



Jak współcześnie badać frazeologię w tekstach dawnych? Możliwości wykorzystania narzędzi cyfrowych w pracy historyka języka

How to study phraseology in historical texts using contemporary methods?
The potential to use digital tools in the work of a language historian

Abstract: The research aims to overcome the limitations of the systemic examination of phraseology in historical texts using available digital tools. The article seeks to answer the question posed in the title by presenting a functional study of phraseology using the *Korpus dawnych polskich tekstów dramatycznych (1772–1939)* [A corpus of Polish dramatic texts (1772–1939)] and the intonation analyzer *Wydźwięk*. I conduct pragma- and sociolinguistic analysis in the context of language functions identified by Michael Halliday; I also propose using selected sociological factors included in the corpus that may correlate with the choice of phraseology in a given context. I verify the evaluation of the intonation of expressions by comparing the tool's results with my linguistic intuition. The research results are presented in a separate section, juxtaposed with the state of research (systematic analysis) and the researcher's linguistic competencies. The prospects for using digital tools in phraseological research in ancient texts are presented in the conclusion section.

Key words: phraseology, historical texts, pragmalinguistic analysis, sociolinguistic analysis, digital tools

Abstrakt: Celem prezentowanych badań jest przekroczenie ograniczeń związanych z systemową analizą frazeologii w tekstach dawnych dzięki wykorzystaniu wybranych narzędzi cyfrowych: *Korpusu dawnych polskich tekstów dramatycznych (1772–1939)* oraz analizatora emocji tekstu *Wydźwięk*. Prezentowana analiza stanowi próbę odpowiedzi na postawione w tytule artykułu pytanie. Zgromadzony materiał jest badany na dwóch poziomach: pragmalingwistycznym (w kontekście funkcji językowych wyróżnionych przez Michaela Hallidaya i wydźwięku emocjonalnego wypowiedzi) oraz socjolingwistycznym (z wykorzystaniem wybranych czynników socjologicznych zawartych w korpusie, które mogą korelować z wyborem frazeologii w określonym kontekście). Wyniki przedstawiono w osobnej sekcji, zestawiając je z aktualnym stanem badań (analizą systemową) oraz poczuciem językowym badacza. Perspektywy wykorzystania narzędzi cyfrowych w badaniach nad frazeologią w tekstach dawnych sformułowano w zakończeniu.

Słowa kluczowe: frazeologia, teksty dawne, analiza pragmalingwistyczna, analiza socjolingwistyczna, narzędzia cyfrowe

Współczesna humanistyka od prawie dwóch dekad pozostaje w ścisłym związku z cyfrowością¹, której jednym z przejawów, zwłaszcza w praktyce lingwistycznej, stały się elektroniczne korpusy tekstów. Zestawienie oraz możliwości wykorzystania zbiorów w historycznojęzykowych badaniach polszczyzny były kilkakrotnie prezentowane międzynarodowej społeczności akademickiej (por. Gruszczyński, 2020; Gruszczyński i in., 2022; Łaziński i in., 2023; Mitrenga i in., 2021; Opaliński i in., 2019; Pastuch i in., 2018; 2024a i b). Stan ten, jak również prace śląskich językoznawczyń nad pierwszym polskim korpusem diachronicznym znakowanym socjopragmatycznie – *Korpusem dawnych polskich tekstów dramatycznych (1772–1939)* (dalej: KorTeDa) (Mitrenga i in., 2021)² skłoniły mnie do refleksji nad możliwością rozszerzenia studiów nad frazeologią dawną o ujęcie pragmalingwistyczne. Trudności w dostępie do źródeł rejestrujących leksykę, a zwłaszcza frazeologię potoczną, powodują bowiem, że w obszarze tym dominują sondy słownikowe (zob. zwłaszcza: Jawór, 2009; Piel, 2023; Przymuszała, 2018) oraz badania etymologiczne (Ignatowicz-Skowrońska, 2020; Kosek i Zawilska, 2012; Piel, 2018; Puda-Blokesz, 2021). Być może szansę przezwyciężenia tych ograniczeń stanowią nowoczesne technologie służące gromadzeniu i analizie tekstów? Próbą odpowiedzi na to pytanie są niniejsze rozważania. Składają się one z dwóch zasadniczych części:

- przedstawienia założeń metodologicznych i charakterystyki materiału;
- analizy materiału w ujęciach systemowym i funkcjonalnym.

Konkluzje oraz perspektywy dalszych badań znajdują się w *Zakończeniu*.

Założenia metodologiczne

Bogata oferta narzędzi cyfrowych oraz konieczność włączenia w obszar lingwistyki cyfrowej historycznojęzykowych analiz materiału frazeologicznego dały asumpt do sformułowania pytania postawionego w tytule artykułu. Postulując konieczność wyznaczenia nowego kierunku diachronicznych analiz nad frazeologią, pragnę zwrócić uwagę na trzy kluczowe zagadnienia – cele mojej propozycji:

¹ Początków badań w tym zakresie upatruje się już w połowie XX w., jednak to w XXI w. John Unsworth, Susan Schreibman i Ray Siemens (Schreibman i in., 2004) zwrócili uwagę na konieczność rozszerzenia dziedziny także na projekty, w których wykorzystuje się nowoczesne technologie cyfrowe do (ponownych) analiz tradycyjnych artefaktów (w tym danych językowych – A.R.).

² Korpus jest tworzony w Instytucie Językoznawstwa Uniwersytetu Śląskiego w Katowicach. W zakresie doboru źródeł i anotacji narzędzie wzorowano na *A Corpus of English Dialogues* (CED) – składa się ono z 50 oryginalnie polskich, niewierszowanych tekstów dramatycznych o powszechnym i popularnym charakterze, na które naniesiono informację socjopragmatyczną, tj. taką, w której uwzględnia się informacje zarówno socjologiczne (por. CED), jak i pragmalingwistyczne (znakowano m.in. przysłówia i frazeologizmy). Korpus znajduje się w fazie testowej; na potrzeby badań został mi udostępniony przez Autorki – Barbarę Mitrengę, Magdalenę Pastuch i Kingę Wąsińską (stan na 24.04.2023). Szczegółowe dane dotyczące tworzenia narzędzia, w tym charakterystyka źródeł i kryteria doboru materiału, opisano w: Mitrenga i in., 2021 (zob. zwłaszcza s. 169–171, 174–176).

- przekroczenie ograniczeń związanych z systemowym badaniem frazeologii;
- funkcjonalne badanie frazeologii z wykorzystaniem narzędzi cyfrowych (na podstawie dawnych tekstów pisanych przybliżających mowę danego okresu – tzw. *speech-related text types*³;
- obranie perspektywy badawczej zorientowanej na analizę materiału pozwalającą uwzględnić czynniki zewnętrzne (np. wybrane zmienne socjolingwistyczne przypisane interlokutorom).

Implementację wskazanych założeń przedstawię z wykorzystaniem KorTeDa oraz narzędzia do analizy nacechowania emocjonalnego wypowiedzi (sentymetu)⁴.

Zawarte w korpusie teksty składają się z tzw. dialogów konstruowanych (*speech-related text types*) – zgodnie z założeniem, że zapis języka mówionego bliższy jest rzeczywistego języka danego czasu niż język pisany niebazujący na języku mówionym (Mitrenga i in., 2021, s. 170), można traktować je jako wiarygodne archiwum dawnych praktyk komunikacyjnych. Trafność tezy poświadczają liczne badania opierające się na CED (por. Csulich, 2022; Grund i Walker, 2021; Smitterberg, 2023; Walker i Kytö, 2022; Wang, 2019).

W pierwszym etapie zebrany materiał zbadam na poziomie pragmatycznym. Rozpatrywanie frazeologizmów (szerzej: frazemów) na podstawie paradygmatu pragmatycznego sugerował już w latach dziewięćdziesiątych XX w. Wojciech Chlebda (1993), w związku z czym uwaga badacza powinna być zwrócona zwłaszcza na intencję mówiącego i funkcję jednostki w wypowiedzi. Tym samym proponuję przeanalizowanie zgromadzonego materiału w kontekście trzech funkcji języka wyróżnionych przez Michalea Hallidaya (2002)⁵. Na potrzeby artykułu przyjmuję ich następujące rozumienie:

- funkcja interpersonalna: nadawca używa jednostki frazeologicznej w celu nawiązania, podtrzymania bądź zakończenia kontaktu, także w czynnościach perswazyjnych nakierowanych na odbiorcę;
- funkcja ideacyjna (treściowa): użycie jednostki frazeologicznej w wypowiedzi wyraża wyobrażenia nadawcy o rzeczywistym obiekcie lub sytuacji, w tym jego stosunek do nich;
- funkcja (meta)tekstowa: użycie jednostki frazeologicznej jest „ponad tekstem”; ujawnia się, przykładowo, w segmentacji toku komunikatu, przeformułowaniach (tworzeniu kontekstowych, sytuacyjnych innowacji frazeologicznych), sygnalizowaniu dygresji.

³ Tu: w dawnych dialogach konstruowanych, tj. literackich wypowiedziach naśladowujących żywą mowę badanego odcinka czasowego (zob. Mitrenga i in., 2021).

⁴ Działanie narzędzi do badania sentymetu wypowiedzi polega na identyfikacji i klasyfikacji nastawienia autora do tekstu. Metoda ta z powodzeniem jest stosowana nie tylko w branży komercyjnej, ale i na gruncie akademickim, zwłaszcza w naukach ekonomicznych, matematyczno-informatycznych czy socjologicznych. W ostatnich latach swoje zastosowanie odnajduje także w lingwistyce korpusowej (Gózdź-Roszkowski, 2018; Modzelewska i Węglarska, 2020; Smykała, 2020), co wraz z rosnącymi zasobami analizatorów sentymetu prognozuje przyrost opracowań językoznawczych w tym zakresie. Konsorcjum CLARIN-PL proponuje otwarte systemy sieciowe, których zaletą jest powszechna dostępność oraz łatwa obsługa. Należą do nich m.in. wykorzystujące głębokie modele uczenia *Multiemo* (Kocoń i in., 2021) i bazujące na polskim wordnecie (sic!) *Wydźwięk* (Janz i in., 2017).

⁵ Sugerowane kryteria wykorzystywali także inni badacze, por. Awdiejew, 2004; Bartmiński i Niebrzegowska-Bartmińska, 2009; Chlebda, 1993.

Znakowanie korpusu pozwala na pozyskanie informacji o cechach socjologicznych osób mówiących, które to cechy mogą być wykorzystane jako uzupełnienie badań o zmienne socjolingwistyczne, takie jak płeć, wiek czy status społeczny⁶. W kolejnym etapie przeprowadzoną analizę pragmalingwistyczną rozszerzę tym samym o dane na temat relacji pomiędzy użyciem jednostki frazeologicznej w określonej funkcji a pozycją społeczną nadawcy wypowiedzi. Posunięcie to stanowi egzemplifikację socjopragmatycznych badań zleksykalizowanych połączeń w tekstach dawnych, nieobecnych w dotychczasowej literaturze przedmiotu.

Ostatnim poziomem analizy jest weryfikacja nacechowania emocjonalnego komunikatów z wykorzystaniem narzędzia do automatycznego znakowania tekstów informacją o ich wydźwięku. Wśród dostępnych analizatorów warto zwrócić uwagę na otwarty system CLARIN-PL: *Wydźwięk*, który opiera się na polskim słowniku relacyjnym i stosowanej w nim anotacji emocji⁷. Przedstawiając spektrum zastosowania narzędzia, członkowie konsorcjum proponują instruktażowe badanie sentymentu trzech tekstów literackich (*Pana Tadeusza*, *Quo vadis* oraz *Nad Niemnem*). Czas powstania (XIX w.) i charakter utworów (m.in. obecność dialogów i słownictwa nacechowanego chronologicznie) skłoniły mnie do zweryfikowania poziomu zgodności analizy wydźwięku dokonanej przez system sieciowy oraz własne tzw. poczucie językowe (por. Janowska i Pastuchowa, 1995, s. 14).

Analizę nacechowania emocjonalnego zebranych wypowiedzi prowadzę na dwóch próbkach badawczych (podstawowej i kontrolnej), wyłączając z nich teksty, w których połączenia o charakterze frazeologicznym występowały jako przerywniki leksykalne, a zatem pojawiały się w większości kwestii tego samego nadawcy, a nie w określonych sytuacjach komunikacyjnych. Przygotowane fragmenty nie są „wyjęte” z kontekstu: ich długość pozwala ocenić wydźwięk komunikatu badaczowi, który dysponuje dostępnymi mu wiedzą językową, poczuciem językowym i kompetencjami językowymi⁸.

Charakterystyka materiału i ustalenia terminologiczne

Charakterystykę materiału poprzedzam fragmentem poświęconym ustaleniom terminologicznym, co jest konieczne ze względu na wieloaspektowość frazeologii i jej niejednorodne postrzeganie (Lewicki, 1974, s. 151). Przyjęcie tradycyjnego rozumienia *związku frazeologicznego* (którego znaczenie globalne różni się od sumy znaczeń elementów wchodzących w jego skład) wymagałoby nie tylko ograniczenia materiału do idiomów, lecz także ujęcia systemowego (stylistyczno-syntaktycznej interpretacji zdania), w którym proces komunikacji międzyludzkiej schodzi na dalszy plan. W odniesieniu do przedstawionych celów

⁶ Wszystkie dane zostały zaczerpnięte z KorTeDa. Moim celem nie jest szczegółowe referowanie kryteriów przyjętych przez Autorki korpusu, a jedynie wykorzystanie istniejących anotacji w prowadzonych badaniach.

⁷ *Wydźwięk* służy przetwarzaniu tekstów w trzech kategoriach nacechowania wypowiedzi: ogólnego nacechowania emocjonalnego, emocji podstawowych (Ekman, 1992; Plutchik, 1980; Wierzbicka, 1992a, b, c, d) oraz wartości uniwersalnych (Puzynina, 1992).

⁸ W próbkach – zarówno podstawowej, jak i kontrolnej – nie ma przykładów wypowiedzi, w których ocena nacechowania emocjonalnego wymagałaby znajomości innych fragmentów utworów.

badania trafniejsze wydaje się zatem przyjęcie szerszego – komunikacyjnego – postrzegania frazeologii, które zbliżone jest do definicji frazemu (jednostki odtwarzalnej sytuacyjnie). Proponowane tu połączenia (określane dalej wspólnym terminem *frazem*) charakteryzują się odtwórczością stałej struktury, wnoszeniem specyficznego kontekstu do wypowiedzi (za sprawą wyboru nadawcy z reguły dysponującego wielością środków językowych), a tym samym stylistyczną modyfikacją wypowiedzi. Choć nie ma konieczności, by tworzyły sens metaforyczny, stanowią jednak całość syntaktyczną i semantyczną, wnosząc nowe treści w stosunku do przypisanych elementów (por. Herman, 2004, s. 20–29).

Poza granicami terminologicznymi dyskusyjne pozostaje z całą pewnością odtworzenie okoliczności i momentu leksykalizacji danego połączenia wyrazowego. Trudne, a niejednokrotnie nierozstrzygalne jest bowiem stwierdzenie, że jednostka uważana dziś za s frazeologizowaną miała ten sam charakter dwa stulecia wcześniej. Z pomocą przychodzą źródła leksykograficzne, jednak i te mogą okazać się niewystarczające, jeśli porównać współczesne możliwości dostępu do źródeł z tymi, jakimi dysponował badacz w pierwszej połowie wieku XX⁹. Co więcej, każdorazowa koncentracja badań wokół odtwarzania genezy, autorstwa czy kanonicznej postaci frazeologizmu mogłaby skutkować ich nieustannym ograniczaniem do sond słownikowych i poszukiwań etymologicznych. Mając na uwadze przywołane głosy oraz trudności, jakie towarzyszą analizie dawnego materiału językowego, za kryterium uznania danej jednostki za frazeologiczną przyjmuję:

- względnie stabilny skład leksykalny: dana jednostka powinna być notowana w źródłach zawierających rejestrowaną frazeologię (szerzej: frazematykę) w tej samej, wariantywnej lub zbliżonej postaci;
- znaczenie, które wyraża właściwości uogólnione, tzn. że można odnieść je do różnych obiektów z poszczególnych klas.

Przedstawione wyznaczniki pozwalają zatem na analizę zarówno tzw. (potencjalnych) wariantów kanonicznych (notowanych w słownikach lub podawanych w nich jako pierwszy wariant), jak i tekstowych realizacji inwariantów, czyli konstrukcji synonimicznych dla formy podstawowej frazemu, które zachowują co najmniej częściowe podobieństwo w planie wyrażania z zachowaniem dowolności stylistyczno-pragmatycznej, np. *zdatny jak Żyd do roli* – *zdarny jak Żyd do wiązki siana* lub są kontaminacją dwóch stabilnych w języku połączeń, por. *żyć jak święty* – *żyć jak święty turecki*, w celu stworzenia zwrotu, który znaczeniowo odpowiada konstrukcji: *goły jak święty turecki*.

Ze zgromadzonego w korpusie materiału zostały wyizolowane dwie próbki badawcze. Grupa podstawowa składa się z 14 konstrukcji z komponentem etnonimicznym (wskazującym na przynależność narodową/etniczną)¹⁰. Kryterium doboru stanowi obecność etnonimu, dwu- lub wielowyrzawowość jednostki, jej częściowa metaforyzacja oraz frazeologiczność,

⁹ Nie znaczy to jednak, że są one niedoskonałe – szacuje się, że bazę materiałową *Słownika frazeologicznego języka polskiego* Stanisława Skorupki (SFJP) tworzy ok. 2000 źródeł pochodzących z SJPDor oraz zbiorów własnych Skorupki (Piotrowski, 2020, s. 517).

¹⁰ Badania frazeologii i frazematyki oparte na materiale etnonimicznym prowadzili m.in. Bernjak, 2013; Blagoeva i Sosnowski, 2022; Fomienkowa, 2009; Mrózek, 1994; Pławski, 2006; częściowo: Pacuła, 2012. Przeprowadzona kwerenda pozwoliła zaobserwować, że nie pojawiło się dotychczas pragmalingwistyczne opracowanie tego materiału.

która może być uzasadniona sytuacyjnie, tzn. odtwarzalna w podobnych okolicznościach. W większości przypadków jednostki są jednorazowo poświadczane w tekstach, tak jak inne frazemy znakowane w korpusie. Fakt ten zaważył na wyborze jednorodnej próbki badawczej. Jej ograniczona wielkość każe wstrzymać się przed formułowaniem wniosków wynikających z wyjątkowych analiz, w związku z czym świadomie pomijam je w swoim projekcie. Grupa kontrolna została zebrana na potrzeby weryfikacji badań z wykorzystaniem analizatorów sentymentu. Tworzą ją 4 warianty konstrukcji *jak Boga Kocham*¹¹. Dobór ten jest motywowany wielowariantowością jednostki oraz wieloma poświadczeniami każdego wariantu (łączenie 47). Dane dotyczące obu próbek znajdują się w tabeli 1 (grupa podstawowa) oraz tabeli 2 (grupa kontrolna) zamieszczonych po prezentowanej charakterystyce.

Analizę materiału rozpoczyna ujęcie systemowe, czyli pozyskanie materiału językowego, który następnie konfrontowany jest z danymi słownikowymi (SL; SWil; SFJP; WSFJP; NKPP i WSF PWN). Dane przedstawiam w formie tabelarycznej (zob. tab. 1). Na podstawie poświadczeń słownikowych rekonstruuje możliwe postaci wariantywne (znajdują się w kolumnie *Potencjalna postać kanoniczna frazemu*). Obok znajdują się dane o statusie społecznym nadawców. Pozyskanie tych ostatnich informacji stało się możliwe dzięki znakowaniu korpusu danymi socjologicznymi. Dostępność tej opcji jest jednym z czynników, które umożliwiają przekroczenie dotychczasowych ograniczeń w diachronicznych badaniach nad frazeologią, ponieważ rozwiązanie to pozwala m.in. na socjolingwistyczne analizy dawnego materiału językowego. Kolumna *Uwagi* zawiera dodatkowe informacje dotyczące badanych połączeń. Dane dotyczące grupy kontrolnej przedstawia tabela 2; dodatkowo, w *Uwagach* znajduje się informacja o liczbie użycia frazemów.

¹¹ Propozycja analizy funkcjonalnej formuły *jak Boga Kocham* i jej najczęściej występujących wariantów tekstowych została przedstawiona w artykule: *Jak frazeologizmy funkcjonują w różnych wariantach i kontekstach? Funkcjonalna analiza materiału historycznojęzykowego na podstawie jednostki jak Boga Kocham* (Ryś, 2023).

Tabela 1

Frazyemy z komponentem etnonimicznym (KorTeDa, dane na 24 kwietnia 2023 r.). Sonda słownikowa

Realizacja tekstowa	Potencjalna postać kanoniczna frazemu	Poświadczenia słownikowe	Nadawca wypowiedzi (status społeczny)	Uwagi (w tym: status społeczny przypisany nadawcy wypowiedzi)
(zadany) jak żyd do wiązki siana	zdatny jak żyd do roli/kosy/ wiązki siana/młocki	NKPP	Pawłowa (niski)	zdatny jak krowa do wierzchu; zdatny jak niedźwiedź do tańca; zdatny jak osieć do tańca; zdatny jak wół do karew; zdatny jak żyd do roli; jest zdatny, tak jak z byka dudy (NKPP III, s. 847)
[...] dzieci kiej u żyda czapek	miec dzieci jak żyd czapek	NKPP	Pawłowa (niski)	—
[...] modlą się do niej jak żydzi do miesiąca w pełni	modlić się do kogoś/czegoś jak żyd do miesiąca w pełni (?)	—	Cecylia (wysoki)	Księżyc: tyś jak miesiąc w pełni (NKPP II, s. 237)
Panny [...] rosną szybko, jak żydowska lichwa	rosnąć szybko jak żydowska lichwa	—	Onufry (średni)	zob. m.in. rosnąc: rosły jak po deszczu; rosną jak konopie; rośnie jak drzewo w lesie; rośnie jak grzyb przy drodze; rośnie jak na drodze; rośnie jak rżyisko na wiosnę (NKPP III, s. 70–71)
Włoczy się iak żyd	włóczyć się jak żyd (wieczny tułacz)	NKPP	Dojrzańska (średni)	włóczy się by cygan po świecie (zob. cygan SL); włóczyć się jak cygan po świecie (SWiI, cz.1, s. 180); żyd wieczny tułacz (SF II, s. 902)
Cygan świadczy się swojemi dziećmi	Świadczy się Cygan swoimi dziećmi/Cyganem	SL, SWiI, SFJP, NKPP, WSFJP	Pan Jowiński (wysoki)	—
Obędzie się cygańskie wesele przez marcypanu	Obędzie się cygańskie wesele bez marcepanów/marcypanów	SL, SWiI, SFJP, NKPP, WSFJP	Pawłowa (niski)	—
(zują) jak święci ale jak turecy	(żyć jak) święty turecki	SL, NKPP	Jędrzejowa (niski)	Turecki święty (SL V, s. 528); święty turecki (NKPP III, s. 549)

cd. tab. 1

Realizacja tekstowa	Potencjalna postać kanoniczna frazemu	Poświadczenia słownikowe	Nadawca wypowiedzi (status społeczny)	Uwagi (w tym: status społeczny przypisany nadawcy wypowiedzi)
[...] <i>goły iak turecki święty</i>	<i>goły jak święty turecki</i>	SL, SWił, SFJP, NKPP, WSFJP, WSF PWN	Aryst (średni)	realizacja w ramach: <i>święty turecki</i>
[...] <i>tak zdatny na (podstarościę), jak ja na sułtana tureckiego</i>	<i>zdatny na (kogos) jak (ktoś) [na sułtana tureckiego]</i>	—	Ekonom (średni)	—
<i>Złapał Kozak Tatarzynę, a Tatarzyn za łeb trzyma</i>	<i>Złapał Kozak Tatarzynę, a Tatarzyn za łeb trzyma</i>	SFJP, NKPP, WSF PWN	Pan Jowiński (wysoki)	—
<i>(czuję się) jak na niemieckim kazaniu</i>	<i>siedzieć/czuć się/być jak na niemieckim/tureckim/francuskim/sanskryckim kazaniu</i>	SL, SWił, SFJP, NKPP, WSFJP, WSF PWN	Pan Jowiński (wysoki) Cześnik (średni) Maryjan (średni)	—
<i>Matka Polka matak-Polek</i>	<i>Matka Polka / Matka-Polka / matka Polka / matka-Polka</i>	WSF PWN	Mączewski (średni)	—
<i>Wyglądasz jak śmierć angielska</i>	<i>wyglądać jak śmierć angielska / na chorągwi</i>	SFJP, NKPP, WSFJP	Juljastewiczowa (średni)	<i>wyglądać jak śmierć (SWił cz. 2, s. 1870); wyglądać jak śmierć na chorągwi (WSF PWN, s. 543)</i>

Źródło: Opracowanie własne.

Tabela 2

Wariantywność jednostki *jak Boga kocham* (KorTeDa, dane na 24kwietnia 2023 r.). Sonda słownikowa

Odtworzona forma frazemu	Obecność w słownikach	Uwagi
<i>jak Boga kocham</i>	SL, SWil, SFJP, WSFJP	<i>jak Bóg na niebie, jak mi Bóg miły</i> (NKPP I, s. 163–164); 19 przykładów realizacji
<i>jak cię kocham</i>	SFJP	10 przykładów realizacji
<i>jak Bożię kocham</i>	SFJP, WSFJP	8 przykładów realizacji
<i>jak babcię kocham</i>	SFJP	euf. zamiast <i>jak Boga kocham</i> (SF, s. 92); 5 przykładów realizacji

Źródło: Opracowanie własne.

Wyniki analizy

Wśród pozyskanych frazemów obserwuję 11 konstrukcji notowanych w słownikach: (zdarny) *jak Żyd do wiązki siana, dzieci kiej u żyda czapek, włoczy¹² się iak żyd, Cygan świadczy się swojemi dziećmi, wyglądasz jak śmierć angielska, Obędzie się cygańskie wesele przez marcypanu, (zyją) jak święci ale jak tureccy, goły iak Turecki święty, Złapał Kozak Tatarzyna, a Tatarzyn za łeb trzyma, czuję się jak na niemieckiem kazaniu / iakbym był na niemieckiem kazaniu/byłem jak na niemieckiem kazaniu, Matka Polka / matek-Polek, wyglądasz jak śmierć angielska* oraz 3 jednostki nienotowane w źródłach, które:

- są swobodnymi wariantami związków powszechnie znanych i poświadczonych przez opracowania frazeograficzne: *rosnąć szybko jak żydowska lichwa* (na wzór: *rosły jak po deszczu; rosną jak konopie; rośnie jak drzewo w lesie; rośnie jak grzyb przy drodze; rośnie jak na drożdżach; rośnie jak rżysko na wiosnę* (NKPP)) lub
- wyrażają właściwości uogólnione, tj. typowe dla różnych obiektów danej klasy lub dla obiektów innych klas, por. *modlą się do niej jak żydzi do miesiąca w pełni, zdatny na (podstarościgo), jak ja na sułtana tureckiego*.

Można zatem zauważyć, że analizowane połączenia są swobodnymi realizacjami (potencjalnych) postaci kanonicznych, których odrębność polega na:

- użyciu komponentu dialektalnego: *zdatny* → *zdarny* (dawniejsza forma), *jak* → *kiej* (tu także odmienny zakres użycia, por. *kiej* 'kiedy' (Wronicz, 2010, s. 99)), *bez* → *przez* (por. dawniejsza forma przyimka, zachowana w gwarach SJPDor), *żyją* → *zyją* (ślady mazurzenia);
- zastosowaniu krótszej postaci frazemu: *włóczyć się jak Żyd wieczny tułacz* (NKPP) → *włoczy się iak żyd*;
- zmianie kolejności elementów związku: *goły jak święty turecki* (SL; SWil; SFJP; NKPP; WSFJP; WSF PWN) → *goły iak Turecki święty*;

¹² We wszystkich cytatach i realizacjach tekstowych zachowano oryginalną pisownię.

- skonstruowaniu wariantu tekstowego na podstawie znanego modelu składniowego: *modlą się do niej jak żydzi do miesiąca w pełni, tak zdatny na (podstarościego), jak ja na sułtana tureckiego, (panny) rosną szybko, jak żydowska lichwa*, w tym tzw. kontekstowej innowacji frazeologicznej (*żyć jak święty* → *żyć jak święty turecki*, w celu stworzenia zwrotu, który znaczeniowo odpowiada konstrukcji: *goły jak święty turecki*) powstałej wskutek żartu sytuacyjnego, por.

MARYSIA. Prawda że za wszyckich robi, ale tez ojcowie nic nie robią, a *zyją jak święci*.

JĘDRZEJOWA. A juści *jak święci ale jak tureccy* – bo nic nie mają (KorTeDa)

Mając na uwadze zmiany w rozwoju polskiej pisowni oraz ortografii, za zgodne z wariantem kanonicznym uznaję realizacje: *Cygan świadczy się swojemi dziećmi, Złapał Kozak Tatarzyna, a Tatarzyn za łeb trzyma, (czują się) jak na niemieckiem kazaniu / iakbym był na niemieckiem kazaniu / byłem jak na niemieckiem kazaniu, Matka Polka / matek-Polek, wyglądasz jak śmierć angielska*. Można zatem zauważyć, że w badanej próbie przeważającą część materiału stanowią modyfikacje połączeń traktowanych dziś jako normatywne. Trudno jednoznacznie stwierdzić, czy jest to wynik intencjonalnego i świadomego działania nadawcy, czy też dopiero proces kształtowania się postaci kanonicznych.

Kierowanie się kryteriami przyjętymi w ramach analizy funkcjonalnej (por. *Założenia metodologiczne*) pozwala zauważyć, że w badanej próbie frazemy najczęściej używane są w celu wyrażenia przez nadawcę wyobrażeń o rzeczywistym obiekcie lub sytuacji, a zatem w funkcji ideacyjnej (*zdatny jak Żyd do roli / kosy / wiązki siana / młocki; mieć dzieci jak Żyd czapek; modlić się do kogoś / czegoś jak Żyd do miesiąca w pełni; rosnąc szybko jak żydowska lichwa; włączyć się jak Żyd (wieczny tułacz); goły jak święty turecki; (żyć jak) święty turecki; wyglądać jak śmierć angielska / na chorągwi; siedzieć / czuć się / być jak na niemieckim/tureckim/francuskim/sanskryckim kazaniu; Świadczy się Cygan swoimi dziećmi/Cyganem; Obejdzie się cygańskie/żydowskie wesele bez marcepanów/marcypanów; Złapał Kozak Tatarzyna, a Tatarzyn za łeb trzyma; Matka Polka / Matka-Polka / matka Polka / matka-Polka; zdatny na podstarościego jak (ktoś) na sułtana tureckiego*). Wynika to ze specyfiki strukturalnej związków; dominują tu bowiem konstrukcje porównawcze, których celem jest przedstawienie zależności między obiektem porównywanym i porównującym (por. Greszczuk, 1988; Ścibek-Trypus, 2017). Warto zauważyć, że w badanym kontekście pojawiają się pełniące funkcję pełnej wypowiedzi: konstrukcja komparatywna (*zdatny jak Żyd do roli / kosy / wiązki siana / młocki*) i trzy połączenia tradycyjnie klasyfikowane jako przysłowia (*Obejdzie się cygańskie/żydowskie wesele bez marcepanów/marcypanów; Świadczy się Cygan swoimi dziećmi / Cyganem; Złapał Kozak Tatarzyna, a Tatarzyn za łeb trzyma*)¹³, a także jedno wyrażenie (*Matka Polka / Matka-Polka / matka Polka / matka-Polka*), za którego pomocą autor przenosi archetypiczne wyobrażenia z frazemu na kobietę, o której się wypowiada.

Materiał korpusowy uprawnia do stwierdzenia, że użycie przytoczonych konstrukcji w omawianym kontekście jest charakterystyczne zwłaszcza dla średniej warstwy społecznej

¹³ Na potrzeby analizy przyjmuję szerokie rozumienie frazeologii; do materiału włączam zatem połączenia, które tradycyjnie są przez niektórych badaczy uznawane za przysłowia – przedmiot paremiologii.

(8 jednostkowych użyc) – stanowi dwukrotność przykładów pochodzących z wypowiedzi klas niskiej (4) i wysokiej (3). Warto przy tym zwrócić uwagę na to, że w komunikatach reprezentantów zarówno niższych, jak i wyższych warstw funkcja ta ujawnia się w jednostkowych wypowiedziach. W pierwszej grupie na 4 użycia aż 3 realizuje Pawłowa. Analogiczną sytuację obserwuje się w drugiej grupie – 3 z 4 przykładów są wyrażane przez Jowialskiego. Cechą charakterystyczną obu postaci jest częste używanie frazemów, w związku z czym dane dotyczące rzeczywistej realizacji jednostek przez reprezentantów niskiej i wysokiej klasy społecznej należałoby rozszerzyć o dodatkowe przykłady.

Funkcja interpersonalna, odpowiedzialna za fatyczne elementy kontaktu, ujawnia się w użyciu konstrukcji, które występują jako niezależne wypowiedzenie (*Złapał Kozak Tatarzyna, a Tatarzyn za łeb trzyma*) lub kończą komunikat: w badanych dialogach zostały użyte w kontekście opiniującym wygląd (*wyglądać jak śmierć angielska / na chorągwi*) i zachowanie rozmówcy (*Świadczy się Cygan swoimi dziećmi/Cyganem*). Funkcja ta wyraża się w wypowiedziach osób o wysokiej (dwukrotnie, Pan Jowialski) i średniej (raz, Juljasiewiczowa) pozycji społecznej. Wynik ten pokazuje, że używanie konstrukcji z elementem etnonimicznym w funkcji interpersonalnej jest rzadkie i, być może (!), bliższe mowy średnich i wyższych warstw społecznych. Ze względu na niewielką próbkę materiałową oraz bardzo częste przywoływanie frazemów przez postać Jowialskiego obszar ten domaga się pogłębionych badań i na obecnym etapie każe wstrzymać się przed formułowaniem wniosków.

Na poziomie wewnątrztekstowym, tu w charakterze przeformułowania, występuje jedno połączenie (*żyć jak*) *święty turecki*, którego realizacja jest, jak zostało to zasygnalizowane (zob. *Założenia metodologiczne*), kontaminacją dwóch porównań: *żyć jak święty* i *goły jak święty turecki*. Pojawia się ono w wypowiedzi postaci o niskim statusie społecznym – chłopki Jędrzejowej.

Można zatem zauważyć, że w przedstawionej próbce frazemy najczęściej realizowane są w funkcji treściowej, na co wpłynęła m.in. przewaga konstrukcji porównawczych (6). Co ciekawe, najczęściej pojawiają się one w wypowiedziach postaci o średnim statusie społecznym, nie zaś, jak można by się było spodziewać, niskim¹⁴.

Równolegle do analizy pragmalingwistycznej badam frazemy pod kątem wydzwieku emocjonalnego wypowiedzi, do których są przywoływane. Kierując się własnym poczuciem językowym, obserwuję zgodność wyników analizy z prawdopodobnym rzeczywistym wydzwiekiem wypowiedzi na poziomie ok. 56% dla analizatora sentymentu *Wydzwiek*. Rezultaty ewaluacji są zatem niższe od szacunkowych obliczeń poprawnej identyfikacji wydzwieku w tekstach dawnych (ok. 70%), jaką sugerują twórcy narzędzia¹⁵. W grupie podstawowej na 17 wypowiedzi badanych za pomocą narzędzia ocena 6 z nich (zawierających konstrukcje: *modlić się do kogoś/czegoś jak Żyd do miesiąca w pełni, Obejdzie się cygańskie/żydow-*

¹⁴ Porównania są nieodłącznym elementem konkretności – jednego z podstawowych wyznaczników polszczyzny potocznej, zwłaszcza w wymiarze historycznym (por. Umińska-Tytoń, 1992, s. 230). Cecha ta manifestuje się także w wartościowaniu czy unikaniu pojęć abstrakcyjnych – myśleniu konkretnym – a zatem fundamentalnych składnikach językowo-kulturowego obrazu wsi (zob. Piechnik, 2009, s. 38–39).

¹⁵ Jest to wartość szacunkowa, którą Piotr Miłkowski zasugerował w trakcie *XII warsztatów CLARIN w praktyce badawczej* (Miłkowski, 2022), odpowiadając na pytanie: *Czy narzędzie może być wykorzystane w badaniu tekstów dawnych?*

skie wesele bez marcepanów/marcypanów, (żyć jak) święty turecki, zdalny na (kogoś) jak (ktoś) na [sułtana tureckiego], siedzieć / czuć się / być jak na niemieckim/tureckim/francuskim/sanskryckim kazaniu, wyglądać jak śmierć angielska / na chorągwi) wzbudza wątpliwości, sugerując wyższe oceny w kategoriach pozytywnego wydzźwięku, niż mogłoby to wynikać z kontekstu. Z kolei w grupie kontrolnej wśród badanych 35 wypowiedzi w odniesieniu do 10 można podejrzewać błędny wynik (6/19 – 32% w realizacji wariantu *jak Boga kocham*, 6/10 – 60% – wariant *jak cię kocham*, 6/8 – 75% – wariant *jak Bożę kocham*). Ponadto w tej grupie występują także wypowiedzi (9), którym nie zostały przypisane wielkości ogólnego nacechowania, poziomów emocji podstawowych ani wartości uniwersalnych – zbiór ten stanowi 24% materiału w grupie kontrolnej. Brak parametrów, takich jak poziomy emocji podstawowych czy wartości uniwersalnych, uniemożliwia tym samym traktowanie takich wypowiedzi jako neutralnych. Uśrednione wyniki analizy¹⁶ przedstawia tabela 3:

Tabela 3

Uśrednione wyniki oceny sentymentu wypowiedzi z wykorzystaniem analizatora Wydzźwięk

Grupa podstawowa (badane konstrukcje w postaci rekonstruowanej)	Średnia przybliżonego wyniku oceny sentymentu	Grupa kontrolna (badane konstrukcje)	Średnia przybliżonego wyniku oceny sentymentu
<i>zdalny jak Żyd do roli / kosy / wiązki siana / młocki</i>	-0,5		
<i>mieć dzieci jak Żyd czapek</i>	-0,5		
<i>modlić się do kogoś/czegoś jak Żyd do miesiąca w pełni</i>	+1,5*		
<i>rosnąć szybko jak żydowska lichwa</i>	+0,5		
<i>włóczyć się jak Żyd (wieczny tułacz)</i>	-5	<i>jak Boga kocham</i>	+0,31 (6*)
<i>Świadczy się Cygan swoimi dziećmi/Cyganem</i>	-1	<i>jak cię kocham</i>	? (6*)
<i>Obejdzie się cygańskie/żydowskie wesele bez marcepanów/marcypanów</i>	+1,5*	<i>jak Bożę kocham</i>	+0,33 (6*)
<i>(żyć jak) święty turecki</i>	+0,5*		
<i>goły jak święty turecki</i>	-1		
<i>zdalny na (kogoś) jak (ktoś) na [sułtana tureckiego]</i>	+0,5*		
<i>Złapał Kozak Tatarzyną, a Tatarzyn za łeb trzyma</i>	+0,5		
<i>siedzieć / czuć się / być jak na niemieckim/tureckim/francuskim/sanskryckim kazaniu</i>	-0,17 (1*)		
<i>Matka Polka / Matka-Polka / matka Polka / matka-Polka</i>	+3		
<i>wyglądać jak śmierć angielska / na chorągwi</i>	+1*		

Źródło: Opracowanie własne.

Przedstawione przybliżone wartości sentymentu wypowiedzi są oszacowane na przez analizator *Wydzźwięk* na podstawie średniej arytmetycznej ogólnego nacechowania emo-

¹⁶ Średnia arytmetyczna wartości uzyskanych dzięki zastosowaniu narzędzia.

cjonalnego tekstu. Symbol „+” oznacza dominację nacechowania pozytywnego, „-” – negatywnego, a „?” – niejednoznacznego, tzn. ambiwalentnego lub niemożliwego do zidentyfikowania. Znakiem „*” opatruję pozycje, w których ewaluacja wydźwięku wypowiedzi przeprowadzona za pomocą narzędzia odbiegała od tej, na jaką wskazywałoby poczucie językowe; jak sygnalizowałam, jest to 6 przypadków w grupie podstawowej i 19 – w grupie kontrolnej. Obserwuje się zatem trafność ewaluacji dawnego materiału dramatycznego na poziomie odpowiednio 65% i 46%. Przyczyn rozbieżności oceny prowadzonej za pośrednictwem narzędzia z przewidywanym stanem rzeczywistym można upatrywać z jednej strony w niedoskonałościach systemu, np. braku umiejętności rozpoznawania ironii, zmianie nacechowania emotywnego lub znaczeń wyrazów (zwłaszcza ich rozszerzenia lub zawężenia), z drugiej – w braku odpowiedniej kompetencji językowej badacza: wiedza o użyciu języka we wcześniejszych stadiach rozwoju jest ograniczona do tzw. poczucia językowego (por. Janowska i Pastuchowa, 1995).

Zakończenie

Rozwój technologii cyfrowych oraz prace nad korpusem służącym gromadzeniu dialogów naśladujących dawny język mówiony dają asumpt do rozszerzenia diachronicznych studiów nad frazeologizmami o funkcjonalną analizę związków wyrazowych. Skierowanie perspektywy badań na użycie frazeologizmów w dialogach konstruowanych (*speech-related text types*) umożliwia identyfikację intencji nadawców. W analizowanej próbie najczęściej realizowana jest funkcja ideacyjna (co najmniej jednokrotne użycie każdej konstrukcji), co może wynikać z jej rozumienia i założeń (przedstawienie wyobrażeń nadawcy o danym obiekcie lub sytuacji i jego stosunku do nich) oraz obecności wielu porównań w zbiorze. Materiał poświadcza także reprezentację dwóch pozostałych funkcji – interpersonalnej (3) i (meta)tekstowej (1). Zidentyfikowanie wszystkich wskaźników pragmatylingwistycznych na tak niewielkiej próbie tekstowej może zatem wskazywać na zasadność wykorzystania zaproponowanej przez Hallidaya (2002) klasyfikacji.

Zwrócenie uwagi na płaszczyznę socjojęzykową może z kolei prowadzić do uzyskania cennych wniosków o zależnościach między kontekstem użycia jednostki a poszczególnymi wskaźnikami socjologicznymi interlokutorów. Obserwacja materiału korpusowego daje podstawy do konkluzji, że używanie frazemów w funkcji ideacyjnej jest charakterystyczne dla średnich warstw społecznych. Bieżący stan badań nie pozwala na wysunięcie podobnych wniosków w odniesieniu do pozostałych funkcji czy warstw społecznych. Obszar ten wymaga pogłębionych analiz: z jednej strony powinny one bazować na materiale wykraczającym poza KorTeDa, z drugiej zaś – spełniać kryteria, które przyjęły Autorki korpusu (Mitrena i in., 2021, s. 174–176).

Choć kontekstowa analiza wydźwięku z wykorzystaniem systemu do anotacji emocji tekstu nie przyniosła satysfakcjonujących wyników (trafność oceny < 70%), dostrzegam szansę we włączeniu podobnych systemów do przyszłych badań nad historycznym materiałem językowym. Główną tego przesłankę stanowi tworzenie narzędzi na podstawie modeli uczenia głębokiego, których efektywność jest skorelowana z ich eksploatacją oraz aplikacją dawnego materiału językowego do możliwości narzędzi technologicznych. Działania zmie-

rzające ku doskonaleniu systemów, jak również prace nad znakowanym socjopragmatycznie KorTeDa otwierają perspektywę funkcjonalnych analiz dawnych jednostek frazeologicznych.

Słowniki

- NKPP – Krzyżanowski, J. (red.). (1969–1978). *Nowa księga przysłów i wyrażeń przysłowiowych polskich* (t. 1–4). Państwowy Instytut Wydawniczy.
- SFJP – Skorupka, S. (1967–1968). *Słownik frazeologiczny języka polskiego* (t. 1–2). Wiedza Powszechna.
- SJPDor – Doroszewski, W. (red.). (1958–1969). *Słownik języka polskiego*. Wiedza Powszechna. <https://doroszewski.pwn.pl/>
- SL – Linde, S. B. (1807–1814). *Słownik języka polskiego* (t. 1–6). Drukarnia Zakładu Ossolińskich.
- SWil – Zdanowicz, A., Bohusz Szyszko, M., Filipowicz, J., Tomaszewicz, W., Czepieliński, F., Korotyński, W., Trentowski, B. (red.). (1861). *Słownik języka polskiego*. Wydane nakładem i staraniem Maurycego Orgelbranda.
- WSF PWN – Kłosińska, A., Sobol, E., Stankiewicz, A. (red.). (2010). *Wielki słownik frazeologiczny PWN z przysłowiami*. Wydawnictwo Naukowe PWN.
- WSFJP – Müldner-Nieckowski, P. (2004). *Wielki słownik frazeologiczny języka polskiego*. Świat Książki.

Literatura

- Awdiejew, A. (2004). *Gramatyka interakcji werbalnej*. Wydawnictwo Uniwersytetu Jagiellońskiego.
- Bartmiński, J., Niebrzegowska-Bartmińska, S. (2009). *Tekstologia*. Wydawnictwo Naukowe PWN.
- Bernjak, E. (2013). Etonimi v slovenski in madžarski frazeologii. *Studia Slavica Savariensia*, 1–2, 100–110. <https://doi.org/10.17668/SSS.2013.1-2.100>
- Blagoeva, D., Sosnowski, W. (2022). Lakunarność we frazeologii na przykładzie języka bułgarskiego i polskiego. *Językoznawstwo*, 2(17), 49–56. https://doi.org/10.25312/2391-5137.17/2022_04wsdb
- CED – Kytö, M., Walker, T. (2006). *Guide to A Corpus of English Dialogues, 1560–1760*. Uppsala Universitet.
- Chlebda, W. (1993). Frazematyka. W: J. Bartmiński (red.), *Encyklopedia kultury polskiej XX w.* (t. 2: *Współczesny język polski*, s. 327–334). Wiedza o Kulturze.
- Csulich, G. (2022). *(Im)politeness and power in the Early Modern English courtroom (1560 to 1639)* [rozprawa doktorska. Birmingham City University]. <https://www.open-access.bcu.ac.uk/13441/>
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3–4), 169–200.
- Fomienkowa, M. (2009). *Językowy obraz świata Polaka przez pryzmat frazeologii z komponentem etnonimicznym* [manuskrypt]. <https://web.archive.org/web/20180502034241/http://elib.bsu.by/bitstream/123456789/89915/1/Marina%20Fomienkowa%20Jezykowy%20obraz%20swiata%20polaka%20przez%20pryzmat%20frazeologii%20z%20komponentem%20etnonimicznym.pdf>
- Gózdź-Roszkowski, S. (2018). Analiza języka wartościowania w dyskursie sądowym z perspektywy językoznawstwa korpusowego: Możliwości i ograniczenia. W: P. Stalmaszczyk (red.), *Metodologie językoznawstwa. Od diachronii do panchronii* (s. 83–102). Wydawnictwo Uniwersytetu Łódzkiego.
- Greszczuk, B. (1988). *Konstrukcje porównawcze i ich rozwój w języku polskim*. Wydawnictwo Wyższej Szkoły Pedagogicznej.

- Grund, P. J., Walker, T. (2021). Speech representation in the history of English. Introduction. W: P. J. Grund, T. Walker (red.), *Speech representation in the history of English. Topics and approaches* (s. 1–28). Oxford University Press.
- Gruszczyński, W. (2020). Korpusy językowe narzędziem pracy historyka języka. W: K. Kłosińska, A. Hącia, P. Zbróg (red.), *Polszczyzna w dobie cyfryzacji* (s. 255–265). Polska Akademia Nauk.
- Gruszczyński, W., Adamiec, D., Bronikowska, R., Kieraś, W., Modrzejewski, E., Wieczorek, A., Woliński, M. (2022). The electronic corpus of 17th- and 18th-century Polish texts. *Language Resources and Evaluation*, 56, 309–332. <https://doi.org/10.1007/s10579-021-09549-1>
- Halliday, M. A. K., (2002). *On grammar*. Continuum.
- Herman, W. (2004). Frazem – próba definicji. *Poradnik Językowy*, 7, 20–29.
- Ignatowicz-Skowrońska, J. (2020). Założenia słownika etymologicznego polskich frazeologizmów. *Język Polski*, 100(4), 118–124.
- Janowska, A. Pastuchowa, M. (1995). *Niebezpieczna kompetencja*. Poradnik Językowy, 8, 11–20.
- Janz A., Kocoń J., Piasecki M., Zaśko-Zielińska, M. (2017). plWordNet as a basis for large emotive lexicons of Polish. Referat wygłoszony w ramach LTC'17 8th Language & Technology Conference. Poznań.
- Jawór, A. (2009). Zmiany znaczeniowe związków frazeologicznych. *LingVaria*, 1(7), 91–99.
- Kocoń, J., Miłkowski, P., Kanclerz, K. (2021). *MultiEmo. Multilingual, multilevel, multidomain sentiment analysis corpus of consumer reviews*. <http://hdl.handle.net/11321/798>.
- Kosek, I., Zawilska, K. (2012). Osobliwości leksykalne w składzie związków frazeologicznych (na wybranych przykładach). *Język Polski*, 92(4), 260–268.
- Lewicki, A. M. (1974). Aparat pojęciowy frazeologii. W: L. Ludorowski (red.), *Z badań nad literaturą i językiem* (s. 135–151). PWN.
- Łaziński, M., Górski, R. L., Woźniak, M. (2023). Korpus XIX w. Uniwersytetu Warszawskiego i IJP PAN. *LingVaria*, 35(1), 125–134. <https://doi.org/10.12797/LV.18.2023.35.09>
- Miłkowski, P. (9–10 grudnia 2022). *Narzędzia analizy wydziwisku emocjonalnego* [referat]. XII warsztaty CLARIN-PL w praktyce badawczej, Katowice, Polska.
- Mitrenga, B., Pastuch, M., Wąsińska, K. (2021). Możliwości i ograniczenia historycznych badań pragmatyngwistycznych. W: M. Mączyński, E. Horyń, E. Zmuda (red.), *W kręgu dawnej polszczyzny*, 7 (s. 163–181). Wydawnictwo Naukowe Akademii Ignatianum w Krakowie.
- Modzelewska, M., Węglarska, K. (2020). Emocje w Sieci. Analiza internetowych dyskusji związanych z rocznicą odzyskania niepodległości. W: K. Kłosińska, A. Hącia, P. Zbróg (red.), *Polszczyzna w dobie cyfryzacji* (s. 47–57). Wydawnictwo Naukowe PWN.
- Mrózek, R. (1994). Frazemy z etnonimami i chrononimami w świetle kulturowej teorii języka. *Opera Slavica*, 4(4), 35–40.
- Opaliński, K., Potoniec, P., Eder, M., Kieraś, W., Woliński, M., Górski, R., Derwojedowa, M., Gruszczyński, W., Król, M. (2019). Narodowy Korpus Diachroniczny Polszczyzny. Projekt. *Język Polski*, 99(1), 92–101.
- Pacufa, J. (2012). Polskie i rosyjskie egzoetnonimy i przezwiska Żyda w kontekście stereotypu językowego. *Linguarum Silva*, 1, 135–148.
- Pastuch, M., Duda, B., Lisczyk, K., Mitrenga, B., Przyklenk, J., Sujkowska-Sobisz, K. (2018). Digital humanities in Poland from the perspective of the historical linguist of the Polish language. Achievements, needs, demands. *Digital Scholarship in the Humanities*, 33(4), 857–873. <https://doi.org/10.1093/llc/fqy008>
- Pastuch, M., Mitrenga, B., Wąsińska, K. (2023). From adverb to intensifier. Corpus-based research in diachronic linguistics on the example of the Polish words okrutnie ('cruelly'), strasznie ('terri-

- bly') and szalenie ('madly'). *Journal of Historical Pragmatics*, 23(6), 285–326. <https://doi.org/10.1075/jhp.19009.pas>
- Pastuch, M., Mitrenga, B., Wąsińska, K. (2024a). Anotacja socjolingwistyczna w Korpusie dawnych polskich tekstów dramatycznych (1772–1939). *LingVaria*, 19, 1(37), 151–170. <https://doi.org/10.12797/LV.19.2024.37.10>
- Pastuch, M., Mitrenga, B., Wąsińska, K. (2024b). *Przerywnik leksykalny w historycznojęzykowych badaniach socjopragmatycznych (na materiale Korpusu dawnych polskich tekstów dramatycznych (1772–1939))*. *Język Polski*, 104(1), 93–110. <https://doi.org/10.31286/JP.00488>
- Piechnik, A. (2009). *Wizerunek kobiety i mężczyzny w językowym obrazie świata ludności wiejskiej (na przykładzie gminy Zakliczyn nad Dunajcem)*. Wydawnictwo Uniwersytetu Jagiellońskiego.
- Piela, A. (2018). *Słownik frazeologizmów z archaizmami. Pamiątki przeszłości*. Wydawnictwo Naukowe PWN.
- Piela, A. (2023). Nazwy miejscowe Kresów Wschodnich w zapomnianej polskiej frazeologii. *Język Polski*, 103(3), s. 101–113. <https://doi.org/10.31286/JP.00222>
- Piotrowski, T. (2020). *Słowniki polskie: Słownik frazeologiczny Stanisława Skorupki*. W: W. Moki-jenko, J. Tarsa, *Język i pamięć. Księga jubileuszowa dedykowana Panu Profesorowi Wojciechowi Chlebdzie z okazji 70. urodzin* (s. 509–519). Wydawnictwo Uniwersytetu Opolskiego.
- Plutchik, R. (1980). *Emotion. A psychoevolutionary synthesis*. Harper and Row.
- Pławski, M. (2006). O tłumaczeniu jako czynności uwarunkowanej kulturowo na przykładzie frazeologizmów zawierających etnonimy. *Rocznik Przekładoznawczy*, 2, 181–190. <https://doi.org/10.12775/RP.2006.014>
- Przymuszała, L. (2018). O potrzebie badań nad frazeologią historyczną. W: M. Pastuch, M. Siuciak (red.), *Historia języka w XXI wieku. Stan i perspektywy*. Wydawnictwo Uniwersytetu Śląskiego, s. 299–311.
- Puda-Blokesz, M. (2021). Mitologizmy frazeologiczne w najnowszych ujęciach lingwistycznych. W: E. Horyń, E. Młynarczyk, P. Żmigrodzki (red.), *Język polski – między tradycją a współczesnością. Księga jubileuszowa z okazji stulecia Towarzystwa Miłośników Języka Polskiego* (s. 150–161). Wydawnictwo Naukowe Uniwersytetu Pedagogicznego im. Komisji Edukacji Narodowej w Krakowie.
- Puzynina, J. (1992). *Język wartości*. Wydawnictwo Naukowe PWN.
- Ryś, A. (2023). Jak frazeologizmy funkcjonują w różnych wariantach i kontekstach? Funkcjonalna analiza materiału historycznojęzykowego na podstawie jednostki *jak Boga Kocham*. *LingVaria*, 18(2), 323–337.
- Schreibman, S., Siemens R., Unsworth J. (2004). The digital humanities and humanities computing. An introduction. W: S. Schreibman, R. Siemens, J. Unsworth (red.), *A companion to digital humanities*. Blackwell Publishing. https://companions.digitalhumanities.org/DH/?chapter=content/9781405103213_intro.html
- Smutterberg, E. (2023). Using very large corpora to teach modern English (1500–1945). *Nordic Journal of English Studies*, 93–112.
- Smykała, M. (2020). Style perswazji w Europie na przykładzie komentarzy w prasie codziennej. W: H. E. H. Lenk, H. W. Giessen (red.), *Persuasionsstile in Europa. Strategien und Mittel des Überzeugens in Zeitungskomentaren aus kontrastiver Sicht* (s. 377–386). Frank & Timme.
- Ścibek-Trypuz, J. (2017). Konstrukcje porównawcze – przegląd ujęć typologicznych. W: U. Sokółska (red.), *Socjolekt, idiolekt, idiostyl. Historia i współczesność* (s. 387–409). Prymat.
- Umińska-Tytoń, E. (1992). *Polszczyzna potoczna XVIII wieku*. Wydawnictwo Uniwersytetu Łódzkiego.

- Walker, T., Kytö, M. (2022). Survival or death: mine/my and thine/thy variation in Early Modern English medical writing. W: T. Hiltunen, I. Taavitsainen (red.), *Corpus pragmatic studies on the history of medical studies* (s. 105–126). John Benjamins Publishing Company.
- Wang, Y. (2019) A corpus-based study of composite predicates in Early Modern English dialogues. *Journal of Historical Pragmatics*, 1, 20–50.
- Węgrzynek, K. (2007). Związki frazeologiczne w *Wielkim słowniku języka polskiego*. W: P. Żmigrodzki, R. Przybylska (red.), *Nowe studia leksykograficzne* (t. 1, s. 87–100). Lexis.
- Wierzbicka, A. (1992a). Are emotions universal or culture specific? W: A. Wierzbicka, *Semantics, culture and cognition. Universal human concepts in culture-specific configurations* (s. 119–138). Oxford University Press.
- Wierzbicka, A. (1992b). *Are there basic human emotions?* [manuskrypt].
- Wierzbicka, A. (1992c). Defining emotion concepts. *Cognitive Science*, 16, 539–581.
- Wierzbicka, A. (1992d). Describing the indescribable. W: A. Wierzbicka, *Semantics, culture and cognition. Universal human concepts in culture-specific configurations* (s. 139–179). Oxford University Press.
- Wronicz, J. (2010). *Mały słownik gwar polskich*. Lexis.