazeologicznych we współczesnej polszczyźnie;
- 2019–2021 – członek kluczowej kadry projektu Narodowego Centrum Badań i Rozwoju (program Szybka Ścieżka) „Technologia kontekstowego rozumienia języka pisanego na potrzeby poprawy błędów oraz automatycznej oceny zrozumiałości tekstu” (członek), partner biznesowy: Ermlab Software (<https://goodwrite.pl/o-projekcie/>);
- 2021–2023 – współpraca z Politechniką Wrocławską, Katedrą Sztucznej Inteligencji w ramach projektu: „CLARIN – Wspólne Zasoby Językowe i Infrastruktura Technologiczna finansowany w ramach Programu Operacyjnego Inteligentny Rozwój 2014–2020, Priorytet IV: Zwiększenie potencjału naukowo-badawczego, Działanie 4.2: Rozwój nowoczesnej infrastruktury badawczej sektora nauki, nr POIR.04.02.00-00C002/19”, kierownik zadania: G3.2: Słownik jednostek wieloczłonowych ze zintegrowanym opisem fleksyjnym;
- Od 12.09.2022 r. – aktualizacja informacji fleksyjnej w *Wielkim słowniku języka polskiego PAN* (wsjp.pl).

.....

 (podpis wnioskodawcy)