

Uniwersytet im. Adama Mickiewicza w Poznaniu

Wydział Filologii Polskiej i Klasycznej

mgr Marta Kunegunda Witkowska

Analiza warsztatu pisarskiego Julii Hartwig

metodami lingwistyki komputerowej

Promotor:

prof. dr hab. Zygmunt Vetulani

Promotor pomocniczy:

prof. UAM dr hab. Marcin Telicki

Poznań, 2021

*Moim Babciom – Marcie i Kunegundzie,  
Mateuszowi i naszym dzieciom: Helence, Janince i Felkowi.*

# SPIS TREŚCI

STRESZCZENIE	6
ABSTRACT	7
RESUMÉ	8
PRZEDSŁOWIE. DLACZEGO JULIA HARTWIG?	9
WSTĘP. PROBLEM I METODY BADAWCZE	13
ROZDZIAŁ 1	
Twórczość Julii Hartwig na tle metodologii badań nad językiem i stylem pisarzy	16
1.1. Badania nad językiem pisarzy	16
1.2. Metody komputerowe w badaniach nad językiem pisarzy	18
1.2.1. Dane językowe – korpusy tekstów	19
1.2.2. Wybrane metody inżynierii języka	20
1.2.3. Metody ilościowe	21
1.3. Uwagi o języku Julii Hartwig – rozpoznanie	21
ROZDZIAŁ 2	
Przymiotnik wyznacznikiem warsztatu Julii Hartwig	23
2.1. Opis języka naturalnego metodami formalnymi	24
2.2. Różne realizacje predykatu w zdaniu	26
2.2.1. Elementy struktury predykatywno-argumentowej	28
2.2.2. Formy predykcji przymiotnikowej	32
2.3. Świat Julii Hartwig odzwierciedlony w przymiotnikach	39
ROZDZIAŁ 3	
Korpus prozy i korpus poezji Julii Hartwig	44
3.1. Korpusy tekstowe i korpusy elektroniczne pod kątem lingwistyki korpusowej	44
3.1.1. Pojęcie korpusu i jego znaczenie	44
3.1.2. Podstawowe cechy korpusów	47
3.1.3. Korpusy tekstowe na świecie	51
3.1.4. Elektroniczne korpusy tekstowe dla języka polskiego	52
3.2. Po co nam elektroniczny korpus twórczości Julii Hartwig?	54
3.3. Charakterystyka korpusu prozy i poezji Julii Hartwig	55
3.3.1. Korpus prozy	56
3.3.2. Korpus poezji	57
3.3.3. Reprezentatywność i zrównoważenie korpusu prozy i poezji Julii Hartwig	59

3.4. Tworzenie korpusu prozy i poezji Julii Hartwig	60
3.4.1. Metoda pozyskiwania korpusu	60
3.4.2. Opracowanie danych początkowych	61
3.4.3. Etapy normalizacji korpusu prozy i poezji Julii Hartwig	62
3.4.4. Złożoność normalizacji i jakość korpusu	66
3.5. Rozwiązane problemy związane z pozyskiwaniem korpusu w formie elektronicznej	66
3.5.1. Pozyskanie elektronicznego korpusu twórczości Julii Hartwig	66
3.5.2. Skanowanie materiału książkowego	67
3.5.3. Sprawdzanie zgodności oryginału z wersją OCR	69
<b>ROZDZIAŁ 4</b>	
Wykorzystanie metod automatycznego przetwarzania tekstu do analizy przymiotnikowej strukturalizacji predykatu w korpusie prozy i poezji Julii Hartwig	70
4.1. Wykorzystane narzędzi przetwarzania języka naturalnego	70
4.1.1. Formalizm i lematyzator POLEX	71
4.1.2. Biblioteka NLTK	73
4.1.3. Korpusomat	73
4.1.4. Text SubCorpora Creator 1.3. (TSCC)	73
4.2. Zastosowanie narzędzi informatycznych w procesie przetwarzania korpusu prozy i poezji Julii Hartwig	74
4.2.1. Oznaczenie (lematyzacja) przymiotników w korpusach prozy i poezji Julii Hartwig oraz pozyskanie list referencyjnych przymiotników	75
4.2.1.1. Lematyzacja	75
4.2.1.2. Tworzenie list referencyjnych	79
4.2.2. Badania ilościowe przymiotników w korpusie prozy i poezji Julii Hartwig	80
4.2.2.1. Formy przymiotnikowe o najwyższej frekwencji	81
4.2.2.2. Frekwencja leksemów przymiotnikowych w korpusie prozy	83
4.2.2.3. Frekwencja leksemów przymiotnikowych w korpusie poezji	86
4.2.3. Zastosowanie Korpusomatu – pozyskanie konkordancji dla form przymiotnikowych	89
4.2.3.1. Etapy analizy	89
4.2.3.2. Charakterystyka wybranych predykatów przymiotnikowych – przykład wyników z analizy konkordancji	97
4.2.4. Zastosowanie TSCC 1.3. do badań nasycenia korpusów prozy i poezji Julii Hartwig formami przymiotnika	105
4.2.4.1. TSCC 1.3. – metoda	105
4.2.4.2. Stopień wysycenia korpusu formami przymiotnikowymi	109

4.2.5. Badanie rozwoju zasobu leksykalnego wykorzystywanego przez Julię Hartwig (przymiotniki)	112
4.2.6. Zastosowanie narzędzi informatycznych do analizy kompetencji językowej	112
4.2.6.1. Pojęcie leksykonu wirtualnego	112
4.2.6.2. Zastosowanie metody przyrostowej do badania kompetencji i leksykalnej pisarza	113
4.3. Problemy i korzyści wynikające z automatycznej ekstrakcji przymiotników wraz z kontekstami	122
4.3.1. Trudności i ograniczenia wynikające z zastosowanej metody	122
4.3.2. Korzyści wynikające z zastosowanej metody	123
ROZDZIAŁ 5	
PODSUMOWANIE	124
5.1. Wnioski	124
5.2. Dalsze prace	127
ANEKS. System kodowania fleksji polskiej POLEX	128
SŁOWNIK TERMINÓW	133
INDEKS	136
BIBLIOGRAFIA	139
ILUSTRACJE	149

## STRESZCZENIE

Głównym celem badawczym, którego realizację inauguruje badania prezentowane w tej rozprawie jest ustalenie, jakie elementy warsztatowe w tekstach literackich Julii Hartwig można uznać za najważniejsze. Przedstawiamy analizę dotyczącą warsztatu pisarskiego Hartwig rozumianego jako sposób posługiwania się przez pisarza językiem, a dokładniej jego zasobem leksykalno-gramatycznym. Podstawę przedstawionych badań stanowi zebrany i opracowany elektroniczny korpus całości publikowalnej prozy i poezji Julii Hartwig z lat 1954-2016. Stosujemy metodologię badań empirycznych, przeprowadzonych z wykorzystaniem narzędzi informatycznych. Pozyskane wyniki ilościowe pozwoliły na sformułowanie wniosków dotyczących stylu Hartwig. Stwierdzamy między innymi tendencję Poetki do posługiwania się konkretnymi przymiotnikami oraz wskazujemy na formy dominujące w jej instrumentarium. Wykazujemy również podobieństwa i różnice w dynamice przyrostu nowych leksemów przymiotnikowych w korpusie prozy i poezji. W rozprawie sygnalizujemy kolejne, rozpoczęte już działania mające za przedmiot szczegółową analizę semantyczno-składniową przymiotników charakterystycznych dla Julii Hartwig.

## ABSTRACT

The main research goal, the implementation of which is inaugurated by the research presented in this dissertation, is to determine which text elements of Julia Hartwig's literary workshop can be considered the most important. We present the research on Julia Hartwig's writing technique understood as a writer's way of using a language, and more specifically its lexical and grammatical resources. The analysis is based on the collected and elaborated electronic corpus of all of Julia Hartwig's published prose and poetry from 1954-2016. We use the methodology of empirical research conducted with the use of IT tools. The obtained results allowed for the formulation of conclusions about the Hartwig's style. Among other things, we find the Poet's tendency to use specific adjectives and indicate the forms that dominate her writing's instruments. Moreover, we show similarities and differences in the dynamics of the growth of new adjective lexemes in the prose and poetry corpus. We plan to extend the conclusions and observations presented in the dissertation with a detailed semantic and syntactic analysis of adjectives characteristic of Julia Hartwig.

## RESUMÉ

L'objectif de cette dissertation est de présenter la recherche qui initie la réalisation d'un projet à long terme visant l'identification des éléments de l'atelier littéraire de Julia Hartwig qui peuvent être considérés comme les plus importants. Nous présentons une analyse de sa technique d'utilisation de la langue et plus précisément une analyse des moyens lexico-grammaticaux qu'elle utilise. Notre recherche prend comme point de départ la collecte et le développement du corpus de prose et de poésie de l'artiste entre 1954 et 2016. Dans notre analyse nous nous servons de la méthodologie empirique empruntée aux sciences naturelles ainsi que des outils informatiques. Les résultats quantitatifs obtenus nous ont permis de concurer sur le style propre de Julia Hartwig. Nous observons, entre autres, l'inclination pour l'emploi privilégié de l'adjectif, tout en notant certaines formes dominantes dans ses choix lexicaux. Nous notons également les similarités et les différences dans l'évolution de l'emploi des formes adjectivales entre la prose et la poésie. Dans la thèse, nous indiquons aussi d'autres activités de recherche, déjà commencées, visant une analyse sémantique et syntaxique détaillée des adjectifs caractéristiques pour le style de Julia Hartwig.

## PRZEDSŁOWIE. DLACZEGO JULIA HARTWIG?

„Wierzę w zdanie *W przystanek, który szuka formy*  
*składnej i skromnej jak codzienna mowa*  
*Wszystko we mnie pragnie tej chwili kiedy kształt*  
*obejmie bezkształtne [...]*”  
Julia Hartwig

Cytowane wersy pochodzą z wiersza *Potrzeba*, który jest chyba jednym z częściej komentowanych w literaturze przedmiotu utworów poetyckich Julii Hartwig dotyczącym fenomenu języka, w którym wskazuje się na potrzebę nazywania oraz na zawarte w języku reguły pozwalające określić stan rzeczy, a także sformułować uczucia i myśli, co według krytyki stanowi warsztatowe credo Hartwig. Zagadnienia formy poetyckiej i językowej odpowiedzialności były jednymi z wielu tematów zajmujących poetkę, ale powiedzieć o Julii Hartwig poetka to za mało.

W pierwszym okresie swojej aktywności literackiej, tuż po wojnie, była przede wszystkim tłumaczką – z wykształcenia romanistką – która dokonała przekładów wybranych utworów Aragona, Rimbauda, Apollinaire’a, Cendrasy, Jacoba, Michaux i Reverdey’ego, będących pokłosiem jej pobytu na stypendium w Paryżu w latach 1947-1950. Była wówczas również eseistką i reporterką opisującą socrealistyczną Polskę w zbiorze zatytułowanym *Z niedalekich podróży* z 1954. Pierwszą książkę poetycką *Pożegnania* wydała w 1956 roku. Później, w latach pięćdziesiątych i sześćdziesiątych, pisywała słuchowiska radiowe i książki dla dzieci, jednocześnie pracując nad dwoma obszernymi monografiami literackimi dotyczącymi Apollinaire’a i Gerarda de Nerval’a. Z biegiem lat jej twórczość rozwijała się i rozrastała intensywnie, aż do ostatniego tomu *Spojrzenie* z 2016 roku. Mamy zatem do czynienia z twórczością, która obejmuje blisko osiemdziesiąt lat pracy ze słowem, jeśli uświadomimy sobie, że pierwsze „dziewczęce” wiersze opublikowane były w szkolnym czasopiśmie „Słońce” w 1938 roku.

Dorobek Julii Hartwig został już dobrze opisany przez krytyków i badaczy literatury, czego dowodem jest bardzo długa i bogata bibliografia rzetelnie zbierająca wszystko, co o Hartwig napisano od jej debiutu i uzupełniana do 2010 roku przez Ośrodek Brama Grodzka – Teatr NN w Lublinie, miasta, z którego pochodziła Autorka

(<https://teatrnn.pl/leksykon/artykuly/julia-hartwig-bibliografia>). Już pobieżny ogląd zamieszczonych tam tytułów, zarówno samych dzieł, jak ich opracowań, dowodzi wszechstronności tematyki i treści, jakiej poświęcona była działalność pisarska Hartwig, zawarta w ponad dwudziestu tomach poezji, monografiach literackich, reportażach z podróży, dziennikach, felietonach i esejach, a także szczególnie przez nią ulubionych prozach poetyckich, oryginalnych krótkich zapiskach w formie *błysków* (właściwy tylko Hartwig gatunek!) czy we wspomianej twórczości dla dzieci (również książki pisane wspólnie z mężem Arturem Międzyrzeckim po narodzinach ich córki). Ponadto była autorką różnych wstępów i komentarzy do książek innych autorów, w tym do albumów z artystycznymi fotografiami Ziemi Lubelskiej Edwarda Hartwig a, jej brata, z którym łączyło ją szczególne zamiłowanie do sztuki.

Rozmaitość gatunków, które podejmowała Hartwig, wymagała z założenia umiejętności operowania słowem w zależności od celu formułowanego przekazu, wybranej formy czy oczekiwanego odbiorcy. Ponadto podkreślić należy spory wkład w rozwój jej warsztatu pisarskiego wieloletniej pracy translatorskiej, nie tylko związanej ze wspomnianą już literaturą francuską, ale również przekłady z poezji amerykańskiej. Ceniła bardzo, między innymi, Allena Ginsberga i Williama Carola Williama, Walta Whitmana oraz Emily Dickinson. Pobyt w Stanach Zjednoczonych w latach 1970-1974 wywarł wpływ na poetykę i tematykę jej utworów zbliżoną w tamtym okresie raczej do obserwacji reporterskich niż lirycznych doznań.

Hartwig była poetką umiaru, która operowała słowem w sposób odważny, ale wyważony, zawsze przemyślany. Bogactwo leksyki wynikało u niej z zanurzenia w kulturze europejskiej i amerykańskiej, dalekie jednak od aspektów popularnych i rejestrów potocznych. Mimo to preferowała język prosty i jasny niezależnie od rodzaju dyskursu jaki prowadziła. Wskazywała również na ironię i żart, które są nieodzownymi elementami działalności pisarskiej i pozwalają zdystansować się do opisywanych sytuacji czy relacji. Ceniła również cytat. Ład i harmonia formy, pozorna prostota jej różnogatunkowych wypowiedzi pochybiana jest powszechnie za atut jej utworów, a nie dowód na wyeksploatowanie warsztatu. Stosowała się konsekwentnie do podstawowej zasady pisarskiej, że literatura powinna mieć związek z rzeczywistością, która wynika ze sztuki uważnej obserwacji.

Ogromną wartość stanowi również spuścizna diarystyczna i epistolograficzna: *Wspólna obecność. Korespondencja Julii Hartwig i Artura Międzyrzeckiego z Anną i Jerzym Turowiczami z 2012 roku, Julia Hartwig, Artur Międzyrzecki, Czesław Miłosz.*

*Korespondencja z 2012 oraz Julia Hartwig, Artur Międzyrzecki – Katarzyna i Zbigniew Herbertowie z 2013 roku.* To między innymi wśród listów przygotowanych przez poetkę do druku odnajdujemy doświadczenia polskiej inteligencji powojennej odwołujące się do takich momentów historycznych jak: działalność konspiracyjna z czasów okupacji, wydarzenia marcowe i podpisanie listu '68, po którym została objęta zakazem druku, również moment podpisania w styczniu 1976 roku „Memoriału 101” wymierzonego przeciwko zmianom w konstytucji. Listy oraz dzienniki dają świadectwo stałemu wsparciu dla kultury niezależnej z lat 1976-1989 oraz zaangażowaniu się w ruch niepodległościowy związany z Komitetem Obywatelskim przy Przewodniczącym NZSS „Solidarność”. Twórczości poetki jest wyjątkowa nie tylko ze względu na długi czas jej trwania, ale również dlatego, że dekady, które obejmuje, to często ważne przełomowe momenty w historii naszego kraju i świata, kiedy następują zmiany stosunków społecznych, których poetka była świadkiem, i którym dała wyraz. Szczególnie w dziennikach pojawia się motyw zapamiętywania i utrwalania dziejowych momentów oraz prezentowania sylwetek osób ważnych dla naszej kultury, czasem niezłomnych, czasem zagubionych. Wielu współczesnych wybitnych pisarzy wpisywało się w ten nurt, ale osiągnięcia Hartwig w sztuce ocalania od zapomnienia wydają się niedoścignione.

O uznaniu jej talentu i umiejętności świadczą przyznane liczne wyróżnienia: Nagroda Foundation d'Hautvillers „Prix de Traduction” (Francja 1978), Nagroda Literacka im. Alfreda Jurzykowskiego (USA, 1981), Nagroda Polskiego PEN Clubu (1979,1997), Nagroda Thornton Wilder Prize (USA, 1986), Nagrody im. Georga Trakla (Austria, 1991), Nagroda Ministra Kultury za całokształt twórczości (2001), Nagroda Wielkiej Fundacji Kultury(2002), Nagroda Polskiego PEN Clubu im. Jana Parandowskiego (2009) i Nagroda im. Czesława Miłosza (2009). Ponadto we Włoszech otrzymała Premio Acerbi za całokształt twórczości (2011), w kraju nagrodę dzieła życia „Dzieło życia” w XII edycji Nagrody im. C.K. Norwida (2013) oraz Nagrodę Poetycką im. Wisławy Szymborskiej (2014), a także Doktorat Honoris Causa nadany przez Uniwersytet im. Adama Mickiewicza w Poznaniu (2015). Doczekała się również honorów państwowych, takich jak: Krzyż Wielki Orderu Odrodzenia Polski (2011), Krzyż Oficerski Odrodzenia Polski (1997), Medal „Zasłużony Kulturze Gloria Artis” (2005). Jednak największym zaszczytem było przyznanie Jej przez Rząd francuski w 2007 roku Narodowego Orderu Legii Honorowej V klasy.

W rozprawie zajmuje nas warsztat pisarski Julii Hartwig wyjątkowej, znanej i cenionej Autorki, należącej do generacji wojennej, która wniosła ogromny wkład

w kształtowanie się tożsamości i kultury naszego kraju. Przyjaźniła się z Czesławem Miłoszem, Zbigniewem Herbertem, Antonim Słonimskim i Jerzym Turowiczem, ceniła Wisławę Szymborską, miała słabość do liryki Konstantego Ildefonsa Gałczyńskiego, lubiła Jarosława Iwaszkiewicza. Podczas pobytu na stypendium w Paryżu w latach 1947-50 poznała Pabla Picassa i Louisa Aragona, w swym warszawskim mieszkaniu gościła Jeana Paula Sartre'a, Simone de Bauvoire i wielu innych. Z drugiej strony do jej kręgu należeli Marek Edelman, Leszek Kołakowski, Tadeusz Mazowiecki, Bronisław Geremek czy Adam Michnik. Wymienionych wybitnych artystów, ludzi kultury i społeczników znała osobiście i w swoim długim życiu uważnie czerpała z tych relacji wiedzę i inspirację. Uczestnictwo w bogatym życiu literackim było tym bardziej intensywne, że jej mąż Artur Międzyrzecki, również poeta, tłumacz i eseista – absolwent École des Sciences Politique w Paryżu, uczestnik bitwy pod Monte Cassino – był organizatorem życia literackiego w powojennej Polsce. Pracował w wielu redakcjach, Związku Literatów Polskich oraz polskim PEN Clubie (w latach 1991-1996 był jego prezesem i wiceprezesem światowego PEN). Z łatwością możemy zauważyć, że twórczą osobowość Julii Hartwig, niezwykle złożoną, tworzyły rozmaite miejsca i środowiska.

Julia Hartwig była ważną postacią w historii literatury polskiej nie tylko dlatego, że była uznaną poetką, ale również przez swój wkład w dbałość o kulturę języka polskiego i języka poetyckiego. Cechowała ją ogromna erudycja oraz, jak zostało powiedziane, znajomość kultury, języka francuskiego i angielskiego, dlatego zasadna i interesująca wydaje się refleksja nad warsztatem pisarskim tej autorki. Hartwig w referacie, wygłoszonym podczas nadania jej honorowego doktoratu na Uniwersytecie im. Adama Mickiewicza w Poznaniu w 2015 roku nawiązywała do Guillaume'a Apollinaire'a i Gérarda de Nerval'a by skonstatować, „że początkiem każdej prawdziwej sztuki jest wyrazista osobowość, że jej kształt wpływa w sposób najbardziej wiarygodny na kształt dzieła” (Hartwig 2015). To, że Julia Hartwig jest *wyrazistą osobowością* nie budzi wątpliwości, a *kształt jej dzieła* intryguje i jest źródłem językoznawczych dociekań.

## WSTĘP. PROBLEM I METODY BADAWCZE

Głównym dalekosiężnym celem badawczym, którego realizację inaugurują badania prezentowane w tej rozprawie jest ustalenie, jakie elementy w wypowiedziach literackich Julii Hartwig można uznać za najważniejsze. Podzielając opinię niektórych badaczy<sup>1</sup>, twierdzimy, że do tego typu elementów należy zaliczyć przede wszystkim przymiotnik - jedną z podstawowych kategorii semantyczno-logicznych w zdaniu, zgodnie z podejściem *predykatowo-argumentowym* w gramatyce<sup>2</sup>. Analizując warsztat pisarski Julii Hartwig, pragniemy zatem zwrócić uwagę na dominującą w jej twórczości rolę tej kategorii gramatycznej, przy użyciu której ta wybitna Autorka opowiada o świecie. Przekonanie o słuszności powyższej hipotezy wynika z przemyśleń samej poetki, która twierdzi, że przymiotnik stanowi dla niej najważniejszą część mowy i jest wykładnikiem „literackości” dzieła, a zdanie buduje wiersz. Dla potwierdzenia tej uwagi warto przywołać w tym miejscu ciekawy komentarz Hartwig, dotyczący twórczości amerykańskiego poety W.S. Merwina, a mianowicie: „Ostatnie wiersze Merwina cechuje pewien ascetyzm, który wyraża się w pozbywaniu się literackości, za czym idzie oszczędność w używaniu przymiotników” (Hartwig 2014: 323). Dlatego wydaje się uprawnione stwierdzenie, że w przymiotnikach „zaszyta” jest część Hartwigowego obrazu świata. Opisane w pracy metody badawcze, tzn. zarówno sposób pozyskania korpusu, jak i zastosowane narzędzia analizy z dziedziny NLP, o których będzie mowa później, stanowią jedynie tło do wykazania wyróżnionej roli tej kategorii gramatycznej w zdaniach budowanych przez Julię Hartwig.

Analiza warsztatu pisarskiego Julii Hartwig odbywa się z zastosowaniem metod z dziedziny lingwistyki komputerowej, co w praktyce oznacza odwołanie się do teorii budowania korpusów elektronicznych w formatach przystosowanych do ich przeszukiwania i dalszego przetwarzania. W niniejszej pracy przyjmujemy założenia lingwistyczne wykorzystane w opisie danych językowych w postaci spójnego formalizmu, będącego podstawą słownika POLEX (Vetulani Z., Walczak, Obrębski, Vetulani G., 1998) oraz stosujemy techniki automatycznego analizowania tekstu.

---

<sup>1</sup> Legeżyńska (2017:135), komentując reportaże *Z niedalekich podróży*, stwierdza: „W tego rodzaju deskrypcjach, nietkniętych wirusem socu, Julia Hartwig jest sobą: poetką obrazu materialnego świata, celnie utrwalonego w literackim opisie za pomocą przymiotnika (o którym powie kiedyś, że jest dla niej najważniejszą częścią mowy) lokowanego w zdaniu o harmonijnej strukturze”.

<sup>2</sup> Zob. Rozdział 2 niniejszej pracy.

Warunkiem koniecznym do właściwej analizy warsztatu pisarskiego Julii Hartwig było przygotowanie na potrzeby tej pracy znormalizowanego elektronicznego korpusu jej twórczości, na który składają się utwory poetyckie (korpus poezji) i utwory pisane prozą, reportaże, dzienniki, eseje i felietony – włączone do korpusu prozy. Niezależnie od głównego celu pracy opracowanie korpusu w formie cyfrowej jest wartością samą w sobie. Pozostała część opisanego w Przedśłowiu bogatego dorobku wymagałaby osobnego omówienia, co mogłoby nastąpić w ewentualnej kolejnej fazie prac nad rozszerzeniem dotychczas zebranego materiału.

Rozprawa wpisuje się w nurt humanistyki cyfrowej i jest próbą pokazania pewnego modelu pracy z tekstem literackim w celu dopełnienia ugruntowanej wiedzy na temat twórczości Hartwig, a także rodzajem raportu z aplikowania narzędzi do badań lingwistyczno-literackich, które w praktyce naukowej zdają się być coraz mniejszym wyzwaniem dla filologa ze względu na przystępność stosowanych w badaniach interfejsów oraz możliwości współpracy z informatykami.

Użycie narzędzi do analizy tekstu gwarantuje obiektywność otrzymanych wyników. Ważne były dla nas, z jednej strony, funkcjonalności zastosowanych aplikacji w zakresie przeszukiwania bazy danych, jaką stanowi korpus prozy i poezji Hartwig, które dotyczyły możliwości generowania list referencyjnych przymiotników, list frekwencyjnych, oglądu wyników ze względu na metadane, a z drugiej, szybkiego pozyskania kontekstów wystąpienia interesującej nas kategorii gramatycznej. Narzędzia informatyczne okazały się zatem kluczowe dla pozyskania wyników początkowych, dotyczących przymiotników u Hartwig, takich jak wielkość korpusu, listy referencyjne i frekwencyjne, zestawienia użyć, które w kolejnych krokach podlegały analizom jakościowym omawianej kategorii. Możliwość szybkiego zliczania różnych elementów analizowanego tekstu okazała się kolejnym atutem przyjętej metody badawczej. Sprawą nadrzędną było zyskanie możliwości szybkiego sprawdzania i analizowania korpusu, co nie byłoby możliwe, gdyby praca przebiegała na tekstach w tradycyjnej wersji papierowej. Przeprowadzona próba zastosowania podejścia inżynierskiego do korpusu twórczości Hartwig stanowi – jak nam się wydaje - wartościowy eksperyment badawczy.

Jednym z celów szczegółowych, które niniejsza rozprawa antycypuje, jest rozpoznanie relacji pomiędzy przymiotnikiem a wchodzącym z nim w związek wyrazowy czasownikiem. Charakterystyka semantyczno-syntagmatyczna przymiotnika z czasownikiem w utworach Hartwig jest świadectwem dokonywanych przez Julię Hartwig wyborów leksykalnych i zarazem wyznacznikiem jej warsztatu pisarskiego.

W rozdziale pierwszym pt. *Twórczość Julii Hartwig na tle badań nad językiem i stylem pisarzy* omawiamy badania nad językiem pisarzy zarówno w perspektywie tradycyjnych metod badawczych, jak również metod komputerowych i ilościowych. W tym kontekście przywołujemy prace badawcze dotyczące twórczości Julii Hartwig i zwracamy szczególną uwagę na zagadnienia dotyczące językowych aspektów jej utworów.

W rozdziale kolejnym (*Przymiotnik wyznacznikiem warsztatu Julii Hartwig*) przedstawiamy opis języka naturalnego metodami formalnymi i opisujemy różne realizacje predykatu w zdaniu osadzone w koncepcji *struktury predykatywno-argumentowej* i przyjmujemy założenie, że do najważniejszych elementów logiczno-semantycznych w zdaniu stosowanych w twórczości Hartwig należy zaliczyć przymiotnik. Szczegółowo analizujemy formy predykcji przymiotnikowej i odwołujemy się do przykładów z dzieł Julii Hartwig, w ten sposób interpretujemy obraz świata odzwierciedlony w używanych przez Poetkę przymiotnikach.

Rozdział trzeci (*Korpus prozy i korpus poezji Julii Hartwig*) poświęcony jest korpusom tekstów pod kątem lingwistyki korpusowej. Omawiamy zawartość oraz metody i zasady tworzenia korpusu prozy i korpusu poezji wraz z rozwiązaniami problemów dotyczących pozyskiwania korpusu w formie elektronicznej.

Czwarty z kolei rozdział (*Wykorzystanie metod automatycznego przetwarzania tekstu do analizy przymiotnikowej w korpusie prozy i poezji Julii Hartwig*) poświęcony jest zastosowaniu wybranych narzędzi przetwarzania tekstów języka naturalnego do analizy korpusu Julii Hartwig. Przedstawiamy badania oparte na *listach referencyjnych* form i leksemów przymiotnikowych oznaczonych w korpusie według wspólnego paradygmatu POLEX. Prezentujemy również analizę *list frekwencyjnych* wraz z przykładami użyc zestawionych na listach *konkordancji*, badanie rozwoju zasobu leksykalnego wykorzystywanego przez Poetkę oraz badanie kompetencji leksykalnej pisarza z wykorzystaniem metody przyrostowej do oszacowania słownictwa wirtualnego Julii Hartwig dla przymiotników.

Końcowy rozdział piąty (*Podsumowanie*) zawiera omówienie wniosków oraz zaanonsowanie dalszych prac będących planowaną kontynuacją przedstawionych w niniejszej rozprawie badań.

Rozprawę zamykają aneks, indeks i bibliografia.

# ROZDZIAŁ 1

## Twórczość Julii Hartwig na tle metodologii badań nad językiem i stylem pisarzy

Przedmiotem naszego zainteresowania jest warsztat pisarski autora rozumiany jako sposób posługiwania się przez pisarza/poetę językiem, a dokładniej jego zasobem leksykalno-gramatycznym. Tak zarysowana problematyka wpisuje się w długą historię naukowych omówień i zobowiązuje do przedstawienia po krótko stanu wiedzy na temat języka pisarzy oraz metod stosowanych przez badaczy.

### 1.1. Badania nad językiem pisarzy

Badania nad językiem pisarza mają bogatą tradycję w polskim językoznawstwie. W pracach podejmujących problem języka twórcy, czyli *języka osobniczego* (zwanego również *językiem indywidulowanym* albo *idiolektem*), odnajdujemy co najmniej dwa sposoby rozumienia tego terminu. Pierwszy z nich ma znaczenie wąskie i wskazuje na język twórcy jako materiał ilustrujący procesy rozwojowe języka. Tekst literacki stanowi wówczas rodzaj dokumentu, świadectwa danego czasu (Doroszewski 1949, Brzeziński 1975, Wilkoń 1978, Walczak 1988, Lubaś 1992). Drugie ujęcie, szersze, traktuje język osobniczy jako język danej jednostki wyróżniający się na tle innych. Inaczej mówiąc, język twórcy to odmiana, czyli „modyfikacja systemu języka ogólnego, system w użyciu” (Puzynina 1988, Kozłowska 2009). W rozprawie przyjmujemy tę drugą perspektywę, gdyż przedmiotem naszych zainteresowań jest autor (twórca) widziany poprzez swój język, a precyzyjnej Julia Hartwig postrzegana przez pryzmat dysponowania słownictwem.

W kontekście języka danego autora należy wziąć pod uwagę to, co w badaniach określa się jako *styl indywidualny* lub *styl pisarza* czy też *idiostyl*, których granice są płynne a znaczenia niesprecyzowane i dlatego są one wciąż przedmiotem ożywionej dyskusji (np. Klemensiewicz 1961, Wilkoń 1978, Fliciński 2004, Gajda 2013). Badacze wskazują, że „klasyczna” definicja stylu ugruntowana została przez szkołę francuską (Ch. Bally) i stanowi, iż styl to zasada wyboru i organizacji środków językowych w tekście. Inna definicja – sformułowana przez strukturalistów – odzwierciedlała

myślenie o stylu jako strukturze pozbawionej piętna jej twórcy. Badania strukturalistyczne polegały na wykazywaniu sformalizowanego katalogu cech tekstu i określaniu ich funkcji (praska szkoła stylistyki funkcjonalnej) niezależnie od realnych kontekstów kulturowych, jak również autora oraz sytuacji komunikacyjnej. Współczesne teorie humanistyczne na powrót wiążą pojęcie stylu z autorem (podmiotem) i jego sposobem widzenia świata. Styl jest „rezultatem indywidualnych działań twórcy i wyraża określoną ideę” (Sławkowa 2016).

Główne zagadnienia badawcze dotyczące słownictwa pisarzy referowane w artykułach i monografiach skupiają się w co najmniej czterech perspektywach. Pierwsza, historyczno-językowa, rejestruje stan języka w danym momencie, druga, stylistyczna, opisuje sposób wykorzystywania w dziele zasobów leksykalnych danego języka. Kolejna, semantyczno-syntagmatyczna, porządkuje system językowy ze względu na znaczenia słów i ich konteksty. Kierunek semantyczno-syntagmatyczny korzysta z metod wypracowanych w badaniach pól wyrazowych (Tokarski 2001) czy też z teorii konotacji leksykalnej (Pajdzińska, Tokarski 2001). Czwarta perspektywa, statystyczna, zajmuje się analizą ilościową słownictwa autora w oparciu o metody matematyczne, które z powodzeniem stosowano już pod koniec XIX wieku (Lutosławski 1896).

„Historyczne” metody statystyczne w badaniach nad słownictwem pisarzy polskich były wykorzystywane przez polskich badaczy w pracach dotyczących słownictwa Tadeusza Gajcego (Wyka 1962), oraz nad słownictwem w eposie *Pan Tadeusz* (Sambor 1969), a także w *Dziadach* Adama Mickiewicza (Rudnicka-Fira 1986). Badano także słownictwo polskich naturalistów (Stachurski 1989), tworzone listy frekwencyjne słownictwa współczesnej poezji (Zgólkowie 1992) itd. Nowe możliwości, jakie współcześnie dają narzędzia informatyczne, wykorzystywane w analizach słownictwa, prezentują krakowscy badacze Jan Rybicki (2014, 2017) i Maciej Eder (2014), którzy opracowali stylometryczną mapę literatury polskiej, używając do tego celu programu do analizy stylometrycznej *Stylo*, który został włączony do pakietu Websty (dostępnego na platformie Clarin-pl). Zenon Klemensiewicz (1961), w klasycznym już artykule *Jak charakteryzować język osobniczy?* stwierdza, że najtrudniejszym i niewykonalnym (ówcześnie) zadaniem jest zestawienie słownictwa autora i porównanie go ze słownictwem ogólnym języka polskiego po to, by wykazać ilościowe i jakościowe cechy języka osobniczego. Blisko sześćdziesiąt lat później, jak pokazali Rybicki i Eder, zestawianie słownictwa autorów jest możliwe, zatem z wykorzystaniem nowych technologii, zadanie opisane przez Klemensiewicza jest wykonalne.

Powyższe zagadnienia wpisują się w szeroko pojęty nurt humanistyki cyfrowej<sup>3</sup>, będącej w kręgu zainteresowań lingwistów, filologów, socjologów, historyków sztuki itd. Tomy czasopism *Teksty Drugie* (2014)<sup>4</sup> czy *Forum Poetyki* (2017)<sup>5</sup> poświęcone humanistyce cyfrowej oraz ukazanie się na rynku polskim w 2016 roku książki Franco Morettiego *Wykresy, mapy, drzewa* są tego jasnymi dowodami. Historia literatury widziana bardziej racjonalnie kusi, z jednej strony, perspektywą makroanalityczną (Jockers 2013), a z drugiej, możliwością tzw. *czytania z daleka* (Moretti 2016). Współczesne, informatyczne techniki badawcze, umożliwiają analizę wielkich zasobów literackich, których samodzielne czytanie zajęłoby wiele lat albo byłoby niemożliwe.

Przydatność metod z dziedziny lingwistyki komputerowej do badań językoznawczych i literaturoznawczych, choć budzi olbrzymi entuzjazm wśród naukowców, wymaga od obydwu stron poszerzenia kompetencji własnych w celu wzajemnego porozumienia i wypracowania wspólnych lub równoległych pytań i założeń badawczych (zob. Centrum Humanistyki Cyfrowej przy Instytucie Badań Literackich PAN). Dlatego też na styku tychże zainteresowań powstały konsorcja służące rozwojowi infrastruktury badawczej dla humanistyki, z których wymienić należy, tytułem przykładu, polskie projekty, takie jak *Dariah.pl* czy *Clarín.pl* (zob. Maryl 2017).

## 1.2. Metody komputerowe w badaniach nad językiem pisarzy

W niniejszej rozprawie pragniemy przedstawić zastosowanie wybranych metod komputerowych do reprezentatywnej części dzieła Julii Hartwig. W pierwszej kolejności zwrócimy jednak uwagę na wybrane prace nad językiem polskim prowadzone metodami lingwistyki komputerowej. W trzech kolejnych podrozdziałach skoncentrujemy się na metodach, które zostały zastosowane przy badaniu korpusu Julii Hartwig

---

<sup>3</sup> Maciej Maryl w następująco wypowiada się na temat praktyki humanistyki cyfrowej: „Kluczowe jest tu uznanie humanistyki cyfrowej nie tyle za konkretną dyscyplinę czy metodę, ile zespół różnych praktyk badawczych, wykorzystujących cyfrową (pierwotnie bądź wtórnie) formę przedmiotu badań. A zatem chodzi tu nie tylko o badanie kultury cyfrowej, ale także o stosowanie metod cyfrowych do badania lub udostępniania dzieł kultury analogowej.” (Maryl 2017:289) Maryl, odnosząc się do dorobku światowego w tej dziedzinie, proponuje następującą periodyzację praktyk cyfrowych: „pierwsza fala humanistyki cyfrowej dotyczy digitalizacji i remediacji tradycyjnego warsztatu badawczego dyscyplin (zaliczymy do niej skanowanie tekstów i obiektów, budowę cyfrowych archiwów i katalogów, tworzenie korpusów, adnotowanie tekstów itd.); fali drugiej towarzyszy rozpoznanie i wykorzystywanie możliwości dalszego przetwarzania materiału cyfrowego pozyskanego w fazie pierwszej, czyli wykorzystanie nowego medium do tworzenia nowych metod badawczych i gatunków komunikacji naukowej (narzędzia do badania i wizualizacji, badania na dużych zbiorach materiałów), oraz fala trzecia, czyli krytyczna analiza ograniczeń epistemologicznych medium i samych założeń humanistyki cyfrowej”.(ibidem, 291)

<sup>4</sup> *Teksty Drugie*, 2014, nr 2

<sup>5</sup> *Forum poetyki – Makropoetyka*, (2017), <http://fp.amu.edu.pl/forum-poetyki-jesien-2017/>, [05.04.2018]

przygotowanego na potrzeby tej pracy bądź są przewidziane do wykorzystania w dalszych pracach.

### 1.2.1. Dane językowe – korpusy tekstów

Na wagę danych językowych jako podstawę przyszłego *przemysłu językowego* (ang. *language industry*) zwrócił uwagę Antonio Zampolli, założyciel i pierwszy prezes stowarzyszenia ELRA (European Language Resources Association), które postawiło sobie za cel gromadzenie i dystrybucję zasobów językowych (ang. *language resources*) dotyczących języka mówionego (ang. *speech*), tekstów, słownictwa i terminologii jako podstawy umożliwiającej rozwój lingwistyki wg metodologii nauk empirycznych oraz – co za tym idzie – stworzenie technologii, a w dalszej perspektywie tzw. inżynierii języka na skalę przemysłową. Tym samym Zampolli przyczynił się do zbliżenia nauk humanistycznych i lingwistyki komputerowej – o czym piszą Bogdan Hojdis i Adam Cankudis w artykule poświęconym edycjom cyfrowym literackim i naukowym, traktowanym jako podstawa humanistyki cyfrowej (Hojdis, Cankudis 2020: 15)<sup>6</sup>. Zgromadzenie zbiorów danych językowych, w których istotną rolę odgrywają korpusy literackie, umożliwiło wykorzystanie potencjału obliczeniowego informatyki do badania dorobku pisarskiego twórców metodami lingwistyki komputerowej.

O ile odpowiednio zebrane korpusy danych językowych pozwalają na tworzenie podstaw empirycznych dla wielu obszarów badań lingwistycznych, w szczególności przez uzyskiwanie modeli języka umożliwiających przetwarzanie informatyczne, o tyle pozyskiwanie i opracowywanie właściwych modeli obliczeniowych na potrzeby badań nad twórczością literacką stwarza wiele problemów już na etapie gromadzenia odpowiednich korpusów danych. Dotyczy to w szczególności dzieł twórców współczesnych, gdzie na przeszkodzie stoi, poza problemami technicznymi związanymi z dostępem do źródeł i ich jakością, szereg problemów natury prawnej. Trudnym

---

<sup>6</sup> W jednym z przypisów do wspomnianego artykułu czytamy: „Należy mieć na uwadze, że lingwistyka komputerowa nie zawsze jest utożsamiana z humanistyką cyfrową. Podział ten będzie się uwidaczniał dopóty, dopóki humaniści cyfrowi nie uświadomią sobie korzyści, jakie dają im prace lingwistów i możliwości zautomatyzowanego przetwarzania tekstu. Próby pogodzenia obu środowisk podejmowano w przeszłości: „Językoznawstwo komputerowe zawsze rozwijało się niezależnie od informatyki humanistycznej i mimo wysiłków Dona Walkera w TEI Steering Committee nadal była ona odrębną dyscypliną. Walker i Antonio Zampolli z Instytutu Lingwistyki Komputerowej w Pizie bardzo starali się połączyć nauki humanistyczne z lingwistyką komputerową, ale z niewielkim skutkiem” (Hojdis, Cankudis 2020:15)

wyzwaniem w tym względzie okazało się zestawienie reprezentatywnego korpusu dzieł Julii Hartwig<sup>7</sup>.

### 1.2.2. Wybrane metody inżynierii języka

Obiektem zainteresowań lingwistyki komputerowej jest język naturalny i jego przetwarzanie (ang. *Natural Language Processing /NLP/*). W ramach inżynierii językowej, wspólne wysiłki lingwistów oraz informatyków doprowadziły do powstania różnego typu narzędzi, spośród których niektóre okazują się przydatne przy badaniu twórczości wybitnych pisarzy (por. m. in. Eder 2014, 2016, Rybicki 2014, 2017). W niniejszej pracy podążamy tym tropem w analizie warsztatu pisarskiego Julii Hartwig.

*Przetwarzanie języka naturalnego* jest możliwe, gdy traktujemy język jako zorganizowaną strukturę. Sposób postrzegania języka jako struktury nawiązuje do tradycji eksplikacji tekstu rosyjskiej szkoły strukturalnej (Jacobson) oraz gramatyki generatywnej Noama Chomsky'ego, traktującej składnię jako główny czynnik organizujący wypowiedź (zob. np. Vetulani Z. 2004). Kognitywna teoria języka w znacznym stopniu wskazuje na twórczy charakter języka i zajmuje się, między innymi, badaniem znaczenia słów (Bartmiński 1990, Pajdzińska, Tokarski 2001, Taylor 2007).

Prace w zakresie analizy znaczenia słów i relacji zachodzących między nimi zaowocowały stworzeniem semantyczno-leksykalnych baz danych typu wordnet, których pierwowzorem był system Princeton WordNet George A. Millera /PWN/ (Miller i inni, 1990). Pośród tego typu systemów są: Leksykalna Baza Danych PolNet (Vetulani Z., Kubis, Obrębski 2010) opracowana dla języka polskiego na Uniwersytecie im. Adama Mickiewicza w Poznaniu oraz system Słowosieć (Maziarz, Piasecki, Rudnicka 2014) opracowany na Politechnice Wrocławskiej. W obu przypadkach, na wzór systemu Princeton WordNet, struktura logiczna systemów odwołuje się do pojęcia synonimii i relacji hipo- i hiperonimii. (Wczesne prace w projekcie WordNet inspirowane były ideą rekonstrukcji „słownika umysłowego” człowieka i takiego sformalizowania siatki pojęć, by była ona przetwarzalna przez narzędzia informatyczne (Kurcz 2005)). W praktyce ta pierwotna inspiracja często ustępuje miejsca potrzebom konkretnych zastosowań tam, gdzie zachowanie konceptualizacji właściwej użytkownikowi danego języka odgrywa

---

<sup>7</sup> Sytuacja jest odmienna w przypadku dzieł dawnych. Sprzyja temu również polityka wolnego dostępu do dorobku literackiego twórców, co do którego wygasły już prawa autorskie. Dzięki platformom takim, jak wolnelektury.pl badacze mają możliwość pozyskiwania wysokiej jakości danych językowych do różnego typu analiz dzieł klasyków literatury.

rolę drugorzędną. Ma to często miejsce tam, gdzie systemy sztucznej inteligencji wykorzystujące bazy typu wordnet nie wymagają precyzyjnego (głębokiego) rozumienia tekstu. Wykorzystanie informacji semantyczno-leksykalnej gromadzonej w bazach typu wordnet, o których była mowa w Rozdziale 1.2.2., będzie kolejnym krokiem badania aspektów warsztatowych Julii Hartwig, dla którego prace przedstawiane w niniejszej rozprawie stanowią punkt wyjścia.

### **1.2.3. Metody ilościowe**

Badania nad językiem naturalnym odbywają się również w oparciu o analizy ilościowe. Typowym przykładem zastosowań metod ilościowych (a w szczególności statystycznych) są zagadnienia związane z ustalaniem autorstwa. Z kolei zastosowanie rachunku prawdopodobieństwa dla przewidywania następstwa konkretnych liter, sylab, wyrazów czy fraz, pozwala na wykrywanie praw językowych, dotyczących określonych, stałych zależności w systemie języka, takich jak np. prawo Zipfa (Zipf 1943, Guiraud 1954). Prace Zipfa i innych zainspirowały badaczy z kręgu lingwistyki kwantytatywnej do badań ilościowych sekwencyjnej struktury tekstu oraz wpłynęły na uporządkowanie pojęć i metodologii tej dyscypliny (Pawłowski 2010).

### **1.3. Uwagi o języku Julii Hartwig – rozpoznanie**

Językową charakterystykę dorobku Hartwig należy rozpocząć od przywołania najważniejszych opracowań tej twórczości w postaci trzech obszernych monografii: pierwszej, Marcina Telickiego (2009), dotyczącej poetyckiej antropologii, drugiej, autorstwa Flakowicz-Szczyrby (2014), sytuującej poezję Hartwig wobec egzystencji i sztuki oraz dzieła Elżbiety Dutki (2016), prezentującego twórczość poetki w kontekstach autobiograficznych i geograficznych, a także pracy zbiorowej *Pochwała istnienia. Studia o twórczości Julii Hartwig* (2015). Ponadto francuskie inspiracje, zainteresowanie poematem prozą omówili Telicki (2005), Kosowska (2014), Kuczyńska-Koschany (2015). Na szczególny charakter dzieł Hartwig i jej wyrazisty ton, który nie poddaje się żadnym modom i programom, wskazywali Kwiatkowski (1995), Łuszczkiewicz (1998), Hull (2010), Nasiłowska (2011). Stankowska (2015) z kolei pisała, że poezja Hartwig jest klasyczna, ponieważ dąży do odnajdywania ładu i rytmu istnienia, ale że jej poetyka wymyka się pisanim regułom, jak np. nowy gatunek wypowiedzi lirycznej w formie *błysków* (Szaruga 2002, Baran 2018). Podkreślano też

dostojność języka i frazy (Miłosz 2000, Kielar 2000), które zanurzone w sferze kultury okazały się najsprawniejszym i najbliższym systemem znaków (Barańczak 2000, Telicki 2005, Kosowska 2014). Znaczący dorobek Julii Hartwig zauważyli ponadto zmiany w poetyce jej utworów po pobycie w Stanach Zjednoczonych i pisali o tzw. „idiomie amerykańskim” (Legeżyńska 2015, Winiecka 2015: 113, Dutka 2016).

Dla zilustrowania uwag odnośnie języka Julii Hartwig pochodzących od w/w autorów i jednocześnie uwiarygodnienia naszych intuicji związanych z kategorią przymiotnika, przytaczamy dwa fragmenty obrazujące wysoką frekwencję użycia tej kategorii gramatycznej w prozie i poezji:

**Długi jasny** korytarz z podłogą wyłożoną **niemalowanymi** deskami. Po lewej rząd okien, po prawej szereg drzwi. Cicho. **Żadnych** zapachów kuchennych, **żadnych** gospodarskich odgłosów. A przecież akustyka musi tu być świetna. Przez  **cienkie drewniane** drzwi słyhać zapewne **każdy** krok, **każde** skrzypnięcie. Mimo to, zastukawszy, dość długo czekam. Otwiera mi siwy pan o **bladych niebieskich** oczach. Krok ma **nierówny, lewą, sztywną** nogę ciągnie po podłodze. Uśmiecha się **nieśmiałym** uśmiechem człowieka o **gołębim** sercu, **niepewne, badawcze** spojrzenie, śledzące wargi rozmówcy, zdradza wadę słuchu.

[Wprowadzenie do monografii *Apollinaire* (Hartwig1962)]

Z korpusu poezji Julii Hartwig przytaczamy wiersz z tomu *Zapisane* z roku 2013:

TO CO SIĘ DZIEJE już wcześniej się działo  
w myślach i w **ciemnych** znakach **nocnych**  
Nasza przeszłość jest opowieścią której nikt nie odczyta  
Poddana jest jednak osądowi naszego sumienia  
które zawsze mówi: **winny**  
i jest po stokroć **sprawiedliwym** sędzią  
Ale wyławiamy chwile **sześcieliwe**  
Wszak należy nam się jakaś nagroda  
za zbrodnie jakich dokonuje na nas czas

## ROZDZIAŁ 2

### Przymiotnik wyznacznikiem warsztatu Julii Hartwig

Jak zostało już powiedziane w Przedśłowiu, zdanie stanowiło dla Hartwig podstawową strukturę (Przybylski 2015), która porządkowała jej wypowiedź. Autorka przyznawała:

Z czytelnością wierszy łączy się wybór składni. Co do mnie, buduję wiersz na zdaniu, niemal nigdy na wersie, który nazwałabym rozsypanym, typowym na przykład dla futuryzmu.

(Hartwig: 2015: 33).

W niniejszej pracy interesuje nas charakter relacji między elementami języka w obrębie zdania, których wstępnego rozpoznania dokonał Skibski, pisząc o „synatagmatycznej wartości wersów” (2015: 136, 2017: 223). Omówił on zaobserwowane zjawiska składniowe, takie jak: *szereg* (konstrukcję złożoną głównie z rzeczowników, jak np. *owad, ptak, motyl*, nieoddzielonych w tekście Hartwig przecinkami, który przechodzi w apozycję, czyli wyrażenie złożone z rzeczowników o tożsamy formach przypadkowych, czy parataksę. Według Skibskiego tego typu zabiegi składniowe wpływają na „niespieszność” stylu Autorki. Komentując tajniki warsztatu pisarskiego Hartwig, Skibski wskazał na dynamikę funkcji semantyczno-składniowych i wyraźnie podkreślił (jednocześnie zapraszając do badań na ten temat), że to właśnie zależności gramatyczne wyznaczają porządek czytania (o czym najpełniej przekonamy się w rozdziale piątym niniejszej pracy).

Na wstępie zostało powiedziane, że Julia Hartwig uznaje przymiotnik za najważniejszą część mowy. Z tego powodu będziemy przyglądać się cechom i właściwościom przynależnym przymiotnikowi. Innym powodem, dla którego analizujemy przymiotnik, jest to, że jest on jedną z czterech głównych kategorii pojęciowych (zob. np. Grzegorzczkowska 2012: 128). W rozprawie będziemy dążyć do odpowiedzi na pytanie, jak konkretyzuje się „wyróżniona rola” przymiotnika w języku poetki.

## 2.1. Opis języka naturalnego metodami formalnymi

We współczesnym językoznawstwie funkcjonują różne propozycje odnośnie formalnego opisu zdania, spośród których wyróżnić należy rozpowszechnioną *gramatykę składników bezpośrednich* oraz *formalizm predykatywno-argumentowy*. Pierwsza z nich nawiązuje do gramatyki generatywnej Noama Chomskiego (1957) i traktuje zdanie, jako strukturę liniową złożoną z komponentów zwanych „frazą nominalną” i „frazą werbalną” (*Sentence: NP + VP*). Druga – do której odwołujemy się w niniejszej rozprawie – posiłkuje się zapożyczonym z logiki formalnej modelem zdania opartym na przywołanej powyżej strukturze predykatywno-argumentowej (zob. Vetulani Z., 2004; str. 53-65).

Odwołanie się do metodologii zaczerpniętej z nauk matematycznych nawiązuje do tradycji wywodzącej się z przedwojennej polskiej szkoły logiczno-matematycznej, której jedną z czołowych postaci był Alfred Tarski. Jego fundamentalne prace dotyczyły pojęcia prawdy, a w szczególności związków pomiędzy wyrażeniami języka matematyki a pojęciami do których te wyrażenia się odnoszą (Tarski 1933). Prace te zainteresowały badaczy interesujących się semantyką języka naturalnego. Wśród nich nie można pominąć Richarda Montague czy Charlesa Fillmora, Montague, doktorant Tarskiego na Uniwersytecie Berkeley, podjął śmiałą i uwieńczoną sukcesem próbę wskazania nowego kierunku badań, opisanego w pracy, której tezą jest brak istotnej różnicy pomiędzy językiem naturalnym, a sztucznymi językami formalnymi *English as Formal Language* (Montague 1970).<sup>8</sup> Idee Tarskiego i Montague zostały podchwyczone przez wielu wpływowych w skali światowej lingwistów, jak np. Barbara H. Partee (1990) czy Irena Bellert (1989), a także informatyków interesujących się językiem naturalnym jak np. Alain Colmerauer pionier programowania logicznego w systemach analizy języka (1978) czy jego doktorantka Veronica Dahl (1981).

Badania w zakresie opisu lingwistycznego dla języka polskiego prowadzone w drugiej połowie XX wieku w znacznej mierze pozostały pod wpływem polskiej szkoły logicznej. Obok Tarskiego, należy tu wymienić Ajdukiewicza, Grochowskiego, Karolaka, Świdzińskiego i innych. Model zdania odwołujący się do inspirowanego logiką formalną podejścia predykatywno-argumentowego został wykorzystany w podstawowej monografii pt. *Gramatyka współczesnego języka polskiego. Składnia* pod redakcją

---

<sup>8</sup> "There is in my opinion no important theoretical difference between natural languages and the artificial languages of logicians; indeed, I consider it possible to comprehend the syntax and semantics of both kinds of language within a single natural and mathematically precise theory. On this point I differ from a number of philosophers, but agree, I believe, with Chomsky and his associates." (Montague 1970).

Zuzanny Topolińskiej (1984). Zgodnie z tym podejściem funkcje poszczególnych elementów w zdaniu elementarnym są ściśle określone: *predykat* wskazuje na własność lub relację i otwiera pozycje dla jednego lub więcej *argumentów*. Podkreślmy, iż mówimy tu o zdaniu, które - jak zostało zasygnalizowane w poprzednim rozdziale - jest istotne w kontekście charakterystyki warsztatu Hartwig<sup>9</sup>. Zdanie elementarne ma jeden predykat, czyli podstawę logiczną zdania realizowaną przez różne kategorie gramatyczne, w tym przymiotnik (Vetulani G. 2000: 15, Szumska 2006). Za zdania elementarne przyjmujemy za Karolakiem (patrz Topolińska (red.) 1984: 13) takie wyrażenia, które są samodzielnymi jednostkami informacyjnymi, na które składają się tylko składniki semantycznie konieczne. Innymi słowy takie modele zdań „dają pełną informację o wybranym fragmencie rzeczywistości pozatekstowej. Tym, co konstytuuje zdanie elementarne, co determinuje zdaniowy charakter poszczególnych ciągów wyrażen, jest zatem ich sens, a nie właściwości struktury formalnej” (*ibid*, 19-20).

Zauważmy, że Julia Hartwig operuje często ciągiem przymiotników, z których każdy jest predykatem. Rozważmy przykład:

- (1) *Teraz widać było, jak jest wysoki, ogromny, brodaty, w niebieskiej mycce na głowie i w butach sztywnych [05zawsze powroty]*<sup>10</sup>

Tak skonstruowane zdanie jest de facto równoważne sekwencji zdań elementarnych:

*jak jest wysoki,*  
*(jak jest) ogromny,*  
*(jak jest) brodaty)*  
*itd.*

---

<sup>9</sup> Zob. Rozdział 1.3. niniejszej pracy.

<sup>10</sup> Wszystkie cytaty w niniejszej pracy z twórczości Hartwig pochodzą z elektronicznego korpusu prozy i korpusu poezji. W nawiasach kwadratowych podajemy źródło cytatu, którego forma (numer + tytuł) odpowiada notacji stosowanej w korpusie prozy i korpusie poezji opisanych szczegółowo w rozdziale 3 niniejszej pracy.

## 2.2. Różne realizacje predykatu w zdaniu

W tradycji językoznawczej oraz klasycznej gramatyce szkolnej, analizując zdanie od strony logiczno-gramatycznej, zwraca się przede wszystkim uwagę na czasownik jako element organizujący wypowiedź (K. Polański (red.), 1980). Tymczasem orzekanie, czyli wyrażanie sądów o podmiocie może się realizować również poprzez inne kategorie gramatyczne, tzn. rzeczowniki, przymiotniki i przysłówki (Vetulani G. 2000, 2012; Szumska 2006; Kaliska 2018). Zilustrujmy to zjawisko następującymi przykładami:

- (2) *Rousseau zdobywa sławę* [05zawsze powroty]
- (3) *Jestem tu tylko trochę* [011błyski]
- (4) *Niestety, Daniela nie oddzwoniła* [05zawsze powroty]
- (5) *Próby pojednania, a próby takie zawsze są bolesne, gdy dawny stopień intymności staje się nieosiągalny...* [02apollinaire]

Zauważmy, że zdania (2-5) to przykłady formułowania sądów, których sens realizowany jest odpowiednio przez użycie rzeczownika (*sława* (2)), przysłówka (*trochę* (3)), czasownika (*nie oddzwoniła* (4)), wreszcie przymiotników (*bolesne, nieosiągalny* (5)).

Niekiedy jeden predykat (niosący jeden sens, oddający jedną ideę) realizuje się w użyciu pod postacią różnych struktur gramatycznych (jeśli dany system językowy na to pozwala). W przypadku różnorodności gramatycznej predykatu, wszystkie formy (wszystkie realizacje gramatyczne, tj. czasownikowa/rzeczownikowa, jak np. *zwyciężyć/odnieść zwycięstwo* lub rzeczownikowa/ przymiotnikowa, jak np. *czynić dobro/być dobrym* itd.) niosą ten sam sens (Jędrzejko, 1988, G. Vetulani, 2012), co oznacza, że tożsamość semantyczna zachodzi na poziomie *struktury głębokiej*. Ten głęboki sens oddawany jest w dyskursie poprzez różne *struktury powierzchniowe*.<sup>11</sup> Porównajmy w tym miejscu dwa następujące przykłady:

---

<sup>11</sup> Pojęcia zaczerpnięte z koncepcji semantyki interpretacyjnej J.J. Katza, która stanowiła element włączony do gramatyki generatywno-transformacyjnej Chomsky'ego. (zob. Śliwiński 1990: 16-17). Katz skrytykował założenie Chomsky'ego, że można wygenerować nieskończony zbiór zdań poprawnych gramatycznie na podstawie tylko reguł składniowych bez wnikania w znaczenie komponentów zdania. Dlatego do teorii generatywnej dołożony został „składnik” semantyczny i wówczas rolę znaczeniowórczą przypisano strukturom podstawowym, których zbiór nosi nazwę *struktury głębokiej*. W takim ujęciu jak pisze Śliwiński: *zadanie reguł transformacyjnych polega obecnie na wyprowadzaniu /generowaniu/ struktur powierzchniowych ze struktur głębokich*. Innymi słowy można powiedzieć, że *struktury głębokie* są bardziej abstrakcyjne, dlatego są bliższe sensu wypowiedzi, natomiast *struktury powierzchniowe* to konkretne realizacje wypowiedzi danego języka.

- (6) *Sekwana zieleni się*
- (7) *Sekwana jest gęsto zielona* [05zawsze powroty]

W (6) rolę predykatu pełni czasownik *zieleni się*, natomiast w (7) predykatem właściwym jest przymiotnik *zielona* (odsyłający do cechy, właściwości Sekwany), który współwystępuje w strukturze z czasownikiem *być*, tworząc swoisty związek wyrazowy (tradycyjne orzeczenie imienne). Przyjrzyjmy się jeszcze zdaniom:

- (8) *Czesław doskonale orientuje się w ówczesnym stanie literatury w Polsce*  
[10dziennik 2]
- (9) *Czesław jest doskonale zorientowany w ówczesnym stanie literatury w Polsce*

W zdaniach tych mamy do czynienia z predykatem realizowanym przez czasownik *orientować się* (8) oraz przez imiesłów w funkcji przymiotnikowej *zorientowany* (9). Podane przykłady dowodzą, że mimo różnic strukturalnych, chodzi o tę samą treść predykatywną, o czym może świadczyć wybór jednakowych argumentów (elementów na pozycji podmiotu i dopełnienia), a mianowicie:

*orientować się* (Czesław, stan literatury w Polsce)  
*zorientowany* (Czesław, stan literatury w Polsce)

Skoro typ i liczba argumentów pokrywają się w dwóch różnych strukturach, tzn. w dwóch różnych realizacjach gramatycznych predykatu, podane przykłady można uznać za tożsame. Odpowiada im jeden (ten sam) schemat:

predykat (arg 1, arg2)

Jak zostało powiedziane, omawiane powyżej przykłady to predykaty dwuargumentowe, w których *explicite* wystąpiły argumenty (Czesław, stan literatury w Polsce). Nadmienimy jednak, że niektóre elementy wypowiedzi są z różnych powodów opuszczane. Tym niemniej, pełny opis sensu predykatu wymaga przedstawienia „najdłuższego ciągu argumentów wyznaczających ten sens, tzn. argumentów walencyjnych, w tym również

tych, które w konkretnych realizacjach bywają pomijane” (Vetulani G., 2000: 22). Tak jest w poprawnych składniowo konstrukcjach (11) do (15), eliptycznych w stosunku do (10).

- (10) *Na lotnisku czekał na nas kierowca z Instytutu* [10dziennik2]
- (11) *Na lotnisku czekał kierowca z Instytutu*
- (12) *Na lotnisku czekał kierowca*
- (13) *Czekał na nas kierowca*
- (14) *Kierowca czekał*
- (15) *Czekał*

Przedstawione rozmaite konstrukcje zdaniowe są niezbędne do rekonstrukcji modelu struktury wedle założenia, że dana konstrukcja jest odpowiednikiem jednego sensu. Ponadto własności kategoryjno-morfologiczne, które determinują cechy powierzchniowo-kombinatoryczne „dopuszczają tzw. *użycie absolutne* czyli bez wykładników pojęć uzupełniających (argumentów).” (Szumska 2006: 76). W języku Hartwig, jak zobaczymy w kolejnej części tego rozdziału, przymiotniki są udziałem tego typu zjawisk.

### 2.2.1. Elementy struktury predykatywno-argumentowej<sup>12</sup>

Przejdźmy teraz do charakterystyki głównych elementów logiczno-semantycznych w zdaniu ilustrując je przykładami z korpusu prozy i poezji Hartwig. Przyjmujemy za Wiktorem Markiem, że predykat to:

wyrażenie, które opisuje pewną własność lub relację”. (...) predykat opisujący własność nazywa się predykatem jednoargumentowym (ang. *monadic predicate*), predykat opisujący relację nazywa się, dwu-, trzy-, lub więcej- argumentowym (ang. *dyadic predicate, triadic predicate, etc.*), w zależności od liczby członów danej relacji.  
(*Mała Encyklopedia Logiki* (1970: 229)

W poniższych przykładach predykat jest realizowany jako czasownik nieprzechodni (*śpi, krzyczą, opustoszał*), lub przymiotnik (*czarna, dobry*).

---

<sup>12</sup> Patrz także rozdział poświęcony pojęciu modelu predykatywno-argumentowego w monografii (Vetulani Z., 2004: 55-65)

- (16) *Śpi Prozerpina* [016to wróci]
- (17) *Krzyczą mewy* [04dziennik amerykański]
- (18) *Paryż opustoszał* [022spojrzenie]
- (19) (...) *przestrzeń jest czarna* [05zawsze powroty]
- (20) *Świat jest dobry* [022spojrzenie]

Przedstawione przykłady to predykaty jednoargumentowe  $P(x)$ , które opisują własność lub stan. Natomiast w (21) i (22) predykaty wyrażają relację między argumentami:

- (21) *Styczeń pachnie wiosną.* [04dziennik amerykański]
- (22) *W klon trafiła błyskawica.* [011błyski]
- (23) *Emmanuel ofiarował nam starannie wybrane książki.* [05zawsze powroty]

W przykładzie (22) czasownik *trafić* wyraża relację pomiędzy dwoma argumentami *klon* i *błyskawica* realizującą schemat  $P(x,y)$ . Z kolei w przykładzie (23) czasownik *ofiarować* realizuje schemat trójargumentowy  $P(x,y,z)$ . W schemacie tym argument  $x$  odpowiada tradycyjnemu podmiotowi *Emmanuel*, argument  $y$  dopełnieniu bliższemu, którym jest wyraz *nam*, zaś  $z$  – dopełnieniu dalszemu *starannie wybrane książki*. Predykat, ze względu na swoje właściwości semantyczno-logiczne, decyduje o liczbie oraz naturze współwystępujących z nim argumentów. W powyższych schematach informacja o naturze i innych właściwościach argumentów przypisana jest symbolom  $x$ ,  $y$ ,  $z$ . (zob. Vetulani Z., 2004:55-65)

Zwróćmy uwagę na potencjał powierzchniowych strukturalizacji treści, które prezentują zestaw rozmaitych kategorii gramatycznych i kategorii leksykalno-semantycznych, ale mają też swoje granice i motywacje wynikające z gramatyki. Dlatego to właśnie dynamika zachodząca pomiędzy planem treści (*obiektywnym zjawiskiem pozajęzykowym*) a planem wyrażania interesuje nas najbardziej z punktu widzenia charakterystyki języka danego pisarza/poety. Możemy ją zaobserwować najlepiej w momencie aktualizacji treści. Zamiennosc użycia zdań tożsamy semantycznie, a różniących się w sposobach strukturyzacji treści, podyktowana jest zazwyczaj potrzebami komunikacyjnymi wynikającymi z sytuacji i z rejestru języka. Wybór różnych realizacji powierzchniowych jest determinowany również przez styl, wymagania gatunkowe wypowiedzi literackiej czy cel danego aktu komunikacji. Ewa Jędrzejko zwraca uwagę na to, że zróżnicowanie form wypowiedzi ma wymowę stylistyczną:

„użycie różnych form może mieć na celu również zapobieganie monotonii strukturalnej wypowiedzi” (Jędrzejko 1993: 41). Wybór leksykalny zależy również od funkcji wypowiedzi związanej, między innymi, z kontekstem emocjonalnym. System języka dostarcza bogatego repertuaru ekwiwalentnych środków wyrażania, czyli tak zwanych „synonimów funkcyjnych” (ibid.), por. *szanować kogoś / mieć szacunek dla kogoś*, które to synonimy funkcyjne są powszechne w języku polskim:

*orientować się / zorientowany / orientacja*

*zieleńić się / zielony / zieleń*

*chorować / chory / choroba*

*skąpić / skąpy / skąpstwo*

W składni obserwujemy

istnienie różnych wzorców syntaktycznych do wyrażania tej samej struktury semantycznej (...). Za semantycznie równoważne uznajemy składniki wypowiedzenia, którym na głębokim poziomie można przypisać tę samą charakterystykę funkcjonalną predykatu lub argumentu o określonej roli...

(ibidem, 38 i 44).

Jak już zostało powiedziane pewna liczba jednostek predykatywnych funkcjonuje autonomicznie, tj. bez odpowiedników semantycznych realizowanych przez inne kategorie gramatyczne, jak na przykład predykaty przymiotnikowe *duży, drogi, długotrwały*.

(24) *Ci byli duzi, mocni, niegadatliwi* [02apollinaire]

(25) *Mleko jest bardzo drogie* [04dziennik\_ amerykanski]

(26) *Zapoje Konstantego były zazwyczaj długotrwałe* [10dziennik2]

Przymiotniki te nie mają odpowiedników w formie czasowników w rodzaju *\*dużyć, \*drogić, \*długotrwałować*.

W podsumowaniu powyższych uwag, stwierdzamy, że schemat zdania elementarnego tworzą predykaty oraz współwystępujące z nimi argumenty, które łączy związek syntaktyczny. Argumenty przyjmują pozycje otwierane przez predykaty. Typ i liczbę argumentów definiuje dany predykat. W poniżej przedstawionych

przykładach obserwujemy różne typy argumentów, od form prostych po formy złożone, będące niekiedy zdaniem.

- (27) *Lubiłam spotkania.*
- (28) *Lubiłam te nasze spotkania.* [10dziennik2]
- (29) *Lubiłam, jak odbywały się te nasze spotkania.*
- (30) *Lubiłam spotykać się.*

Podsumowując uwagi dotyczące różnych form predykatywnych (Rozdz. 2.2.), powtórzmy za Karolakiem, iż „aktualizacja znaków następuje w konkretnych okurencjach (wystąpieniach), a ponieważ naturalną formą okurencji znaków jest ich współwystępowanie z innymi znakami, aktualizacja znaków dokonuje się w różnych okazach zdań”<sup>13</sup>. W przypadku predykatów czasownikowych ich aktualizowanie następuje poprzez przyjmowanie końcówek osobowych wskazujących na czas, miejsce i osobę. Z kolei dla predykatów przymiotnikowych i rzeczownikowych funkcję dostarczania informacji gramatycznych (na temat czasu, aspektu, osoby) przejmuje czasownik współwystępujący z predykatem właściwym, który w takim użyciu „traci” swoje prymarne znaczenie. Innymi słowy, predykat, będąc kategorią logiczno-semantyczną, realizuje się w dyskursie pod postacią różnych kategorii gramatycznych, czego dowodem są zaprezentowane wyżej przykłady pochodzące z utworów Julii Hartwig. Przytoczmy jeszcze następującą uwagę Agnieszki Kaliskiej:

Powierzchniowa strukturyzacja treści to inaczej organizacja treści w planie wyrażania dokonująca się za pośrednictwem różnych pod względem przynależności do kategorii gramatycznej i kategorii leksykalno-semantycznej wyrażen języka oraz reguł leksykalno-gramatycznych warunkujących ich współwystępowanie w planie wyrażania, a także sposoby w jaki to współwystępowanie się odbywa. Zjawisko to można nazwać także *językową formalizacją treści*, bowiem treść – będąca ze swej natury abstrakcyjną – daje się uchwycić dopiero przy pomocy powołanych do tego konkretnych form językowych, z których każda obok ogólnokategorialnego składnika treści (*genus proximum*) konkretyzuje pewien odcień znaczeniowy (*differentia specifica*), ujmując tym samym obiektywne zjawisko pozajęzykowe z pewnego mniej lub bardziej szczegółowego punktu widzenia. (Kaliska 2016:41)

---

<sup>13</sup> *Encyklopedia Językoznawstwa Ogólnego* (1995: 29)

### 2.2.2. Formy predykcji przymiotnikowej

Przymiotniki doczekały się sporej literatury przedmiotu, która świadczy o ich niejednorodności i o rozmaitych próbach ich klasyfikacji (Nagórko 1987, Śliwiński 1990).

Dorota Szumska we wstępie rozprawy o przymiotniku jako wyrażeniu predykatywnym (Szumska 2006:22-23) stwierdza, że badania dotyczące kategorii przymiotnika „nie tworzą pełnego semantycznego opisu tej klasy”<sup>14</sup>. Coraz konsekwentniej „za cechy definicyjne przymiotnika jako jednostki płaszczyzny powierzchniowej przyjmuje się jego własności morfologiczne i/lub syntaktyczne” (ibid.). Lecz i w tym aspekcie wyznaczenie obszaru właściwości przymiotnika jest dyskusyjne, gdyż leksemu przymiotnikowego oczekiwać można w każdym miejscu zhierarchizowanej struktury zdania. Szumska, analizując przymiotnik w roli predykatu, stosowała koncepcję opisaną na formule definicji otwartej skupionej na kryterium syntaktycznym i morfologicznym, co poszerzyło klasę badanych jednostek o bliskie funkcjonalnie liczebniki porządkowe, zaimki i imiesłowy, które wchodziły w relację z rzeczownikiem i stanowiły element konstytutywny grupy imiennej. Podobne rozwiązania stosowali Saloni (Saloni i in. 2012)<sup>15</sup>, Szupryczyńska (1980), Śliwiński (1990). W tradycyjnym ujęciu M. Szymczak podaje następującą definicję przymiotnika jako części mowy:

część mowy określająca jakość rzeczownika (np. biały) lub relację (np. szkolny) do rzeczownika (określanego); jej podstawową funkcją składniową jest funkcja przydawki (np. drewniany dom) lub orzecznika (np. dom jest drewniany); odmienia się przez przypadki, liczby, rodzaje i stopniuje się.

(Słownik języka polskiego, M. Szymczak (red.), 1978-81):

Jak zostało powiedziane, charakteryzujemy warsztat Julii Hartwig analizując *użycie* przymiotnika, ale nie przymiotnik jako wyabstrahowany z kontekstu komponent zdania, tylko ze względu na związki, jakie tworzy z innymi kategoriami. Problematyka

---

<sup>14</sup> Odnotować należy, że literatura dotycząca przymiotnika stanowi obszerny rozdział historii językoznawstwa polskiego, prócz zacytowanych, wymienić należy za Szumską prace nad: semantyczną charakterystyką przymiotnika (Nagórko 1987, 1988) związanych z barwą (m.in. Tokarski 1995a), emocjami (m.in. Buttler 1977, 1978, Nowakowska-Kempna 1998, Szumska 2000), cechami psychicznymi (m.in. Puzyńska 1991) czy oceną (m.in. Kreisberg 1999), wreszcie wartościami (m.in. Puzyńska 1992).

<sup>15</sup> Saloni, Woliński, Wołosz, Gruszczyński, Skowrońska

łączliwości składniowo-semantycznej przymiotników z rzeczownikami omówiona została, między innymi, przez Śliwińskiego (1990), który pracował na materiale *Małego słownika języka polskiego* (Skorupka, Auderska, Łempicka 1968). Rezultaty analiz zostały zaprezentowane w *Słowniku semantyczno-syntagmatycznego przymiotników polskich* klasyfikującym przymiotniki z punktu widzenia ich właściwości konotacyjnych i akomodacyjnych (Śliwiński 1993). Zastosowana przez Śliwińskiego metodologia stawia przymiotnik w centrum rozważań, czemu daje wyraz w przyjętej przez niego definicji konotacji.

(Konotacja) traktowana jako jeden z podstawowych terminów w teorii składniowej znaczy dosłownie współoznaczanie (...). Np. przymiotnik nie tylko oznacza daną cechę, ale równocześnie 'współoznacza', czyli konotuje, najogólniej przedmiot, któremu cecha ma ta może przysługiwać, i stąd, pojawiwszy się w tekście, otwiera niejako puste miejsca dla rzeczownika. Podobnie przysłówek otwiera miejsca dla czasownika lub przymiotnika itd. Ten typ konotacji sprecyzowanej przez K. Bühlera możemy nazywać konotacją kategorialno-semantyczną, gdyż polega ona na konotowaniu przez jedne kategorie znaczeniowe jednostek innych kategorii znaczeniowych.

Interesuje nas badanie warsztatu pisarskiego Hartwig w oparciu o analizę użycie przymiotnika w jego funkcji predykatywnej. Zwracając uwagę na charakter predykatywny tego elementu w zdaniu, zgadzamy się z Szumską co do jego dominującej funkcji w zdaniu (Szumska 2006).

Dla ilustracji pojęcia struktury predykatywno-argumentowej zdania posłużyliśmy się przykładami, od (31) do (34), zaczerpniętymi z korpusu twórczości Julii Hartwig, dla pokazania funkcjonowania czasowników w predykatywno-argumentowej strukturze zdania. W niniejszej rozprawie, z powodów przedstawionych powyżej, interesujemy się przede wszystkim przymiotnikiem w funkcji predykatywnej i dlatego dalsze rozważania będą ilustrowane przykładami, w których predykatem jest przymiotnik.

Przymiotnik nazywa cechy przedmiotów, a także określa relacje do innego przedmiotu lub czynności i stanowią element konstytutywny grupy imiennej (Heinz 1965, Szupryczyńska 1980, Karolak 2001a, Grzegorzyczkowa 2012). Wspomniane cechy w sposób bezsporny przysługują tylko pewnej grupie przymiotników, które z tego powodu w literaturze określane są *właściwymi* (Karolak, 2001: 42).

W poniższych przykładach 31 – 34 obserwujemy użycie przymiotników w funkcji predykatu.

- (31) *Nasze reakcje na sztukę nigdy nie są czyste.* [06pisane przy oknie]
- (32) *Była to wspaniała osoba, dzielna, bohaterska, bezkompromisowa.*  
[09dziennik1]
- (33) *Dla Apollinaire'a to jednomiesięczne opóźnienie musiało być bolesne*  
[02apollinaire]
- (34) *Charakterystyczne dla wybrzeży bretońskich są ciemne, niemal czarne, ostro  
zarysowane skały [...]* [05zawsze powroty]

Przytoczone powyżej przykłady użycia przymiotnika predykatywnego odwołują się do jego funkcji w wypowiedzi, skomentowanej przez Doroszewskiego w następujący sposób (1972: 400): „przymiotniki pełniące funkcję predykatywną mogą być zawsze użyte w funkcji atrybutywnej, co właściwie oznacza, że określenia: przymiotniki predykatywne, przymiotniki atrybutywne odnoszą się nie do przeciwstawnych sobie klas wyrazowych, ale najwyżej do różnic w zakresie funkcji poszczególnych elementów tych klas” i wskazuje na przykłady takie jak *ten dom jest biały — to jest dom biały*. Różnica między obydwojoma wyrażeniami rysuje się w interpretacji natury rzeczownika poprzez współwystępujący przymiotnik, gdzie w pierwszym przykładzie *ten dom* jest znakiem desygnatu podmiotu, o którym ktoś orzeka, że jest *biały*. W drugim przykładzie – *dom biały* – przymiotnik *biały* pełni rolę wyróżniającą rzeczownik *dom* spośród innych rzeczowników (obiektów) jakimi są domy.

Przymiotnik jako kategoria dominująca w strukturze zdaniowej w pozycji adnominalnej otwiera miejsce dla rzeczownika. Poniżej kilka przykładów z korpusu prozy Hartwig ilustrujących omawiane zjawisko:

- (35) *Na zupełnie **czystym** niebie, **wspaniały** pokazowy zachód słońca.*  
[04dziennik amerykański]
- (36) *I **mila, cicha** żona o **okrągłej i czystej** twarzy dziewczyny.* [04dziennik amerykański]
- (37) *To z **tamtych odległych** czasów zapamiętałam **czule** westchnienie poety*  
[10dziennik2]
- (38) ***Czarne** wyświecone ubranie, **czarny** zarost, **czarne** półbuty włożone na **gole** stopy.* [05zawsze powroty]

Dodajmy jeszcze przykłady z korpusu poezji:

- (39) *na rozległej przestrzeni łagodnych pagórków oblanym mlecznym światłem snu* [010 nie ma odpowiedzi]
- (40) *A dzieje się tak tylko po to, żeby mały najmniejszy, ledwo widoczny pajęczek mógł wykonać swe zadanie* [002wolne ręce]
- (41) *Są łyzy łatwe. I łatwe słowa. I tanie współzucie.* [013mówiąc nie tylko od siebie]
- (42) *Wjeżdżamy w kraj łakomych marzeń.* [001pożegnania]

Mimo predyspozycji przymiotnika do tworzenia związków z rzeczownikiem oraz panującego przekonania o tym, że „najistotniejszą cechą przymiotnika jest jego zależność składniowa od rzeczownika” (Kurkowska 1953), to szczególnie interesuje nas u Hartwig pozycja przyczasownikowa (wg podziału Szupryczyńskiej 1980) i związek przymiotnika z czasownikiem, w którym, przymiotnik pełni rolę predykatu a czasownik jest elementem posiłkowym.

Podstawową cechą takich jednostek przymiotnikowych jest możliwość analizowania ich i klasyfikowania tak jak czasowniki, ponieważ przymiotniki mają własność selekcyjonowania argumentów (jednego lub kilku).<sup>16</sup> Pod tym względem zachowują się jak czasowniki (por. rzeczownik predykatywny Vetulani G. 2000, 2012). Zjawisko to – jak precyzuje Karolak (patrz Vetulani G. 2000: 23) – „opisywane jest mianem *walencji* lub *wartościowości syntaktycznej*”. Przyjmujemy zatem, że w takim przypadku to przymiotniki określają strukturę zdania, w którym zostały użyte.

Najczęściej spotykamy przymiotniki przy czasowniku *być / zostać*, które w tradycyjnej gramatyce szkolnej stanowią łącznik w orzeczeniu imiennym, natomiast przymiotnik pełni rolę orzecznika. Tymczasem, w myśl założeń składni predykatywno-argumentowej, konstrukcje zbudowane ze słowa predykatywnego i wprowadzającego je *czasownika podporowego* (Vetulani G. 2000, 2012) stanowią specjalny przypadek

<sup>16</sup> Por. Na temat cech selekcyjnych przymiotników pisała również Kaliska (2016:142): „Cechy selekcyjne przymiotników są podstawą do wyodrębnienia klasy przymiotników predykatywnych (fr. *adjectifs prédicatifs*) i oddzielenia ich od „wtórnie” predykatywnych przymiotników nazywających cechy (np. *mały, duży, kolorowy, itd.*), których prymarną funkcją jest funkcja określenia w grupie imiennej. W opracowaniach polskich językoznawców można odnaleźć takie rozróżnienia (np. Bańko 2002: 61). Przymiotniki predykatywne są zazwyczaj dwuargumentowe. Na przykład: *Jan był ciekaw* (wyników egzaminu | \*E), *Jan był łaskaw* (pomóc biedakowi | \*E), *Jan był rad* (ze spotkania z Marią | \*E), *Jan jest godzien* (pierwszej nagrody | \*E), *Jan jest winien* (tej kłótni | \*E), *Jan jest świadom* (konsekwencji swego czynu | \*E), *Jan był pewien* (słuszności powziętej decyzji | \*E). W użyciu przyrzeczownikowym, „te same” przymiotniki przybierają nieco inne formy: *ciekawa książka, łaskawy król, wesoły człowiek, godna śmierć, świadomy człowiek.*”

związku przymiotnika z czasownikiem, Inne czasowniki aktualizujące przymiotniki predykatywne to warianty aspektowe i rodzajowe takie jak: *wydawać się vs. zdawać się vs. stawać się vs. pozostawać*.

Jak zostało już powiedziane klasa przymiotników jest niejednorodna i niełatwo wśród jej elementów wyodrębnić jednostki predykatywne. Analiza predykcji przymiotników przeprowadzona przez Szumską (2006: 121) wskazuje, że:

klasa adiektywów wyznaczona na podstawie aplikacji kryteriów morfologiczno-syntaktycznych obejmuje korelaty formalne zarówno jednostek predykatywnych, jak i jednostek niepredykatywnych [...], podstawową funkcją przymiotnikowych wyrażen predykatywnych i niepredykatywnych jest współtworzenie fraz imiennych które funkcjonują jako sekundarnych wykładników propozycji, to znaczy współtworzenie, czyli wyrażen predykatywno-argumentowych tworzonych przez rzeczowniki.

Co więcej Szumska traktuje „test transpozycji z pozycji adnominalnej do orzecznika” jako swoistą miarę sprawdzającą konstytutywność przymiotnika jako wykładnika predykatu, zatem możliwość współwystępowania w związku przymiotnika z czasownikiem świadczy najczęściej o predykatywnej roli tego przymiotnika.

System języka niejako samoistnie wskazuje na pomocne kryterium składniowe, ponieważ przymiotniki predykatywne występują najczęściej z czasownikiem, który jest pozbawiony swego podstawowego znaczenia (por. Vetulani G. 2012: 23). Ponadto obserwacja łączliwości jednostek przymiotnikowych z daną formą czasownikową pozwala na wytyczenie granic pomiędzy określonymi klasami odniesień w ramach pól semantycznych, związanych na przykład z emocjami, opiniami, ocenami itd., co jest dla nas szczególnie ciekawe ze względu na charakterystykę warsztatu pisarskiego Hartwig. Zgadza się z Szumską, że występowanie wariantów realizacyjnych jednostek predykatywnych jest faktem, który prowadzić winien do zastanowienia się „nad uwarunkowaniami i konsekwencjami dokonywanych z ich udziałem operacji redukcji warstwy powierzchniowej przekazu werbalnego”, który rozbudowuje „paradygmat interpretacyjny tego przekazu”, oraz „że użycie różnych form może być semantycznie lub/i pragmatycznie relewantne, a więc pełnić rolę komunikatywną.” (Szumska 2006: 76).

Przyjrzyjmy się przykładom konstrukcji<sup>17</sup>, pochodzących z korpusu tekstów Julii Hartwig, w których realizowana jest przymiotnikowa strukturalizacja predykatu.

Po pierwsze jako orzeczenie imienne:

- (43) *Obrazowanie staje się bardziej wyzwolone, pełniejsze, bogatsze*  
[08podziękowanie za gościnę]

Po drugie, przymiotnik predykatywny może występować jako przydawka orzekająca (inaczej nazywana okolicznikową), która jest elementem dodanym, ale charakteryzującym podmiot i czasownik:

- (44) *Picasso jest w jej wspomnieniach mały, czarny, przysadzisty, niespokojny, oczy ciemne, głębokie, gesty niezręczne, kobiece ręce, źle ubrany, zaniedbany* [02apollinaire]
- (45) *Więc to jest ów przyjaciel Apollinaire'a, oddany, wierny do ostatka, skromny i nienawidzący fałszu.* [02apollinaire]

Przykłady (44) i (45) obrazują charakterystyczną dla Hartwig zwięzłość stylu oddającą swoiste napięcie. Sens tych zdań można wyrazić przy pomocy szeregu zdań prostych (od (46) do (50) dla zdania (44) oraz od (51) do (57) dla zdania (45) nieoddających jednak tego napięcia pomimo wykorzystania tych samych predykatów przymiotnikowych.

- (46) *Picasso jest w jej wspomnieniach*  
(47) *Picasso jest mały*  
(48) *Picasso jest czarny*  
(49) *Picasso ma ciemne oczy.*  
(50) *Picasso ma głębokie oczy.*  
(51) *Więc to jest przyjaciel.*  
(52) *Więc to jest ów przyjaciel.*  
(53) *Więc to jest przyjaciel Apollinaire'a*

---

<sup>17</sup>Przedstawiona typologia – uzupełniona przykładami z korpusu Hartwig – została zaczerpnięta z omówienia Kaliskiej (2016:138), która analizowała predykcję przymiotnikową w języku polskim i francuskim. Również o predykatywnej strukturalizacji przymiotnikowej pisali m.in. Wojtasiewicz (1972) i Węgrzynek (1995). Pierwszy z wymienionych badaczy dokonał wstępnej klasyfikacji przymiotników predykatywnych ze względów semantycznych.

- (54) *Więc to jest przyjaciel oddany.*
- (55) *Więc to jest przyjaciel wierny do ostatka.*
- (56) *Więc to jest przyjaciel skromny.*
- (57) *Więc to jest przyjaciel nienawidzący fałszu.*

Podobne rozumowanie dotyczące strukturalizacji predykatywnej przymiotnika w roli przydawki prezentuje między innymi Szumska (2006)<sup>18</sup>.

Po trzecie, predykat przymiotnikowy w funkcji przydawki dopowiadającej występującej w pre- lub postpozycji w stosunku do członu nominalnego najczęściej oddzielonego przecinkiem, tak jak w przykładach (58) i (59):

- (58) *piętrowy biały dom z kolumnadą, ani piękny, ani brzydki* [01z niedalekich podróży]
- (59) *opinie utrzymane w tonie chłodnym i umiarkowanym* [06pisane przy oknie]

Po czwarte, przymiotnik predykatywny w funkcji przydawki przyorzecznikowej, która określa orzecznik opisywany przez Topolińską (1984: 332) jako „informacyjnie pusty i stanowi tylko formalne oparcie dla określeń”, jak w przykładzie (22)

- (60) (...) *cała ta historia rodzinna jest sprawą bolesną i dyskretną* (...) [02apollinaire],

które można uznać za odpowiednik komunikacyjny zdania *Cała ta historia rodzinna jest bolesna i dyskretna*.

Przytoczone zdania wskazują na szczególne upodobanie autorki do posługiwania się przymiotnikami występującymi w szeregach, gdzie są one oddzielane przecinkami lub spójnikami (por. Szumska 2006 179-181). Z tego powodu w rozprawie zwracamy uwagę na operatywność przymiotników i zgadzamy się ze Skibskim:

---

<sup>18</sup>U Szumskiej (2006: 122) czytamy: „Predyspozycja przymiotnika, jako wyrażenia predykatywnego do formalizowania treści predykatywnej w postaci związku niezdaniotwórczego a z presupowaną asercją, nawet z możliwością pominięcia wykładnika konstytutywnego sprawia, że jest on środkiem do realizacji operacji redukcji powierzchniowej, a grupa imienna, w której skład wchodzi, wyrażeniem predykatywno-argumentowym umożliwiającym stosowanie do potrzeb komunikacyjnych zróżnicowanie zakresu strukturalizacji reprezentowanych treści propozycjonalnych”.

relacje składniowe, jako reguły konkretnego tekstu, (ale też reguły wiersza) stają się podstawową motywacją oryginalnych znaczeń, które jednak nie unieważniają standardowych mechanizmów językowych, a więc powtarzalności, typowości i kategoryalności, jako podstaw sensu konstruowanego przez odbiorcę.

(Skibski 2017: 143)

### 2.3. Świat Julii Hartwig odzwierciedlony w przymiotnikach

Przyjęta przez nas hipoteza badawcza stanowi, że obraz świata artysty determinuje jego warsztat twórczy. Analizując świat Julii Hartwig odzwierciedlony w przymiotnikach, odwołujemy się do Teorii Językowego Obrazu Świata (JOS) (Bartmiński (1990), Pajdzińska (2001, 2004), Grzegorzczkowa (2001, 2015), Tokarski (2001, 2004) i inni.

Przytoczymy w tym miejscu jedną z pierwszych definicji JOS, która wydaje się najlepiej opisywać poruszane przez nas zagadnienia związane z twórczością Hartwig:

JOS to pewien zespół sądów mniej lub bardziej utrwalonych w języku, zawartych w znaczeniach słów lub przez te znaczenia implikowanych, który orzeka o cechach i sposobach istnienia obiektów świata pozajęzykowego. W tym sensie JOS jest utrwaleniem zespołu relacji zawartych w językowym ukształtowaniu tekstu, a wynikających z wiedzy o świecie pozajęzykowym.

(Bartmiński, Tokarski 1986: 72).

W innym miejscu Bartmiński (2010: 158) stwierdza wprost, że językowy obraz świata to „zespół sądów o świecie”.

Podobnego zdania był Doroszewski (1972: 401) pisząc, że „konstrukcje syntaktyczne są świadectwami rozmaitych sposobów różnicowania rzeczywistości”, czyli sposobami opowiadania lub określania otaczającego świata. Spostrzeżenia Doroszewskiego (ibid.), dotyczące form wyrazowych, wpisują się w długą historię dociekań filozoficznych i filologicznych, traktujących język (z całym skomplikowaniem struktury leksykalno-semantycznej i gramatycznej) jako rodzaj klasyfikatora świata. Zatem język, używając formuły Doroszewskiego, *różnicuje rzeczywistość*, utrwala doświadczenie poznawcze człowieka i według określenia Grzegorzczkowej „kształtuje swoisty językowy obraz świata.” (2012: 60).

Relacja między rzeczywistością a nazywającym ją językiem była zagadnieniem ważnym dla Hartwig, ponieważ, według poetki, sposób powiązania przez pisarza

(artystę) obydwu tych płaszczyzn, charakteryzuje jego warsztat. W eseju dotyczącym Pierre'a Reverdy'ego pt. *W niewoli poetyki* stwierdza ona:

O każdym poecie mówi zapewne najwięcej jego stosunek do rzeczywistości. Jest to niewrażliwy punkt każdej poetyki. (Hartwig 2006: 342).

W innym miejscu pisze również:

Rzeczywistość widzialna stanowiła dla mnie zawsze stały punkt odniesienia. (...) Wierzę, że można tak opisać widok, rzecz, sytuację, by czytelnik zrozumiał, że mają one jeszcze inne, ukryte sensy. (Hartwig 2000: 6).

Zależności między rzeczywistością a odnoszącym się do niej językiem omawia współczesna teoria referencji, która traktuje język jako formę poznania i opanowania myślowego świata. Jak wskazuje Grzegorzczkowska (ibid.), *referencja* opisuje właściwość wyrażenia użytego, które odnosi do przedmiotu w świecie, natomiast *denotacja* to odniesienie do klas zjawisk, czyli właściwość wyrażenia nieużytego a zawartego w słowniku. Interesuje nas to, co zostało już wypowiedziane w tekście Hartwig, zatem *użycie* przymiotnika i jego własności referencyjne. Używanie danych wyrażen językowych jest odbiciem mechanizmów mentalnych człowieka, jego postrzegania i kategoryzowania rzeczywistości, która najpierw odbywa się w umyśle człowieka (ibid: 136).

Chcemy wskazać na rzeczywistość Julii Hartwig czyli na światy, jakie wyłaniają się z zapisów poetki, a także na sposób, w jaki te światy konkretyzują się w języku. Komentatorzy dzieła Hartwig podkreślała, że obcowanie przez nią z tym, co ją otacza, odbywa się przede wszystkim poprzez percepcję wzrokową. W ten sposób ogląda Paryż, postrzegając miasto przez ramę okna, niczym obraz. Świat opisany w jej utworach zanurzony jest w historię sztuki poprzez odniesienia do konkretnych malarzy. Poetka przywołuje zmysłowe doświadczenia obcowania ze sztuką, nade wszystko z malarstwem, ale też z poezją czy muzyką (Wiśniewski 2013). W jej utworach odnajdujemy częste odniesienia do pejzaży, portretów, ale też „malarskie” opisy krajobrazu naturalnego (Telicki 2005, Przybylski 2015, Kosowska 2014).

Twórcze związki z Francją, szeroki horyzont wrażeń, związanych również z surrealizmem<sup>19</sup>, powoduje, że niektórzy badacze zwracali uwagę na dwa światy spotykające się w tej twórczości. O spotkaniu dwu światów, a mianowicie „Normy Dnia” i „Pasji Nocy”, omawiając twórczość Hartwig z początkowego okresu, pisał Jerzy Kwiatkowski:

Przemianę, jaka dokonała się w tej poezji, określić można najkrócej jako jej przenicowanie się: z zewnątrz – do wewnątrz. Z jawy – w sen. Z jasności – w ciemność. Z Normy Dnia – w Pasję Nocy.

(Kwiatkowski 1973: 124)

Z kolei Marta Flakowicz, omawiając poematy prozą, zauważa:

To poetka paradoksów i graniczności, klasyk ze swobodą posługujący się mową nocy, spokrewniony z surrealizmem, zamkniętym w precyzyjnej poetyce zdania. Sfera irracjonalna, przeczuciowa, podświadoma współgra z poezją rozumu, intelektu, stanowiąc o często dramatycznym, ale jakże płodnym napięciu tej twórczości. Świat poetycki Hartwig łączy realne z onirycznym, wykracza poza pojmowalne i jednoznaczne. Stąd poematy, osadzone w widzialnym konkrecie, zapełniają się jednocześnie zjawami, serafinami i duchami zmarłych. Zadziwia rozmach tego pisarstwa.

(Flakowicz 2009: 214)

Z czasem w utworach Hartwig, to właśnie świat realny a nie oniryczny przeważa w utworach. Pojawiają się tematy inspirowane pobytem w Stanach Zjednoczonych. Amerykańska podróż była spotkaniem z zupełnie odmienną rzeczywistością społeczno-kulturową Nowego Świata, która wpłynęła na sprozaizowanie frazy oraz zwrot w kierunku tematyki zwykłych, codziennych zajęć sytuacji. Hartwig czerpała pomysły z prasy, jako że jej wiersze amerykańskie częstokroć powstawały przy przeczytaniu aktualnych doniesień z życia społeczeństwa amerykańskiego. Tytuły niektórych wierszy mogłyby z powodzeniem, ze względu na ich syntetyczność, być nagłówkami prasowymi. Wymieńmy kilka tytułów z tomu *Wiersze amerykańskie* (Hartwig 2002):

---

<sup>19</sup> Marcin Telicki (2009: 11) wyróżnia trzy okresy twórczości poetki (1945–1978, 1980–2002, 1999–2007), które odpowiadają triadycznemu ujęciu zagadnień służących jako klucze do zrozumienia wierszy z tych okresów i wskazuje na pamięć, tożsamość i surrealizm.

*Gomez na dachu,*  
*Wino pana Petersona,*  
*Odwołanie zeznań,*  
*Pan Porter się upił,*  
*Murzyńska rodzina,*  
*Odwiedziny u starego poety w Santa Barbara,*  
*Mówi muzyk z Preservation Hall w Nowym Orleanie*  
(por. Legeżyńska 2015, Winiecka 2015).

Ważnym elementem świata budowanego czy opowiadanego przez poetkę są ludzie. Bohaterowie jej utworów to poeci, malarze, muzycy, kloszardzi, ale też osoby przypadkowe. W tym kontekście Magdalena Baran mówi o tomach Hartwig, że są „gęsto zaludnione” (2017:204). Sporo portretów ludzkich odnajdujemy w reportażach i dziennikach oraz monografiach (*Apollinare* (Hartwig 1962) i *Gerard de Nerval* (Hartwig 1972)). Jej świat nazwisk oraz opisów fizjonomii i charakterów zaskakuje starannością, wnikliwością, a czasami demonstracyjną szczerością.

W opisach postaci i charakterów Julia Hartwig szczególnie obficie sięga po przymiotniki:

- (61) *Bardzo chudy, w tweedowej beige marynarce, czarna gęsta czupryna, mała głowa, drobna cofnięta broda, oczy bardzo przenikliwe, czarne, duży, wyrazisty nos* [04dziennik amerykański]
- (62) *Amerykańska studentka o ciężkim spojrzeniu, ubrana w grubą kurtkę i spodnie, leniwie spogląda, nieruchoma, na rysującego ją młodzieńca o twarzy pięknej i zdeprawowanej* [05zawsze powroty]
- (63) *Pamiętam dobrze tę smukłą, nieśmiałą, niezwykle kruchą blondynkę o cienkich palcach, długim nosie, delikatnym zarysie pastelowej twarzy* [05zawsze powroty]
- (64) *Twarze pospolite, ubrania codzienne, uderzająco niegustowne* [04dziennik amerykański]
- (65) *Oboje otyli, ona brzydka, o szerokiej bladej twarzy, on przystojny, dobrze ułożony, ciężki, jak bywają chłopcy włoscy* [04dziennik amerykański]

Wśród opisywanych postaci są nie tylko osoby obecne w ówczesnym życiu Hartwig, ale też często osoby zmarłe, które poetka – poprzez literacki zapis – pragnie ocalić od zapomnienia.

Ostatni okres twórczości pisarskiej Hartwig, głównie poetyckiej, charakteryzuje wypowiedź bardziej refleksyjna, niż referująca z pozycji świadka wydarzenia oraz sytuacje (por. Flakowicz--Szczyrba 2012, Czerska 2014). W rozmowie z Juliuszem Kurkiewiczem na temat tomu *Jasne niejasne* poetka przyznaje:

Odmienność tego tomiku od poprzednich widzę w przewadze refleksji nad obrazem, w upartym poszukiwaniu rozwiązań, nierzadko zakończonych poczuciem bezradności. Bo wiele jest odpowiedzi i w życiu, i w sztuce, które, choć mają pozór jasności, wcale nie są jasne.

(Hartwig 2010).

Jest to też świat lektur i dialogu myślowego z tym, co jest zawarte w twórczości innych pisarzy i poetów.

Wydaje się, że ostatecznym rezultatem konfrontowania się poetki z rzeczywistością jest afirmacja świata, takiego jakim jest. Tytuł mowy, jaką wygłosiła poetka w Poznaniu z okazji nadania jej Doktoratu Honoris Causa, brzmiał: *Dawać do siebie dostęp zachwyтови* (Hartwig 2015). Przemówienie to zostało zawarte w tomie studiów nad jej twórczością zatytułowanym sugestywnie *Pochwała istnienia*. Z kolei Anna Legeżyńska (2017) swojej eseistycznej monografii będącej biografią Julii Hartwig nadała dużo mówiący tytuł *Wdzięczność*.

Reasumując, możemy powiedzieć, że Hartwigowy język konkretyzuje takie obszary otaczającego świata, jak sytuacje, krajobrazy, ludzie, doświadczenia (np. sztuki, podróży<sup>20</sup> czy lektury), które układają się w sieć pojęć określonych wyrażen leksykalnych<sup>21</sup>. Od Renaty Grzegorzczkovej (2012: 133) dowiadujemy się, że „w strukturze semantycznej słownictwa odzwierciedla się nie tylko swoista kategoryzacja świata, ale także subiektywna interpretacja wyróżnionych zjawisk”.

---

<sup>20</sup> Na temat podróży i miejsc ważnych dla poetki z punktu widzenia zarówno biografii jak i twórczości pisała Elżbieta Dutka (2016) w książce zatytułowanej: *Centra, prowincje, zaułki. Twórczość Julii Hartwig jako auto/bio/geo/grafia*.

<sup>21</sup> Patrz teoria pól wyrazowych J. Trierera (1931), która dzieli słownictwo na grupy tematyczne oraz omawia relacje semantyczne, w jakie wchodzi leksyka.

## ROZDZIAŁ 3

### Korpus prozy i korpus poezji Julii Hartwig

Niektóre przedstawione w Rozdziale 2 obserwacje dotyczące instrumentarium leksykalnego poetki zostały poczynione w drodze analizy obszernego materiału w postaci korpusu złożonego z jej prac twórczych (o metodzie badań korpusowych pisaliśmy ogólnie w Rozdziale 1.2.1.). W niniejszym rozdziale omówiona zostanie zawartość utworzonego przez nas korpusu prozy i poezji Julii Hartwig oraz zasady i metody jego zbierania.

#### 3.1. Korpusy tekstowe i korpusy elektroniczne pod kątem lingwistyki korpusowej

##### 3.1.1. Pojęcie korpusu i jego znaczenie

Za Graeme Kennedym (1998) przez *korpus* rozumiemy zbiór tekstów pisanych lub mówionych, służący za źródło do analizy i opisu językoznawczego<sup>22</sup> (Kennedy 1998). Współcześnie podobną definicję przyjmuje np. Elżbieta Hajnicz, która pisze, że korpus stanowi „podzbiór tekstów zorganizowany w taki sposób, by umożliwić sprawne wydobywanie i weryfikowanie informacji na temat dowolnego, dobrze określonego podzbioru języka” (Hajnicz 2011)<sup>23</sup>.

Przyjęte w niniejszej rozprawie podejście metodologiczne wpisuje się w nurt *lingwistyki korpusowej* nawiązującej do metodologii nauk ścisłych, a co za tym idzie do badań materiału dostępnego empirycznie, takiego jak teksty, zdania, słowa (Vetulani Z.

---

<sup>22</sup> „In the language sciences a corpus is a body of written text or transcribed speech which can serve as a basis for linguistic analysis and description.”

<sup>23</sup> Hajnicz (2011) wskazuje na zbiory fiszek, jako metodę gromadzenia danych, poprzedzającą współczesne cyfrowe korpusy danych przetwarzalnych elektronicznie. Fiszki, czyli zaświadczenie użycia wraz z cytatem, były dowodem na występowanie danego zjawiska językowego. Przechowywano je w zbiorach ze względu na jakąś określoną prawidłowość językową. Poprzez analizę materiału zawartego w konkretnych zbiorach fiszek można było np. udowodnić występowanie pewnych wzorców językowych w określonym systemie językowym lub w jego podzbiorze.

i in., 2010: 21)<sup>24</sup>. Oznacza to, że formułowanie reguł i wniosków przebiega na podstawie obserwacji przejawów zjawisk zachodzących w świecie, jak np. występowania określonych form językowych. Podzielamy w tym względzie opinię Witosz (2009), że styl autora badamy obserwując zbiór wypowiedzi jednostkowego nadawcy, czyli o autorze wnioskujemy na podstawie jego działalności pisarskiej.

Należy zaznaczyć, że podejście empiryczne odwołujące się do korpusów jako reprezentacji badanej rzeczywistości językowej sięga połowy XVIII wieku, kiedy to zaczęto wykorzystywać biblię jako źródło do tworzenia i badania konkordancji przez Alexandra Crudena około roku 1736 (Cruden 1769; patrz także Kennedy 1998:14), czy też jako źródło cytowań<sup>25</sup> na potrzeby tekstów religijnych lub religioznawczych.

Z kolei na przełomie XIX-go i XX-go stulecia, w pracach nad słownikiem języka angielskiego *Oxford English Dictionary* (1928), wykorzystano korpus literatury angielskiej od roku 1000. „Badania korpusowe były intensywnie prowadzone już w latach 20-tych XX-go wieku w UK oraz USA, zaś dekadę później stały się podstawą lingwistyki kwantytatywnej rozwijanej w Szkole Praskiej.” (Vetulani Z. 2010 i in.)<sup>26</sup>.

Dla języka polskiego, prekursorskim był referencyjny korpus Idy Kurcz (Kurcz, Lewicki, Sambor, Woronczak. 1974a, b, 1976, 1977, Lewicki, Masłowski, Sambor, Woronczak 1975), który umożliwił utworzenie pierwszych dobrze udokumentowanych list frekwencyjnych dla szeregu odmian stylistycznych polszczyzny. Współcześnie, korpusy języka polskiego (jak np. NKJP, wersja otwarta, udostępniona publicznie; Przepiórkowski i in. 2012) są wykorzystywane między innymi do rozbudowy leksykalnej bazy danych PolNet (Vetulani Z. i inni 2010) o kolokacje werbo-nominalne i inne (Vetulani G. 2012). O ile pierwsze duże słowniki morfologiczne języka polskiego, jak na przykład POLEX /1994-1996/ (patrz np. (Vetulani Z., Walczak, Obrębski, Vetulani G. 1998), także (Vetulani Z, Vetulani G. 2019)) czy niezależnie rozwijany, nieco późniejszy słownik Pollex (Bogacki 1997) tylko pośrednio odwoływały się do danych empirycznych, zaś bezpośrednio czerpały materiał ze słowników tradycyjnych jak np. słownik Doroszewskiego (1958), Szymczaka (1981), czy Słownik PWN (2000) (ten ostatni o charakterze tradycyjnym, lecz na nośniku elektronicznym), to już współczesne słowniki wykorzystywane dla konkretnych aplikacji informatycznych wymagających

---

<sup>24</sup> Vetulani Z., Marciniak, Obrębski, Vetulani G., Dąbrowski A., Kubis, Osiński, Walkowska, Kubacki, Witalewski

<sup>25</sup> Patrz też [https://pl.wikipedia.org/wiki/Wikiprojekt:Biblia/Skr%C3%B3ty\\_i\\_wersety](https://pl.wikipedia.org/wiki/Wikiprojekt:Biblia/Skr%C3%B3ty_i_wersety)

<sup>26</sup> Patrz przypis 25

modelowania kompetencji językowej w zakresie ograniczonej dziedziny i charakteru obsługiwanych potrzeb komunikacyjnych posługują się korpusami uzyskanymi w drodze eksperymentu (Walkowska, 2012), (Vetulani Z., Vetulani G. 2019), względnie otrzymanymi przez ekstrakcję podkorpusów z korpusów referencyjnych (patrz: Vetulani Z, Witkowska, Canbolat 2018). Odpowiednio zestawione korpusy aplikacyjne, są bezpośrednio wykorzystywane do budowy zasobów leksykalnych a także gramatycznych danej aplikacji.

Na przestrzeni ostatnich dwóch dekad wzrosło zapotrzebowanie środowiska naukowego i technicznego na tworzenie przetwarzalnych informatycznie zasobów elektronicznych dla języków do tej pory zasobów takich pozbawionych (ang. Less-Resourced Languages /LRL/). Tematyka ta jest promowana przez szereg organizacji i inicjatyw, jak:

- stowarzyszenie European Language Resources Association (ELRA) i jej agenda ELDA (ELRA 1995; <http://www.elra.info/en/about/>)
- sieć doskonałości FlaReNet
- inicjatywa MetaNet since 2010 (<http://www.meta-net.eu/>),
- organizacja CLARIN (<https://www.clarin.eu/>)

Lingwistyka korpusów dla LRL pojawiała się systematycznie na konferencjach LREC i LTC, a także na workshopach z serii WILDRE (WILDRE 2012) związanych z LREC (od 2012) oraz warsztatów LRL organizowanych każdorazowo (od roku 2011) w ramach konferencji LTC (patrz: Vetulani Z., Choukri, Mariani, Paroubek). Ostatnim doniosłym wydarzeniem międzynarodowym w tym zakresie była organizowana pod patronatem i przy wsparciu UNESCO konferencja Language Technology for All (LT4ALL 2019; <https://en.unesco.org/LT4All>).

Na wydarzeniach międzynarodowych wymienionych powyżej pojawiają się coraz częściej doniesienia o narzędziach technologii informatycznych wspierających badania dorobku literackiego w rozmaitych językach. Jednym z ważnych zadań w tej dziedzinie jest opracowywanie monograficznych, referencyjnych korpusów obejmujących całość działalności twórczej poszczególnych autorów.

Niniejsza rozprawa jest pomyślana jako przyczynek do tego zadania.

### 3.1.2 Podstawowe cechy korpusów

Podstawowe cechy każdego korpusu to reprezentatywność, zrównoważenie a także otwartość (poza przypadkami szczególnymi).

#### Reprezentatywność

W definicji słownikowej terminu *reprezentatywność* (ang. *representativeness*) znajdujemy odesłanie do przymiotnika *reprezentatywny*, czyli taki, który „dobrze reprezentuje kogoś albo coś i ma charakterystyczne cechy reprezentanta jakiejś zbiorowości” (Szymczak 1981). W odniesieniu do korpusu znaczy to tyle, że korpus jest reprezentacją języka rozumianego za de Saussure’em jako *parole*, czyli rzeczywistych słów wypowiedzianych i zapisanych w postaci tekstów charakterystycznych dla ogółu języka *langue*. Zbiór takich tekstów może mieć różny zakres określany ze względu na cel budowania danego korpusu. Intuicyjnie wiemy, że na przykład korpusy narodowe mają oddawać charakter języka jako złożonego systemu zaświadczonego w tekstach i transkrypcjach mowy, a ponadto z założenia jako „dobro wspólne” powinny być dostępne publicznie i bezpłatnie. We wstępie do Narodowego Korpusu Języka Polskiego (NKJP) czytamy:

Od korpusu reprezentatywnego należy oczekiwać odzwierciedlenia języka w całej jego różnorodności, tak by każdy typ tekstu miał jakiś, choćby niewielki udział w korpusie [...] Uznajmy tedy (sic!), że zrównoważenie oznacza tyle, iż żaden z typów tekstów nie obejmuje więcej niż połowę korpusu.  
(Górski, Łaziński 2012: 29)

Ponadto reprezentatywność korpusu jest zależna od przyjętej przez twórców korpusu tekstowego perspektywy, ponieważ zebrane dane mogą być reprezentatywne dla:

- populacji twórców tekstów,
- populacji tekstów,
- produkcji tekstów lub recepcji tekstów (według typologii Rafała L. Górskiego i Michała Łazińskiego *ibid.*: 26).

Z reprezentatywnością korpusu wiąże się również problem jego rozmiaru. Założenie, że duży korpus lepiej reprezentuje język i jego różnorodność niż mały, nie zawsze jest

słuszne, gdyż obserwacja zjawisk językowych wymaga przede wszystkim zbioru tekstów dobrej jakości, o której świadczy dobór elementów do zasobu (por. Kennedy 1998)<sup>27</sup>. W literaturze wskazuje się, że dla analizy fonologicznej wystarczy korpus mniejszy niż dla składniowej (zob. Polański 1995:290), a najobszerniejsze korpusy wymagane są dla badań statystycznych czy stylometrycznych (por. zbiory tekstów literackich wykorzystywanych przez Jana Rybickiego 2017: 6)<sup>28</sup>. Podsumowując powyższe omówienie stwierdzamy, że przez *korpus reprezentatywny* rozumiemy taki zbiór tekstów, który odzwierciedla określoną rzeczywistość językową i reprezentuje język ze względu na jego właściwości gramatyczne i leksykalne. Przygotowany przez nas korpus prozy i poezji Julii Hartwig konstruowaliśmy tak, by odzwierciedlał produkcję tekstów Autorki oraz był reprezentatywny dla istotnych cech warsztatu pisarskiego Hartwig.

## Zrównoważenie

Z zagadnieniem reprezentatywności łączy się pojęcie zrównoważenia (ang. balance), które w niniejszej pracy definiujemy jako taki dobór tekstów w korpusie, by żaden typ tekstu nie dominował nad pozostałymi elementami. Zrównoważenie korpusu wynika z założeń twórcy danego zasobu i najczęściej pod uwagę brane są takie cechy, jak odpowiednia reprezentacja różnych typów tekstu, na przykład: ze względu na kanał dystrybucji, zróżnicowanie tematyczne tekstów, ponadto różnorodność stylistyczną związaną z gatunkiem wypowiedzi, a także rozpiętość chronologiczną – daty publikacji tekstów (por. uwagi na temat ułożenia i oznaczania tekstów metadanymi w tworzonym korpusie tekstów polskich do 1500 roku, Leńczuk 2020).

Przykładem korpusu zrównoważonego może być korpus *Słownika frekwencyjnego polszczyzny współczesnej* (Kurcz i in. 1990)<sup>29</sup>, na który składają się równe próbki z tekstów popularnonaukowych, drobnych wiadomości prasowych, publicystyki, prozy

---

<sup>27</sup> Por. Kennedy (1998: 68): „A vast collection of texts is not necessarily a corpus from which generalizations can be made. A huge corpus does not necessarily 'represent' a language or a variety of a language any better than a smaller. At this stage we simply do not know how big the corpus needs to be for general or particular purposes. Rather than focusing so strongly on the quantity of data in a corpus, compilers and analysts need also to bear in mind that the quality of the data they work with is at least as important.”

<sup>28</sup> „W roku 2014 pięćset polskich tytułów wydawało się całkiem sporym wyborem tekstów. Dziś warto powrócić do tej samej myśli, bo w prezentowanych poniżej badaniach wykorzystano ponad pięciokrotnie większy zestaw, zawierający teksty polskie od *Kazań świętokrzyskich* po Szymickowej *Tajemnicę domu Helclów*, polskie powieści, poezję epicką i dramat oraz polskie przekłady z języka angielskiego (w tym liczne przekłady szekspirowskie), francuskiego, rosyjskiego, niemieckiego, czeskiego, włoskiego, portugalskiego, hiszpańskiego, węgierskiego i tureckiego.” (Rybicki 2017:6)

<sup>29</sup> Kurcz, Lewicki, Sambor, Szafran, Woronczak

artystycznej i dramatu artystycznego z lat 60 XX wieku. Na świecie natomiast za wzorcowy<sup>30</sup> uznawany jest Narodowy Korpus Brytyjski (ang. British National Corpus, BNC). Nadmienmy, że dążenia autorów korpusów do ich zrównoważenia i reprezentatywności są wyzwaniem, które stanowi pole do ożywionej dyskusji wśród badaczy, ponieważ są to cechy, które się często wykluczają.

## Otwartość korpusu

Pożądaną cechą korpusu jest otwartość, czyli możliwość aktualizowania i rozszerzania korpusu zgodnie z potrzebami i oczekiwaniami. Przyjmujemy następującą definicję otwartości korpusu: korpus będziemy nazywali *otwartym*, jeśli będzie zbiorem danych systematycznie aktualizowanych. Korpusy otwarte gromadzą teksty wszystkich odmian i stylów współczesnego języka, jak również języka historycznego. Warto zaznaczyć, że cyfrowa obecnie forma zbiorów korpusowych, ułatwia ich wszelkie modyfikacje, co wpływa na zwiększenie otwartości korpusów na nowe obszary tematyczne lub zbiory powiększane o publikacje literackie, prasowe czy branżowe. Przykładem korpusu otwartego jest korpus komunikatów meteorologicznych TAUM METEO w wersji rozszerzonej stworzony na potrzeby pierwszego wdrożonego systemu tłumaczenia maszynowego (Kittredge 1982)<sup>31</sup>. Kittredge zwrócił uwagę na praktyczną wartość specjalistycznych korpusów reprezentujących warianty języka angielskiego, które nazywał *podjęzykami* (ang. *sublanguage*), często o charakterze specjalistycznym lub technicznym, używane w wyodrębnionych dziedzinach, takich jak prognozy pogody, przepisy kulinarne czy instrukcje techniczne. Ze względu na otwartość podjęzyków wskazywał on jednocześnie na konieczność systematycznych aktualizacji odpowiadających im korpusów.

Korpusy otwarte nazywane są również *korpusami monitorującym* ze względu na strukturę i zawartość zasobów, które służą ciągłej rejestracji i obserwacji najnowszych zmian w języku. „Korpus monitorujący, który przestaje być aktualizowany, pozostaje jedynie zamkniętym korpusem diachronicznym, a jego funkcja referencyjna ulega nieuchronnemu osłabieniu w miarę upływu czasu” zauważa Piotr Pęzik i dodaje, że

---

<sup>30</sup> Określenie za Hajnicz (2011:16): „Najbardziej chyba znanym i najczęściej cytowanym w literaturze jest Narodowy Korpus Brytyjski (ang. British National Corpus, BNC). Jego rola jest tak istotna, że nie będzie wielką przesadą uznać go za korpus wzorcowy.”

<sup>31</sup> Systemem TAUM METEO służył do bieżącego tłumaczenia krótkich komunikatów meteorologicznych angielskiego na język francuski na potrzeby lotnictwa kanadyjskiego.

korpusy monitorujące uznawane są za specjalną odmianę korpusów referencyjnych. Dla przykładu badacz wymienia korpus monitorujący angielszczyzny amerykańskiej COCA, który powiększył się w latach 2010–2017 z około 400 do 560 mln segmentów wyrazowych, natomiast korpus monitorujący angielszczyzny globalnej NOW powiększa się o 140–160 mln słów miesięcznie (Pęzik 2020: 134)<sup>32</sup>. Dla języka polskiego publicznie dostępnym, stale aktualizowanym korpusem monitorującym polszczyzny jest obecnie MoncoPL2 (ibid.)<sup>33</sup>.

## Korpusy referencyjne

Terminem *korpusu referencyjnego* określamy zasób danych tekstowych, który cechuje szerokie pokrycie tematyczne, jest reprezentatywny dla dużej liczby zjawisk oraz może stanowić źródło generowania podkorpusów zorientowanych na aplikacje. Poza wymienionymi powyżej przykładami korpusów referencyjnych-monitorujących, to wielomilionowe korpusy języków narodowych, takie jak Narodowy Korpus Brytyjski (ang. British National Corpus, BNC) czy Narodowy Korpus Języka Polskiego (NKJP), co do zasady pełnią funkcję korpusów referencyjnych dla danego języka naturalnego. Wydaje się, że w przyszłości możliwe będzie zespolenie ogólnych korpusów referencyjnych z reprezentacjami tekstów gromadzonych w ramach korpusów monitorujących<sup>34</sup>.

---

<sup>32</sup> „Podstawowym parametrem ilościowym korpusu monitorującego jest tempo przyrostu (ang. rate of flow, por. Sinclair, 1996), mierzone zazwyczaj w segmentach słów przypadających na jednostkę czasu.” (Pęzik 2020)

<sup>33</sup> Dodajmy dla uzupełnienia, że: „Podstawowym źródłem danych indeksowanych w korpusie MoncoPL są kanały RSS4 udostępniane przez ponad 1500 serwisów informacyjnych, branżowych oraz platform blogowych, z których około 600 było aktywnych w lipcu 2019 roku. Poza źródłami dynamicznymi do indeksu MoncoPL zostały jednorazowo włączone niektóre dostępne na otwartej licencji korpusy, z których największym jest *Polski Korpus Sejmowy* (Ogrodniczuk, 2017).”

<sup>34</sup> Tak w podsumowaniu o korpusie pisze Pęzik: „Celem podjętym w tym artykule było omówienie budowy i funkcji, ale również wykazanie zasadności utrzymywania dużego korpusu monitorującego polszczyzny, jakim jest w pewnym zakresie MoncoPL. Aktualizacja tego korpusu jest planowana przynajmniej do czasu rozpoczęcia prac nad kolejną wersją NKJP. W przekonaniu autora, docelowo to właśnie korpus narodowy powinien pełnić funkcję monitorującą, również dla tych rejestrów i odmian polszczyzny, których odpowiednia reprezentacja wymaga znacznych nakładów i zasobów.” (ibid.)

### 3.1.3. Korpusy tekstowe na świecie

Momentem przełomowym było powstanie w roku 1964 korpusu Uniwersytetu Brown (ok. 1 000 000 słów)<sup>35</sup> – pierwszego wielkiego korpusu cyfrowego (zob. Vetulani Z. i in. 2010:21)<sup>36</sup>. Twórcami tego korpusu byli Henry Kučera i Nelson Francis, którzy tym samym dali początek lingwistyce komputerowej (Kučera, Francis 1967). Korpus Browna zapoczątkował całą rodzinę korpusów zwaną potocznie Rodziną Korpusu Browna (Brown Family Corpus). Jest to rozszerzona wersja pierwotnego korpusu Browna, który obejmuje również Lancaster-Oslo/Bergen Corpus (LOB), brytyjski odpowiednik Browna, a także Frown i FLOB, odpowiedniki Browna i LOB z lat 90. XX wieku. Brown Family Corpus składa się z 6 milionów słów w amerykańskim i brytyjskim angielskim.

Wśród największych istniejących obecnie baz językowych należy wymienić korpusy dla języków:

- angielskiego
  - American National Corpus (11 milionów słów)<sup>37</sup>,
  - British National Corpus (100 milionów)<sup>38</sup>,
  - Oxford English Corpus (2,1 miliardy)<sup>39</sup>,
  - PennTreebank<sup>40</sup>,
  - i wiele innych
- niemieckiego – German Reference Corpus *DeReKo* (50,6 miliarda słów)<sup>41</sup>;
- czeskiego – Czech National Corpus (4 miliardy)<sup>42</sup>;
- rosyjskiego – Russian National Corpus (300 milionów)<sup>43</sup>;
- hiszpańskiego – rodzina korpusów Corpus del Espanol (660 milionów)<sup>44</sup>.

Wiele korpusów zostało utworzonych w celach komercyjnych przez firmy oferujące usługi z zakresu inżynierii języka. Zaslugującym na uwagę przykładem firmy

---

<sup>35</sup> <https://www.sketchengine.eu/brown-corpus/> (dostęp 20.07.2021)

<sup>36</sup> Vetulani Z., Marciniak, Obrębski, Vetulani G., Dąbrowski, Kubis, Osiński, Walkowska, Kubacki, Witalewski

<sup>37</sup> [https://www.sketchengine.eu/oanc\\_masc-corpus/](https://www.sketchengine.eu/oanc_masc-corpus/) (dostęp 20.07.2021)

<sup>38</sup> <https://www.english-corpora.org/bnc/> (dostęp 20.07.2021)

<sup>39</sup> <https://www.sketchengine.eu/oxford-english-corpus/> (dostęp 20.07.2021)

<sup>40</sup> <https://catalog.ldc.upenn.edu/LDC99T42> (dostęp 20.07.2021)

<sup>41</sup> <https://www.ids-mannheim.de/digspra/kl/projekte/korpora/> (dostęp 20.07.2021)

<sup>42</sup> <https://www.korpus.cz/> (dostęp 20.07.2021)

<sup>43</sup> <https://ruscorpora.ru/old/en/corpora-intro.html> (dostęp 20.07.2021)

<sup>44</sup> <https://www.corpusdelespanol.org/size.as> (dostęp 20.07.2021)

specjalizującej się w średniej wielkości korpusach dla wielu języków jest działająca od 2003 r. firma Lexical Computing Ltd. założona przez Adama Kilgariffa). Lexical Company dysponuje pakietem korpusów wielomilionowych (często powyżej 100 milionów słów) dla 52 języków.

Równoległe do powstawania odpowiednio dużych cyfrowych korpusów referencyjnych, uważnych obecnie za niezbędne do tworzenia modeli lingwistycznych poszczególnych języków, na potrzeby zorientowanych tematycznie aplikacji informatycznych tworzone są korpusy dziedzinowe, często małych lub średnich rozmiarów, w ścisłej zależności od modelu języka wykorzystywanego przez daną aplikację.

Lingwistyka korpusowa znajduje się w ciągłym rozwoju i nic nie wskazuje na to, by ta sytuacja miała ulec zmianie. Z jednej strony obserwuje się nieustanne dążenia do powiększania oraz ulepszenia istniejących już korpusów, z drugiej zaś do budowania korpusów dla języków, które nie mają jeszcze (lub mają, lecz w ograniczonej formie) reprezentacji tekstowej w postaci cyfrowej

Od pewnego już czasu, stymulacją do tworzenia korpusów jest też wzrastające zapotrzebowanie na odpowiednio duże bazy tekstów ze strony zorientowanych na konkretne zastosowania aplikacji informatycznych wykorzystujących technologie sztucznej inteligencji, a w szczególności sztucznych sieci neuronowych wymagających dużych ilości danych treningowych. Szczególnym przykładem są prace w zakresie tłumaczenia automatycznego, które w przypadku tekstów technicznych, administracyjnych czy prawniczych osiągnęły stosunkowo wysoki poziom, między innymi poprzez wykorzystywanie sieci neuronowych i korpusów równoległych (Kennedy 1998).

### **3.1.4. Elektroniczne korpusy tekstowe dla języka polskiego<sup>45</sup>**

W historii polskiego językoznawstwa szczególną rolę odegrał półmilionowy korpus opracowany przez zespół badaczy Idę Kurcz, Andrzeja Lewickiego, Jadwigę Sambor i Jerzego Woronczaka przy okazji prac nad polskim słownikiem frekwencyjnym tekstów pisanych z lat 1963-1967. Prace te uchodzą za pierwsze wykorzystanie praktyczne

---

<sup>45</sup> Zestawienie elektronicznych źródeł językowych dla języka polskiego (stan na marzec 2017 roku) prezentuje między innymi *Przewodnik po elektronicznych zasobach językowych dla polonistów* (zob. Pałka, Kwaśnicka-Janowicz 2017) oraz zestawienie na stronie internetowej Instytutu Podstaw Informatyki PAN: <http://clip.ipipan.waw.pl/LRT?action=show&redirect=lrt>

cyfrowego korpusu tekstów. Obliczenia dotyczące zebranego materiału przeprowadzało wówczas Centrum Obliczeniowe Katedry Metod Numerycznych Uniwersytetu Wrocławskiego (Kurcz 1976: 105).

Istotne informacje, dotyczące korpusu Idy Kurcz i jej zespołu, zawiera artykuł Janusza Bienia i Mariusza Wolińskiego, który przytaczamy w dłuższym fragmencie:

Korpus słownika frekwencyjnego to pięć zestawów próbek po 100 000 słów wylosowanych z autentycznych tekstów z lat 1963–1967 należących do 5 stylów — tekstów popularnonaukowych, drobnych wiadomości prasowych, publicystyki, prozy artystycznej i dramatu artystycznego — na potrzeby badań frekwencji słów języka polskiego. Pierwotnie korpus miał formę taśm papierowych wyperforowanych na dalekopisie (czego konsekwencją był brak rozróżnienia małych i dużych liter). Został on wczytany do komputera przez Bronisława Roślawskiego (wówczas na Uniwersytecie Gdańskim) i zapisany na taśmie magnetycznej; niestety, w trakcie tej operacji do korpusu wkradły się pewne przekłamania. Taśma magnetyczna została zapisana na komputerze ODRA 1204 w standardzie, który szybko wyszedł z użycia. W związku z tym taśma z korpusem trafiła w ręce Krzysztofa Szafrana, który w Instytucie Informatyki Uniwersytetu Warszawskiego na podstawie list frekwencyjnych dla poszczególnych stylów (...) opracowywał tzw. tom zbiorczy, opublikowany jako Słownik frekwencyjny polszczyzny współczesnej (...). Krzysztof Szafran za pomocą specjalnie przygotowanego programu odczytał taśmę na komputerze SM-4 i zapisał jej zawartość na bardziej nowoczesnych nośnikach, w wyniku czego korpus stał się dostępny również na dyskietkach stosowanych w komputerach osobistych.

(Bień, Woliński 2003: 1)

Korpus Idy Kurcz posłużył początkowo do utworzenia list frekwencyjnych dla różnych odmian polszczyzny z lat sześćdziesiątych XX-go wieku (Kurcz, Lewicki, Sambor, Woronczak, 1974a, b, 1976, 1977, Lewicki, Masłowski, Sambor, Woronczak, 1975), co z kolei doprowadziło do opracowania polskiego słownika frekwencyjnego (Kurcz i in. 1990)<sup>46</sup>.

Kolejnym ważnym momentem było stworzenie konserwowanego w sposób ciągły i liczącego ponad 200 milionów słów korpusu IPI PAN, który był pierwszym dużym polskim korpusem znakowanym morfosyntaktycznie (Przepiórkowski 2004)<sup>47</sup>. W roku 2012 został sfinalizowany projekt Narodowego Korpusu Języka Polskiego (NKJP)

---

<sup>46</sup> Kurcz, Lewicki, Sambor, Szafran, Woronczak

<sup>47</sup> <http://korpus.pl/> (dostęp 20.07.2021)

(Przepiórkowski, Bańko, Górski, Lewandowska-Tomaszczyk 2012)<sup>48</sup>. NKJP jest rezultatem prac kilku zespołów badawczych skupionych w różnych ośrodkach naukowych w Polsce. Ich wcześniejsze zasoby, a mianowicie: wspomniany korpus IPI PAN, ponadto zasoby Korpusu Języka Polskiego PWN oraz korpusu PELCRA zasiliły NKJP.

Przy tworzeniu złożonych systemów informatycznych, wymagających modelowania kompetencji komunikacyjnej, zakładającej wykorzystanie języka naturalnego, zachodzi potrzeba dysponowania szeregiem wyspecjalizowanych korpusów tekstowych bądź korpusów mowy (speech). Za przykład dla języka polskiego może posłużyć system POLINT -112-SMS, gdzie na potrzeby modelowania komunikacji językowej, poza ogólnodostępną wersją Narodowego Korpusu Języka Polskiego IPI PAN, wykorzystane zostały wytworzone na potrzeby projektu korpusy specjalistyczne jak np.: korpus dialogów pozyskanych w drodze eksperymentu, korpus komunikatów SMS, korpus nagrań interwencyjnych zgłoszeń telefonicznych (Vetulani i inni 2010)<sup>49</sup>. Korpusy takie są rzadko dostępne w przestrzeni publicznej czy opisywane w publikacjach.

Jak piszą twórcy Narodowego Korpusu Języka Polskiego (Przepiórkowski i in. 2012: 4):

Komputery umożliwiają magazynowanie i analizowanie dużo większych zbiorów materiału językowego niż tradycyjne metody językoznawcze. Ponadto, co jest szczególnie ważne, pozwalają na weryfikowanie intuicji językowych, pokazując użycie i frekwencję zarówno form krótszych, jak i wzorców współwystępowania różnych form językowych i ich częstotliwości użycia. Pozwalają także na szeroko zakrojone badania języka różnych grup użytkowników, jak i kwantyfikację wyników badań dialektologicznych czy socjolingwistycznych na dużą skalę w zależności od typu dyskursu, stylu, czy indywidualnych preferencji użytkownika języka.

### **3.2. Po co nam elektroniczny korpustwórczości Julii Hartwig?**

Teksty, wszelkie zapisy literackie, podyktowane określoną sytuacją komunikacyjną, są napisane przez nadawcę (autora) dla odbiorcy (czytelnika). Natomiast korpus jest to zbiór tekstów wykorzystywany w celu analizy określonych zjawisk, poszukiwania

---

<sup>48</sup> <http://nkjp.pl/> (dostęp 20.07.2021)

<sup>49</sup> Vetulani Z., Marciniak, Obrębski, Vetulani G., Dąbrowski, Kubis, Osiński, Walkowska, Kubacki, Witalewski

i rozpoznawania pewnych wzorców czy też badań literaturoznawczych (patrz też Rybicki 2014, 2017, Moretti 2018). Konteksty historyczno-literackie, biograficzne, interpretacja poszczególnych utworów pełnią wówczas rolę drugorzędą. W tendencjach, jakie kształtowały teorię literatury i poetykę, takie postrzeganie tekstu bliskie było strukturalizmowi. Traktowanie zbioru tekstów literackich jako korpusu, i przykładanie do niego specyficznych w badaniach korpusologicznych metod prowadzi do analizy nastawionej na szukanie prawidłowości i wzorców. Rozpoznane cechy, wzorce, formy znane z klasycznych podziałów na przykład na gatunki, rodzaje literackie zyskują szansę „udowodnienia”, przyłożenia ich do „faktów” uzyskanych z zastosowania metod kwantytatywnych. Zaobserwowane cechy stanowią wówczas swoistą „jednostkę miary” do poszukiwania związków pomiędzy poszczególnymi dziełami, których być może dotąd nie zestawiano ze sobą.

Określony podzbiór systemu językowego udokumentowany w tekstach, jakim są teksty literackie, jest specjalistycznym korpusem literackim (lub korpusem języka artystycznego), który pozwala na analizę zjawisk występujących w języku literackim.

Pozyskanie tego typu danych językowych podyktowane było chęcią przyspieszenia wyszukiwania i przeszukiwania niezbędnych danych oraz prezentacji rezultatów badań w uporządkowany i przystępny sposób (Vetulani Z. i in., 2010)<sup>50</sup>.

### **3.3. Charakterystyka korpusu prozy i poezji Julii Hartwig**

Materiał badawczy wykorzystany w niniejszej rozprawie to zbiór tekstów, obejmujący ponad sześćdziesiąt lat (1954-2016) lat różnogatunkowej działalności pisarskiej Julii Hartwig. Stanowi on znakomity przykład posługiwania się językiem literackim. Uzyskany korpus zawiera ok 800 000 słów i składa się z dwóch wydzielonych podkorpusów: dla prozy i poezji. W korpusie uwzględniliśmy teksty artystyczne, publicystyczne i paradokumentalne stanowiące trzon bogatej spuścizny autorki. Poniżej przedstawiamy ogólną charakterystykę materiału włączonego do korpusu, a także wyjaśniamy, dlaczego niektóre z utworów Julii Hartwig nie zostały do korpusu wprowadzone.

---

<sup>50</sup> Vetulani Z., Marciniak, Obrębski, Vetulani G., Dąbrowski, Kubis, Osiński, Walkowska, Kubacki, Witalewski

### 3.3.1. Korpus prozy

Na zasób tekstów zawarty w korpusie prozy składają się dzienniki, reportaże, monografie literackie, zbiór esejów i zbiór felietonów. Zawartość korpusu prozy przedstawia się następująco:

- W całości do korpusu włączone zostało drugie wydanie dziennika pt. *Zawsze powroty* (2005)
- W korpusie znalazły się monografie związane z Francją:
  - *Apollinaire* (wydanie 1972) z materiałem językowym z drugiego wydania oraz
  - *Gerard de Nerval* (1972).
- Reportaże zaliczane do literatury faktu:
  - najwcześniej wydany zbiór pt. *Z niedalekich podróży* (1954), który pomieścił relacje z kilkunastu podróży po Polsce i prezentował odbudowę kraju po wojnie oraz
  - materiał z drugiego wydania książki pt. *Dziennik amerykański* (1980, 2015).
- W korpusie znalazł się również:
  - zbiór felietonów *Pisane przy oknie* (2004), który powstawał od stycznia 2001 do marca 2004 i pierwotnie był drukowany w kolejnych numerach czasopisma „Więź” oraz
  - felietony prezentujące portrety przyjaciół i wybitnych pisarzy zebrane w tomie *Wybrańcy losu* (2006), które były publikowane wcześniej na łamach „Tygodnika Powszechnego” i „Gazety Wyborczej”.
- Ponadto w korpusie znalazł się obszerny dwutomowy *Dziennik* (2011) *Dziennik tom 2* (2014), zawierający opis życia w Polsce, w tym opisy krótkich wyjazdów od marca 2008 do listopada 2009 oraz od stycznia 2011 do grudnia 2013 roku.
- W korpusie prozy uwzględniliśmy ponadto zbiór esejów pt. *Podziękowanie za gościnę. Moja Francja* (2006), w którym zebrane zostały omówienia dotyczące poetów francuskich.

Osobnym problemem jest zaliczenie lub nie do korpusu przypadków celowych powtórzeń własnych utworów (na zasadzie autocytowania) lub przytaczania dzieł innych autorów w całości, jak to miało miejsce w przypadku zbioru esejów pt. *Podziękowanie za gościnę. Moja Francja* (2006), w którym omówieniom utworów poetów francuskich towarzyszy wybór ich wierszy. W tym wypadku materiał zamieszczony w korpusie został okrojony o wiersze francuskich poetów w tłumaczeniu Hartwig, a także o powtórzoną część zapisów dziennikowych z tomu *Zawsze powroty* (zaczepniętych z wydań 2001, 2005).

### 3.3.2. Korpus poezji

Prócz opisanego powyżej zasobu prozy stworzony został korpus tekstów poetyckich składający się z kilkunastu tomików poezji oraz tomów prozy poetyckiej. W sumie na korpus poezji składają się 22 pozycje, które w większości są zróżnicowane gatunkowo w obrębie wiersza, prozy poetyckiej i *błysków*.

Na korpus poezji składają się następujące dzieła:

- Debiutancki tom wierszy *Pożegnania* (1956) w dwóch częściach:
  - *Z różnych lat* (trzydzieści siedem wierszy przy niektórych zanotowane zostały data i miejsce powstania, najstarsze dwa z 1946 roku, również dwa z 1947 i 1948) i
  - *Powitanie córki* (dwanaście wierszy bez dat).
- Kolejny tom zawarty w korpusie nosi nazwę *Wolne ręce*, opublikowany dopiero w roku 1969 w trzynaście lat od debiutu. Tom ten składał się również z dwu części, pierwszej prezentującej nowy dorobek (czterdzieści dwa nowe utwory, w tym poematy poetyckie) i wybór trzynastu wierszy z dawnych lat.
- Kolejny włączony do zbioru tom pt. *Dwoistość* pochodzi z roku 1971 i zawiera nowe utwory, w których dominuje proza poetycka.
- Na tom następny zatytułowany *Czuwanie* (1978), prócz nowych wierszy, składały się poematy poetyckie oraz wydzielona część zatytułowana *Americana*, obejmująca poezję z lat 1970-1974, tj. wiersze powstałe w czasie pobytu w Stanach Zjednoczonych.
- Piąty tom włączony do korpusu, noszący tytuł *Chwila postoju* z 1980 roku ma odmienny charakter, ponieważ zawiera wyłącznie poematy napisane prozą i zapisy dwu- i jednozdaniowych myśli, które zapowiadały nową formę, a mianowicie *błysków*.

- Kolejny tom w korpusie wydany w 1987 roku pt. *Obcowanie* był różnorodny i zawierał nowe wiersze prozę poetycką, *błyski*, a także publikowane już americana z późniejszych (1979-81).
- Następną pozycją z 1992 roku, tom pt. *Czułość*, w którym zestawione zostały ze sobą wiersze, prozy poetyckie i kolejne błyski oraz pojedyncze wiersze datowane na rok 1982, 1984, 1985.
- Kolejną pozycję korpusu stanowi zbiór czternastu wierszy pt. *Nowe wiersze* (1995), będący dodatkiem do wyboru wierszy dawniejszych zatytułowanego *Nim opatrzy się zieleń*.
- Następnym tomem pt. *Zobaczone* wydany został w 1999 roku i składał się głównie z wierszy błysków.
- Należy zaznaczyć, że od roku 2000 wzrasta intensywność publikacji poetyckich i na przestrzeni szesnastu lat ukazuje się aż trzynaście tomów, które znalazły się w omawianym korpusie; w tej liczbie należy wyróżnić cztery zbiory wyłącznie błysków. Są to:
  - *Błyski* (2002),
  - *Wiersze amerykańskie* (2002),
  - *Mówiąc nie tylko od siebie. Poematy prozą* (2003),
  - *Bez pożegnania* (2004),
  - *Zwierzenia i błyski* (2004),
  - *To wróci* (2007),
  - *Trzecie błyski* (2008),
  - *Jasne niejasne* (2009),
  - *Gorzkie żale* (2011),
  - *Zapisane* (2013),
  - *Nowe błyski [w:] Błyski zebrane* (2014) (z tego tomu włączono tylko ostatnią część),
  - *Spojrzenie* (2016).

Pominiętych celem uniknięcia powtórzeń i zachowania równowagi korpusu zostało jedenaście wyborów wcześniej opublikowanej poezji Julii Hartwig (tomiki z okresu od 1980 do 2014).<sup>51</sup>

Nadmienimy, iż osobna analiza mogłaby zostać (w przyszłości) przeprowadzona na tekstach przekładów oraz na takim materiale, jak: komentarze do tłumaczeń, wywiady, opublikowana korespondencja, gościnne wypowiedzi komentujące literaturę i sztukę, które ukazały się w książkach innych autorów oraz wszelkie komentarze i recenzje rozproszone w czasopiśmie. Takie badania byłyby cennym dopełnieniem poczynionych tutaj obserwacji (zob. pełna bibliografia twórczości Julii Hartwig; Dutka 2016).

### **3.3.3. Reprezentatywność i zrównoważenie korpusu prozy i poezji Julii Hartwig**

Reprezentatywność i zrównoważenie są uważane za atrybuty jakości korpusów.

Dobry korpus musi być bowiem kolekcją danych językowych wybranych i zorganizowanych według jawnych kryteriów lingwistycznych w celu ustanowienia próby języka odzwierciedlającej różnicowanie systemu, a ponadto powinien być stale uaktualniany.

(Vetulani G. 2012: 53),<sup>52</sup>

Zależało nam, aby zebrany korpus był przede wszystkim reprezentatywny ze względu na gatunek wypowiedzi i chronologię. Kwestia zrównoważenia składowych korpusu (proza vs poezja) została wzięta pod uwagę podczas analiz ilościowych (zob. rozdział 4).

Włączenie do korpusu 10 tomów prozy oraz 22 tomów poetyckich uzasadnione jest faktem, iż utwory te stanowią indywidualne, autorskie dzieła zaświadczone o wyobraźni twórczej Julii Hartwig.

Jednocześnie zrezygnowaliśmy z włączenia do korpusu bogatego dorobku translatorskiego, ponieważ analiza wyborów leksykalnych tłumaczki wymagałaby

---

<sup>51</sup> Poczynione ograniczenia wynikają z teorii tworzenia korpusów, w myśl której rozmiar korpusu powinien być dostateczny dla reprezentacji zjawisk językowych, będących przedmiotem badań (por. Vetulani Z., Witkowska, Menken, 2015, 2018 a także np. Hajnicz 2011).

<sup>52</sup> W przypadku korpusu dzieł niezującego autora, a z tą sytuacją mamy do czynienia w niniejszej rozprawie, ostatnia uwaga dotyczyć będzie ewentualnie nowo odkrytych dzieł.

drobiazgowych analiz materiału z wykorzystaniem teorii przekładu prozy i poezji, bez bezpośredniego wpływu na osiągnięcie celów niniejszej rozprawy.

Zrezygnowaliśmy również z gromadzenia i digitalizowania wywiadów, komentarzy, recenzji i wystąpień rozproszonych w różnych publikacjach, traktując je jako świadectwo dodatkowej aktywności literackiej, poboczne w stosunku do głównej działalności twórczej Julii Hartwig. Z podobnych powodów postanowiliśmy pominąć opublikowane tomy jej korespondencji z m. in. Herbertem, Miłoszem i Turowiczem.

Nie została także wzięta pod uwagę twórczość Hartwig adresowana do dzieci ze względu na odrębność tego gatunku.

Jednocześnie przyjęto, że powtarzające się w zbiorach poezji wiersze w tej samej formie, bądź przeredagowane, pozostaną w korpusie. Założyliśmy, że decyzje poetki dotyczące powtórzenia utworów w nowych wyborach poezji miały uzasadnienie w tym, że tomiki poetyckie należy uznać za spójną całość. Redundancje zawarte w zgromadzonym korpusie przyjmujemy jako potwierdzenie wyborów leksykalnych dokonywanych przez poetkę, a także jako dopełnienie badań nad kategorią przymiotnika.<sup>53</sup> Jednocześnie wygenerowane w ten sposób powtórzenia tekstu w korpusie zostały wzięte pod uwagę na etapie normalizacji tekstu dla celów analiz ilościowych (patrz rozdział 3.4.3).

## **3.4. Tworzenie korpusu prozy i poezji Julii Hartwig**

### **3.4.1. Metoda pozyskiwania korpusu**

Obecnie niemal każdy tekst w jakiejś fazie powstawania znajduje się na nośniku elektronicznym, dlatego wydawać by się mogło, że pozyskanie elektronicznego korpusu twórczości Julii Hartwig jest stosunkowo łatwym zadaniem, tym bardziej, że w domenie publicznej funkcjonują dzieła wydane w druku bądź elektronicznie. Dlatego najszybszym sposobem pozyskania materiałów było zwrócenie się do wydawców dzieł Hartwig z prośbą o udostępnienie wersji elektronicznych tekstów na potrzeby badań naukowych, co zostało uczynione. Tym niemniej rozwiązanie to na ogół okazywało się

---

<sup>53</sup>Analiza zebranego materiału w perspektywie badania wariantów redundantnych i zmian w publikowanych tekstach zasługuje na osobny projekt. Badania ukierunkowane w ten sposób zostały przeprowadzone dotychczas między innymi na dorobku literackim Czesława Miłosza, w których analiza dotyczyła tzw. poetyki wariantów tekstu i wskazywała na równoważność różnych wersji tego samego utworu, które poszerzają pole interpretacji twórczości danego autora (por. Bem 2017).

niezadawalające, gdyż regułą jest brak możliwości ingerowania w tekst pozyskany w ten sposób oraz brak prawa do jego przetwarzania (zob. Pałka, Kwaśnicka-Janowicz 2017).<sup>54</sup>

W rezultacie naszych działań obszerne dwa tomy *Dziennika* (2011, 2014) oraz *Dziennik amerykański* zostały w wersji elektronicznej PDF udostępnione nam przez Wydawnictwo Literackie i Wydawnictwo Zeszyty Literackie<sup>55</sup>. Natomiast trudno było dotrzeć do wersji elektronicznej dziennika pt. *Zawsze powroty z czasów paryskich*, ponieważ jego wersji elektronicznej nie posiadało żadne wydawnictwo. Dlatego zdecydowaliśmy się na skanowanie materiałów, których nie można było pozyskać w postaci elektronicznej, co jednak wymusiło żmudne czynności doprowadzenia tekstu do postaci umożliwiającej przetwarzanie tekstu na poziomie znaków pisarskich.<sup>56</sup> Taki sposób pozyskiwania danych językowych nie jest niczym nowym i jak można zaobserwować w pracach badawczych, aprobowanym w środowisku. Na przykład Rybicki zgromadził korpus prozy polskiej i charakteryzując materiał przyznawał, że: „Duża część jednak musiała zostać przeniesiona w medium elektroniczne przez skanowanie i OCR” (Rybicki 2017: 12).<sup>57</sup> Działanie to zachęciło nas do tej formy zbierania całości interesującego pod względem badań materiału, składającego się finalnie na korpus prozy i korpus poezji, gdyż dawało możliwość wyboru z bogatej bibliografii podmiotowej.

### 3.4.2. Opracowanie danych początkowych

Jak zostało powiedziane, dwutomowy *Dziennik* oraz *Dziennik amerykański* zostały w wersji elektronicznej PDF udostępnione przez wydawnictwa. Pozostały materiał złożony z siedmiu książek został zeskanowany i zapisane w formacie PDF. Podobnie 22 tomy poetyckie zostały zeskanowane i zapisane w formacie PDF. Następnie dziesięć książek, napisanych prozą, zostały przekonwertowane do formatu MS Word przy pomocy programu typu OCR – ABBY. W ten sposób stworzony został folder *JH KORPUS*

---

<sup>54</sup> Należy zaznaczyć, że będący w powszechnym użyciu termin „wersja elektroniczna” jest rozumiany jako synonimiczny w stosunku do określenia „tekst na nośniku elektronicznym”.

<sup>55</sup> Materiał został przesłany przez wydawnictwa po wcześniejszym oficjalnym zapytaniu wystosowanym przeze mnie i promotora. Stosowna korespondencja została zarchiwizowana w pliku.

<sup>56</sup> Rozwiązanie to wiązało się ze znacznym kosztem czasowym w postaci blisko dwuletniego okresu gromadzenia materiału elektronicznego nadającego się do przeprowadzenia zaplanowanych zadań.

<sup>57</sup> por. też Vetulani G. (2000:116), która podsumowuje: „Korpusy można pozyskiwać w różny sposób, bądź wykorzystując istniejące już teksty (poprzez przepisywanie lub skanowanie), albo też już na etapie tworzenia tekstu, tj. podczas edycji, korekty czy druku. W ten sposób w wielu językach naturalnych (lecz przede wszystkim w języku angielskim), powstały olbrzymie elektroniczne korpusy tekstów, które są archiwizowane i które dają się przetwarzać automatycznie (zob. Habert i in. 1997, Piotrowski 1994).”

*PROZY*, który zawiera 10 podfolderów z wersją elektroniczną każdej książki z osobna na różnych etapach opracowania, to znaczy:

- wersja PDF,
- wersja rozczytana przez OCR z wszystkimi metadanymi, okładką i ewentualnymi zdjęciami i błędami w rozczytaniu tekstu przez program OCR – oznaczana v. 0

Te same kroki zastosowano przy przygotowywaniu tomów poetyckich, które zapisano w postaci 22 plików w folderze *JH KORPUS POEZJI I PROZY POETYCKIEJ*. Pierwsza faza była bardzo czasochłonna, niemniej pozwoliła na przygotowanie planu działań dotyczących kolejnych kroków związanych z normalizacją materiału.

### **3.4.3. Etapy normalizacji korpusu prozy i poezji Julii Hartwig**

Korpus w formie elektronicznej należało przygotować do przetwarzania i porównywania tekstów i plików między sobą na potrzeby badań relacjonowanych w niniejszej rozprawie. Są to w pierwszym rzędzie badania ilościowe. Proces ten nazywamy normalizacją. Na wejściu do tego procesu znajduje się pierwotnie zebrany korpus w postaci zdigitalizowanej (na poziomie znaków pisarskich).

W pierwszej fazie normalizacji zostały wskazane elementy nadmiarowe, jak na przykład powtórzenia utworów, o których była mowa powyżej, nieporządane w analizie ilościowej (patrz. Rozdział 4). Z tego powodu nie są one uwzględniane w kolejnych etapach normalizacji. (Elementy nadmiarowe są zachowane w pełnej wersji zebranego pierwotnie korpusu liczącego ponad 800000 słów tekstu, dzięki czemu będzie on mógł być używany w badaniach wykraczających poza cele tej rozprawy.) Do elementów nadmiarowych zaliczamy w szczególności takie elementy tekstowe, jak:

- zawartość strony tytułowej,
- nazwisko autorki,
- tytuł utworu,
- informacje redakcyjne,
- numery stron,
- elementy tekstowe zawarte w stopce i nagłówku,

- zdjęcia, ilustracje, tabele, etc. i ich podpisy,
- spis treści,
- bibliografię,
- indeksy,
- oraz wszystkie inne elementy, które zawierają metadane w stosunku do tekstu.

W ten sposób uzyskaliśmy wersję bez informacji metajęzykowych, którą w gromadzonych zasobach oznaczyliśmy v.1 (por. Jacobs 2018).

W drugiej fazie sprawdzono zgodność wersji oryginalnych i elektronicznych poszczególnych utworów składających się na oba korpusy – prozy i poezji. Operacja ta była konieczna, gdyż istotna część pierwotnego korpusu powstała z tekstów zeskanowanych i została doprowadzona do postaci pożądanej przez użycie narzędzi OCR i ręczną korektę. W trakcie tych prac poczynione zostały wstępne obserwacje typologii błędów zdarzających się podczas przetwarzania systemem OCR. Rezultatem drugiego etapu normalizacji była zatem tekstowa zawartość tomów pozbawiona wspomnianych powyżej elementów nadmiarowych, po usunięciu ewentualnych rozbieżności w stosunku do oryginałów wynikających z ewentualnych błędów ze strony systemu OCR. Zestawienie plików nr 1 prezentuje fragment korpusu prozy, to znaczy dane dotyczące zbioru reportaży i ilustruje sposób ich archiwizacji, który jest jednakowy dla wszystkich pozostałych utworów zawartych w zebranych materiale.

Nazwa	Data modyfikacji	Typ	Rozmiar
JH Z niedalekich podróży v.0.	16.04.2020 19:52	Dokument programu...	125 KB
JH Z niedalekich podróży v.1	02.05.2019 19:15	Dokument programu...	4 284 KB
JH Z niedalekich podróży v.1a	16.04.2020 19:48	Dokument programu...	606 KB
Z niedalekich podróży 1-19	23.01.2019 22:02	Foxit Reader PDF Do...	14 932 KB
Z niedalekich podróży 20-39	23.01.2019 22:37	Foxit Reader PDF Do...	14 605 KB
Z niedalekich podróży 40-59	23.01.2019 23:36	Foxit Reader PDF Do...	14 889 KB
Z niedalekich podróży 60-79	23.01.2019 23:55	Foxit Reader PDF Do...	17 726 KB
Z niedalekich podróży 80-99	24.01.2019 00:16	Foxit Reader PDF Do...	13 498 KB
Z niedalekich podróży 100-119	24.01.2019 00:30	Foxit Reader PDF Do...	14 722 KB
Z niedalekich podróży 120-139	24.01.2019 00:46	Foxit Reader PDF Do...	16 844 KB
Z niedalekich podróży 140-159	24.01.2019 01:05	Foxit Reader PDF Do...	14 656 KB

Zestawienie 1. Fragment korpusu prozy – przykład archiwizowania poszczególnych wersji korpusu prozy Julii Hartwig.

W trzeciej fazie usunięte zostały wyrazy dzielone na końcu linii oraz oczyszczone ze znaku „-”. Wszystkie tomy zostały przekonwertowane do dokumentów txt oraz zostały utworzone dwa pliki: *JH korpus właściwy 2020-05-11 proza* zawierających pozycje o numerach od 01 do 10 i *JH korpus właściwy 2020-05-11 poezja* złożony z zasobów oznaczonych 001-022.

Kolejnym, czwartym, krokiem było połączenie wszystkich plików 01-10 w jeden, który został nazwany *proza-all* i plików 001-022 w jeden, który został nazwany *poezja-all*.<sup>58</sup> Działanie to było motywowane wymaganiami poszczególnych narzędzi, aby ten sam materiał językowy mógł być analizowany bez względu na typ danych oraz przeszukiwalny w jednym pliku. Ponadto większość działań informatycznych na plikach odbywa się na takim formacie danych i na przykład serwisy wolnelektury.pl lub <https://pl.wikisource.org> umożliwiają pobieranie danych w tej wersji, gwarantując ich rzetelność (zob. Vetulani Z., Witkowska, Kubis 2019).

Poniższe Zestawienie 2 oraz Zestawienie 3 przedstawiają finalną wersję materiału językowego, który stanowił podstawowy tekstowy materiał źródłowy do dalszej pracy.

MARTA > 2021 doktorat > 02. KORPUSY > 2020-05-11 JH KORPUS WLASCIWY > 2020-05-11 proza				
Nazwa	Data modyfikacji	Typ	Rozmiar	
01 z-niedalekich-podrozy	19.05.2020 12:37	Dokument tekstowy	243 KB	
02 apollinaire	14.06.2020 00:18	Dokument tekstowy	907 KB	
03 gerard-de-nerval	14.06.2020 17:53	Dokument tekstowy	621 KB	
04 dziennik-amerykanski	18.07.2020 19:01	Dokument tekstowy	602 KB	
05 zawsze-powroty	14.06.2020 11:15	Dokument tekstowy	705 KB	
06 pisane-przy-oknie	19.05.2020 12:37	Dokument tekstowy	298 KB	
07 wybrancy-losu	19.05.2020 20:41	Dokument tekstowy	145 KB	
08 podziekowanie-za-goscine	19.05.2020 20:43	Dokument tekstowy	303 KB	
09 dziennik-1	14.06.2020 11:10	Dokument tekstowy	666 KB	
10 dziennik-2	19.05.2020 21:33	Dokument tekstowy	696 KB	
proza-all	13.06.2020 19:14	Dokument tekstowy	5 183 KB	

Zestawienie 2. Podsumowanie zawartości korpusu prozy Julii Hartwig.

<sup>58</sup> Ostatecznie pliki *proza-all* oraz *poezja-all* zostały włączone odpowiednio do folderów „JH korpus właściwy 2020-05-11 proza” (Zestawienie plików nr 2) i JH korpus właściwy 2020-05-11 poezja (Zestawienie plików nr 3). Korpus prozy i korpus poezji złożone z zestawów plików txt oznaczonych numerami 01-10 i 001-022 są tożsame pod względem zawartości z jednolitymi dokumentami *proza-all* i *poezja-all*. Zatem przyjęte zostały jako jednakowe materiały językowe i posługiwano się nimi zamiennie w zależności od wymagań narzędzi używanych do ich przetwarzania.

› MARTA › 2021 doktorat › 02. KORPUSY › 2020-05-11 JH KORPUS WLASCIWY › 2020-05-11 poezja

Nazwa	Data modyfikacji	Typ	Rozmiar
001 pozegnania	13.06.2020 18:37	Dokument tekstowy	61 KB
002 wolne rece	25.05.2020 12:07	Dokument tekstowy	45 KB
003 dwoistosc	25.05.2020 12:08	Dokument tekstowy	38 KB
004 czuwanie	25.05.2020 12:10	Dokument tekstowy	42 KB
005 chwila postoju	25.05.2020 12:14	Dokument tekstowy	37 KB
006 obcowanie	25.05.2020 12:16	Dokument tekstowy	94 KB
007 czulosc	23.05.2020 18:27	Dokument tekstowy	86 KB
008 nim opatrzysz sie zielen	20.05.2020 17:46	Dokument tekstowy	8 KB
009 zobaczone	23.05.2020 18:31	Dokument tekstowy	65 KB
010 nie ma odpowiedzi	25.05.2020 12:10	Dokument tekstowy	49 KB
011 blyski	25.05.2020 12:10	Dokument tekstowy	69 KB
012 wiersze amerykanske	25.05.2020 12:10	Dokument tekstowy	59 KB
013 mowiac nie tylko od siebie	10.01.2021 22:56	Dokument tekstowy	75 KB
014 bez pozegnania	25.05.2020 12:10	Dokument tekstowy	49 KB
015 zwierzenia i blyski	25.05.2020 12:10	Dokument tekstowy	86 KB
016 to wroci	25.05.2020 12:10	Dokument tekstowy	20 KB
017 trzecie blyski	25.05.2020 12:10	Dokument tekstowy	99 KB
018 jasne niejasne	25.05.2020 14:47	Dokument tekstowy	23 KB
019 gorzkie zale	25.05.2020 12:10	Dokument tekstowy	26 KB
020 zapisane	20.05.2020 22:20	Dokument tekstowy	16 KB
021 blyski zebrane	24.05.2020 14:59	Dokument tekstowy	4 KB
022 spojrzzenie	06.05.2020 18:07	Dokument tekstowy	12 KB

Zestawienie 3. Podsumowanie zawartości korpusu poezji Julii Hartwig.

Zamieszczone powyżej Zestawienie 3 stanowi ilustrację wielkości korpusu w postaci znormalizowanej, przygotowanej do bezpośredniego stosowania.

Korpus	Znaki bez końca linii	Słowa
proza-all	4 614 710	<b>669 163</b>
poezja-all	985 681	<b>151 794</b>

Tabela 1. Zestawienie ilościowe dla korpusu prozy i poezji Julii Hartwig zebranych w całości w plikach dwóch plikach nazwanych proza-all i poezja-all.<sup>59</sup>

<sup>59</sup> Zestawienie sporządzone dla wersji źródłowej korpusu zawierającej m. in. nazwiska, imiona własne, nazwy geograficzne.

#### **3.4.4. Złożoność normalizacji i jakość korpusu**

Powyższy opis procesu normalizacji korpusów ilustruje złożoność procesu tworzenia dobrego korpusu cyfrowego, to znaczy nie tylko reprezentatywnego pod względem merytorycznym, lecz również użytecznego ze względu na wymagania techniczne późniejszego przetwarzania automatycznego. „Wykorzystywanie podczas badań językoznawczych zinformowanych korpusów, które poddaje się zautomatyzowanej eksploracji, znacznie przyspiesza proces wyszukiwania niezbędnych danych językowych oraz skraca czas ich przetwarzania do pożądanej postaci, np. słowników” (Vetulani G. 2012: 51).

Reasumując, można stwierdzić, że etap przygotowawczy do przeprowadzenia właściwych analiz okazał się czasochłonny, niemniej przy stosunkowo małym korpusie możliwy do samodzielnej realizacji. Ponadto normalizacja przyniosła korzyści, umożliwiając przyspieszenie prac w kolejnych etapach przetwarzania danych.

### **3.5. Rozwiązane problemy związane z pozyskiwaniem korpusu w formie elektronicznej**

Omawiając poszczególne etapy zbierania i normalizacji prezentowanego korpusu, wskazaliśmy pokrótce na różnego typu trudności związane z wybraną metodą badań – to jest gromadzenie korpusu elektronicznego. W tym miejscu zbieramy nasze spostrzeżenia dotyczące problemów metodologicznych oraz technicznych oraz wskazujemy na podjęte decyzje.

#### **3.5.1. Pozyskanie elektronicznego korpusu twórczości Julii Hartwig**

Korzystanie z dorobku pisarzy, poetów regulowane jest zapisami prawa autorskiego. Prawo zabrania powielania dzieł w celu udostępniania ich w domenie publicznej, natomiast praca naukowa dotycząca twórczości danego artysty nie podlega tego typu ograniczeniom. Dlatego zebrany i opracowany przez nas korpus służy wyłącznie do celów badań naukowych prezentowanych w niniejszej rozprawie<sup>60</sup>.

---

<sup>60</sup> O kwestii udostępniania korpusów czytamy również we wstępie do NKJP (Przepiórkowski (red.) 2012:54): „Z powodu praw autorskich nie jest możliwe – nawet niekomercyjne – udostępnianie pełnego NKJP w postaci tekstowej. Inaczej ma się sprawa w wypadku podkorpusu milionowego, składającego się z krótkich fragmentów większych tekstów, w dodatku użytych jako cytaty, będące podstawą znacznie istotniejszych od samych tekstów informacji lingwistycznych.”

### 3.5.2. Skanowanie materiału książkowego

Jedną z większych trudności, z jaką przyszło się zmierzyć, okazała się czasochłonność skanowania i opracowania informatycznego 22 tomów poetyckich i 7 książek pisanych prozą. Poniższe tabele zawierają zestawienie wszystkich 29 pozycji składających się na korpus:

	Tytuł	Rok pierwszego wydania	Wydawnictwo	Ilość wyrazów	Znaki bez końca linii	Gatunek
1	<i>Z niedalekich podróży</i>	1954	Ludowa Spółdzielnia Wydawnicza	33 999	231 598	reportaż
2	<i>Apollinaire</i>	1962	PIW	126 093	868 058	monografia lit.
3	<i>Gerard de Nerval</i>	1972	PIW	86 794	593 926	monografia lit.
4	<i>Zawsze powroty. Z dzienników podróży</i>	2001	Sic!	98 184	676 569	dziennik
5	<i>Pisane przy oknie</i>	2004	Biblioteka „Więzi”	40 949	284 402	felieton
6	<i>Wybrańcy losu</i>	2006	Sic	20 132	137 779	felieton
7	<i>Podziękowanie za gościnę. Moja Francja</i>	2006	słowo/obraz/terytoria	40 993	289 350	esej

Tabela 2. Zestawienie dla korpusu prozy – przedstawione dane liczbowe dotyczą wersji korpusu po normalizacji.

	<b>Tytuł</b>	<b>Rok wydania</b>	<b>Wydawnictwo</b>	<b>Ilość wyrazów</b>	<b>Znaki bez końca linii</b>	<b>Gatunek</b>
1	<i>Pożegnania</i>	1956	Czytelnik	8 635	54 468	wiersz
2	<i>Wolne ręce</i>	1969	PIW	6 704	43 602	wiersz, proza poetycka
3	<i>Dwoistość</i>	1971	Czytelnik	5 811	37 457	wiersz, proza poetycka
4	<i>Czuwanie</i>	1978	Czytelnik	6 328	41 291	wiersz, proza poetycka
5	<i>Chwila postoj</i>	1980	Wydawnictwo Literackie	5 567	36 601	proza poetycka
6	<i>Obcowanie</i>	1987	Czytelnik	13 503	85 765	wiersz, proza poetycka, błyski
7	<i>Czułość</i>	1992	Znak	13 019	84 128	wiersz, proza poetycka, błyski
8	<i>Nowe wiersze [w:] Nim opatrzy się zieleń</i>	1995	Znak	1 159	7 430	wiersz
9	<i>Zobaczone</i>	1999	Wydawnictwo a5	9 142	58 619	wiersz, błyski
10	<i>Nie ma odpowiedzi</i>	2001	Sic!	6 908	44 019	wiersz
11	<i>Błyski</i>	2002	Sic!	9 436	63 828	błyski
12	<i>Wiersze amerykańskie</i>	2002	Sic!	8 268	53 113	wiersz
13	<i>Mówiąc nie tylko do siebie. Poematy prozą</i>	2003	Sic!	10 675	70 093	proza poetycka
14	<i>Bez pożegnania</i>	2004	Sic!	6 806	43 345	wiersz
15	<i>Zwierzenia i błyski</i>	2004	Sic!	11 524	79 037	błyski
16	<i>To wróci</i>	2007	Sic!	2 722	17 372	wiersz
17	<i>Trzecie błyski</i>	2008	Sic!	13 448	92 552	błyski
18	<i>Jasne niejasne</i>	2009	Wydawnictwo a5	3 246	20 168	wiersz
19	<i>Gorzkie żale</i>	2011	Wydawnictwo a5	3 613	22 953	wiersz
20	<i>Zapisane</i>	2013	Wydawnictwo a5	2 301	14 172	wiersz
21	<i>Nowe błyski [w:] Błyski zebrane</i>	2014	Zeszyty Literackie	511	3 561	wiersz
22	<i>Spojrzenie</i>	2016	Wydawnictwo a5	1 601	9 780	wiersz

Tabela 3. Zestawienie metadanych dla korpusu poezji – przedstawione dane liczbowe dotyczą wersji korpusu po normalizacji.

Już samo skanowanie jest czynnością czasochłonną i żmudną. Z naszych obserwacji wynika, że zeskanowanie w sposób ciągły 200 stronicowej książki (przy założeniu ciągłości pracy i braku zakłóceń wyniesie ok 4 godz. 20 min. Do tego należy dodać czas przygotowania stanowiska pracy oraz materiału do procesu skanowania, a także ewentualne błędy systemu skanowania powstałe w trakcie tego procesu i wynikającą z nich konieczność powtarzania niektórych czynności. Zdarzało się pominąć też pojedyncze strony podczas skanowania, ponieważ niektóre z nich uległy zniekształceniu w czasie obróbki informatycznej i należało je poprawić (skan i dołożenie do pliku docelowego) w ostatecznej wersji elektronicznej danej książki. Tego typu przypadki miały miejsce podczas procesu sprawdzania zgodności oryginału z wersją v.1 (zaraz po rozczytaniu za pomocą OCR).

### **3.5.3. Sprawdzanie zgodności oryginału z wersją OCR**

Drugim z najbardziej czasochłonnych procesów było sprawdzanie zgodności oryginału książkowego z wersją v.1. Obecnie narzędzia typu OCR bardzo dobrze radzą sobie z odczytywaniem tekstu drukowanego. Kłopoty pojawiają się przede wszystkim w miejscach niedokładnie zeskanowanych (prześwietlenia, krzywo ułożona kartka, niedokładnie przytrzymany grzbiet książki) oraz znacznie rzadziej, z powodu kroju czcionki czy też wystąpienia wyrazów zapożyczonych bądź neologizmów. Dlatego, mimo początkowo zadowalających rezultatów, okazało się, że po pierwszych próbach przetwarzania, na przykład z wykorzystaniem lematyzatora, na listach wyjściowych pojawiały się słowa ze znakami nierozpoznanymi. Fakt ten spowodował konieczność dokładnego sprawdzenia poprawności tekstu i zgodności pomiędzy wersją elektroniczną każdej pozycji korpusu a wersją książkową.

## ROZDZIAŁ 4

### **Wykorzystanie metod automatycznego przetwarzania tekstu do analizy przymiotnikowej strukturalizacji predykatu w korpusie prozy i poezji Julii Hartwig**

Analizę warsztatu pisarskiego Julii Hartwig metodami lingwistyki komputerowej przeprowadzimy w oparciu o całokształt dzieła Autorki reprezentowany przez zebrany w tym celu referencyjny elektroniczny korpus opublikowanej prozy i poezji z okresu od 1954 do 2016.

Odwołanie się do korpusu stanowi odpowiednik metody empirycznej stosowanej w naukach przyrodniczych opartej na danych pochodzących z obserwacji świata (w tym danych uzyskanych w drodze eksperymentu<sup>61</sup>). Lista dzieł i zasady ich wyboru zostały opisane w rozdziale poprzednim. W tym rozdziale przedstawimy kolejne kroki prowadzące do otrzymania oznakowanego korpusu, pokażemy etapu wyboru, opisu i klasyfikacji charakterystycznych dla Julii Hartwig przymiotników predykatywnych, jako że uważamy je za jeden z ważnych wykładników warsztatu artystycznego literackiego artystki.

#### **4.1. Wykorzystane narzędzi przetwarzania języka naturalnego**

Od dłuższego czasu, dla wielu języków tworzone są odpowiednio zorganizowane bazy językowe z przeznaczeniem do badań naukowych i zadań praktycznych (zob. Rozdział 3). Samo gromadzenie danych korpusowych w pamięci komputera nie stwarza jeszcze wystarczających warunków do prowadzenia badań. Muszą być one w sposób szczególnie zorganizowane z wykorzystaniem *znakowania lingwistycznego* (Przepiórkowski, 2004: 5) wprowadzającego metadane o charakterze ogólnym bądź szczegółowym w przypadkach stosowania korpusu w konkretnym projekcie badawczym lub w dobrze określonej aplikacji. Dlatego też elementy korpusu elektronicznego podlegają anotowaniu (indeksowaniu, tagowaniu) za pomocą właściwych znaczników,

---

<sup>61</sup> W przypadku dysponowania całokształtem dzieła jednego autora pozyskiwanie danych w drodze eksperymentu z oczywistych powodów nie ma zastosowania.

bez których ich dalsze przetwarzanie byłoby mniej efektywne. Typowe znaczniki dotyczą morfo-składni, ale mogą też odnosić się do semantyki, pragmatyki, prozodii itp. Anotowanie zwykle odnosi się do określonej teorii lingwistycznej i/lub formalizacji. Formalizacja w opisie języka naturalnego służy precyzji tego opisu oraz jest pochodną przyjętych przez jego twórców założeń teoretycznych. Z tego powodu za szczególnie ważne uznaliśmy stosowanie w omawianych badaniach narzędzi o dobrze określonych i rozpoznanych podstawach metodologicznych, a w szczególności narzędzi uzyskanych w Zakładzie Lingwistyki Informatycznej i Sztucznej Inteligencji UAM (zob. 4.1.1. i 4.1.2).

W przypadku analizy korpusu prozy i poezji Julii Hartwig zostały zastosowane narzędzia informatyczne do:

- a) oznaczenia przymiotników i pozyskania odpowiednich list referencyjnych,
- b) obliczeń ilościowych, w tym pozyskania list frekwencyjnych form przymiotników, a także list frekwencyjnych dla leksemów,
- c) ekscerpacji kontekstów użycia przymiotników w formie konkordancji,
- d) obliczeń przyrostowych rozwoju zasobu leksykalnego.

Pozyskane wyniki umożliwiły obserwację w zgromadzonym zasobie leksykalnym sposobu dysponowania przez Autorkę kategorią przymiotnika i tym samym pozwoliły wskazać charakterystyczne cechy jej warsztatu pisarskiego. Realizacja wymienionych celów została osiągnięta dzięki zastosowaniu czterech aplikacji (programów) informatycznych: Lematyzatora POLEX, biblioteki NLTK programu Python, Korpusomatu, systemu Text SubCorpora Creator (TSCC1.3.).

#### **4.1.1. Formalizm i lematyzator POLEX**

##### **Elektroniczny słownik morfologiczny języka polskiego – formalizm POLEX<sup>62</sup>**

Słownik POLEX stanowi cyfrową bazę danych obejmującą zasadniczy zasób leksykalny polszczyzny. W nawiązaniu do dorobku klasyków polskiej leksykografii i autorów kluczowych słowników (Doroszewski, Polański, Tokarski, Szymczak, Saloni, Mędak

---

<sup>62</sup> „POLEX – polska leksykalna baza danych” jest nazwą projektu realizowanego w oparciu o grant Komitetu Badań Naukowych w latach 1994 – 1996 pod kierunkiem Zygmunta Vetulaniego.

i inni) opracowane zostało nowe podejście do opisu morfologii polskiej, którego cechą charakterystyczną jest to, że zaproponowany zestaw fleksyjnych klas paradygmatycznych jest jednoznaczny, w tym sensie, że jednemu kodowi (alfa)numerycznemu odpowiada dokładnie jeden (uporządkowany) zbiór końcówek. Z drugiej strony stosowany tu kod jest pozycyjny, tj. pozycje są znaczące i lingwistycznie umotywowane. Cechy te dają możliwość łatwego rozbudowywania i weryfikowania słownika przez eksperta bez konieczności stosowania specjalnych narzędzi (edytora). Jednoznaczność kodu gwarantuje precyzję opisu do tej pory nieosiągniętą w istniejących słownikach (zob. *Jednoznaczne kodowanie fleksji rzeczownika polskiego i jego zastosowanie w słownikach elektronicznych – format POLEX*; Vetulani Z., Walczak, Obrębski, Vetulani G. 1998).

### **Lematyzator POLEX**

Termin *lematyzacja* będziemy stosowali w odniesieniu do procedury znajdowania formy podstawowej (słownikowej) w leksemie, tzn. w zbiorze form fleksyjnych jakie może przyjmować wyraz nie tracąc swojego znaczenia. Za (*Linguistic Problems in the Theory of Man-Machine Communication in Natural Language*; Vetulani Z. 1989: 71) zdefiniujemy pojęcie lematyzacji jako “operacji określonej na korpusie, która polega na zastąpieniu wyrazu tekstowego przez leksem”.

Lematyzator POLEX<sup>63</sup> jest uaktualnioną, na potrzeby przedstawianych w rozprawie badań, wersją lematyzatora znanego pod nazwą LEXAN (zob. *Lexical resources and tools for tagging Polish texts within GRAMLEX*; Vetulani Z., Martinek, Obrębski 1997). Lematyzator wykorzystuje elektroniczny słownik morfologiczny POLEX i jest podstawowym narzędziem stosowanym przez nas do uzyskania opisu morfologicznego dla przymiotnikowych form wyrazowych występujących w analizowanym korpusie Julii Hartwig. W efekcie lematyzacji tekst korpusu został przekształcony na ciąg leksemów odpowiadających poszczególnym wyrazom.

---

<sup>63</sup> Kod źródłowy tego narzędzia został napisany w języku programowania PROLOG (*PROgrammation en LOGique*). Jest to język programowania logicznego ogólnego zastosowania, stworzony pierwotnie do analizy języków naturalnych przez francuskiego badacza Alaina Colemeraure’a (Colemerauer 1972, Vetulani Z. 2004)

#### 4.1.2. Biblioteka NLTK

Do analiz frekwencyjnych wykorzystujemy język programowania Python 3.4 współpracujący z biblioteką Natural Language Toolkit (NLTK; Bird, Klein, Loper 2009). Biblioteka ta stanowi zestaw gotowych podprogramów i narzędzi, które można wykorzystać z poziomu kodu źródłowego i dostosować do swoich potrzeb. Dzięki bibliotece NLTK można na przykład określić wielkości korpusu, policzyć frekwencję wyrazów, znormalizować tekst itp.

#### 4.1.3. Korpusomat<sup>64</sup>

W analizie wykorzystana została również aplikacja do tworzenia i przeszukiwania korpusów Korpusomat (Kieraś, Kobyliński, Ogrodniczuk 2018). Jest to zestaw narzędzi, które powstawały w Zespole Inżynierii Lingwistycznej Instytutu Podstaw Informatyki PAN i nadal są rozwijane oraz zostały włączone do polskich zasobów infrastruktury dla humanistów w ramach projektu Clarin-pl Korpusomat jest z założenia narzędziem przeznaczonym dla użytkowników, którzy nie muszą posiadać specjalistycznej wiedzy ani umiejętności związanych z jego budową. Ma on rozmaite funkcjonalności, które są systematycznie aktualizowane. Program umożliwia wczytanie własnego korpusu oraz wprowadzenie metadanych<sup>65</sup>. Korpusomat służy między innymi do generowania danych usystematyzowanych ze względu na kontekst występowania danej jednostki językowej, wynik możliwy jest do pobrania w arkuszach kalkulacyjnych (format \*csv). Pozyskanie danych z przetworzonego korpusu prozy i poezji Hartwig ze względu na konteksty występowania przymiotnika w badanym materiale było kluczowe dla prezentowanych w niniejszej pracy analiz i wniosków. Ponadto dało możliwość porównywania korpusów, generowania statystyk, pozyskiwania kolokacji itd.

#### 4.1.4. Text SubCorpora Creator 1.3. (TSCC)

Text SubCorpora Creator (TSCC) to program wspomagający ekstrakcję podkorpusów z większych zbiorów danych tekstowych, które dzielone są na części o określonej długości. Aplikacja przetwarza korpus przy użyciu *listy referencyjnej* zawierającej listę

---

<sup>64</sup> Wykorzystany przez nas pakiet narzędzi jest dostępny na stronie <https://korpusomat.pl/> (dostęp 15.07.2021).

<sup>65</sup> Szczegółowe informacje dotyczące języka zapytań zawiera *Instrukcja korzystania z wyszukiwarki korpusowej w Korpusomacie* (z dnia 26 sierpnia 2020) dostępna na stronie [www.korpusomat.pl](http://www.korpusomat.pl).

słów np. listę danej kategorii gramatycznej czy też listę tematyczną. Program przeprowadza analizę nasycenia korpusu ze względu na jakieś zjawisko. W naszym przypadku zjawiskiem tym będzie pojawianie się w korpusie form wyrazowych odpowiadających elementom *listy referencyjnej*<sup>66</sup> przymiotników. Analiza nasycenia odbywa się poprzez obserwowanie pojawiania się nowych form w kolejnych próbkach korpusu. Dane wyjściowe zawierają zestawienie ilościowe dotyczące nasycenia, na podstawie którego można sporządzić reprezentację graficzną stanowiącą wykres *nasycenia leksykalnego*<sup>67</sup> (Vetulani Z. 1989, Vetulani Z., Witkowska, Menken, Canbolat 2018). Wersja aplikacji 1.3. została przygotowana dla języka polskiego w oparciu o słownik POLEX (Vetulani Z., Witkowska, Canbolat 2018).

## **4.2. Zastosowanie narzędzi informatycznych w procesie przetwarzania korpusu prozy i poezji Julii Hartwig**

Celem stosowania narzędzi informatycznych omówionych w tej rozprawie jest stworzenie podstaw do wyciągnięcia empirycznie uzasadnionych wniosków odnośnie charakterystycznych cech warsztatu literackiego artystki.

Ze względu na wagę, jaką Julia Hartwig przypisuje przymiotnikowi i sygnalizuje to w swoich wypowiedziach (por. wstęp), na obecnym etapie badań uważamy za właściwe poświęcenie szczególnej uwagi tej właśnie kategorii leksykalnej. W pierwszym rzędzie interesują nas użycia słownictwa (aspekty ilościowe) i dynamika rozwoju warsztatu twórczego.

Podstawowe znaczenie dla prowadzonych badań mają listy frekwencyjne służące do wyciągnięcia wniosków dotyczących bogactwa warsztatu leksykalnego i preferencji leksykalnych pisarki, odzwierciedlonych w tych listach.

Tym nie mniej, ważne jest zbadanie dynamiki rozwoju tego warsztatu przy wykorzystaniu analizy przyrostowej przeprowadzonej z użyciem wymienionych w rozprawie narzędzi-(patrz 4.2.5).

---

<sup>66</sup> Por. definicja pojęcia *lista referencyjna* w sekcji 4.2.1.

<sup>67</sup>Pojęcie nasycenia leksykalnego (ang. *lexical saturation*) jest to zjawisko zatrzymania przyrostu obserwowanych nowych form językowych (np. leksemów) (lub jego spowolnienia) w badanym korpusie. Innymi słowy korpus jest nasycony leksykalnie, gdy nowe leksemy pojawiają się sporadycznie w kolejnych fragmentach korpusu. Wówczas krzywa reprezentująca graficznie stosunek przyrosty wystąpień nowych form w kolejnych fragmentach ulega wypłaszczeniu.

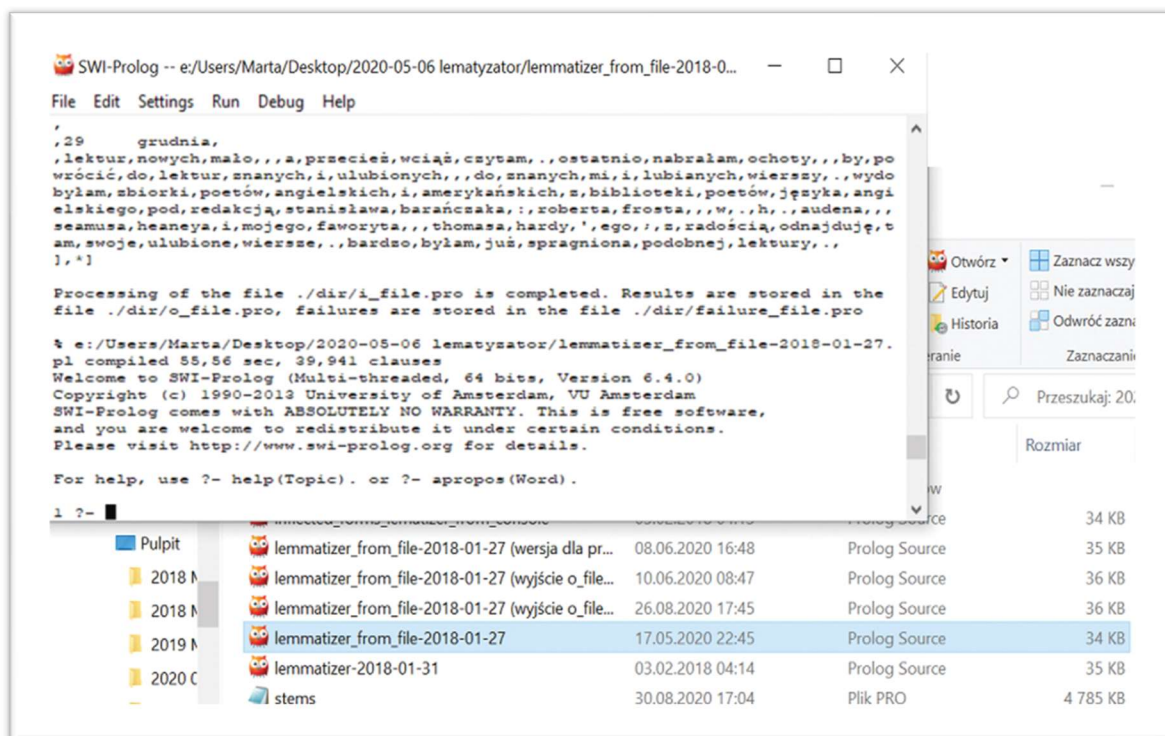
#### **4.2.1. Oznaczenie (lematyzacja) przymiotników w korpusach prozy i poezji Julii Hartwig oraz pozyskanie list referencyjnych przymiotników**

Przypomnijmy, że *korpus* rozumiany jest w tej pracy jako zbiór tekstów stanowiący materiał empiryczny do badań lingwistycznych, natomiast *lista referencyjna* to lista słów utworzona ze względu na określony temat lub cechę języka stanowiąca punkt odniesienia dla prowadzonych badań. Dodajmy, że lista przymiotników stanowi w niniejszej pracy podstawowy punkt odniesienia dla wszystkich przedstawionych badań, dlatego uważamy ją za listę referencyjną.

##### **4.2.1.1. Lematyzacja**

Do oznaczenia przymiotników w korpusie posłużył lematyzator stanowiący wersję uaktualnioną w stosunku do macierzystego lematyzatora POLEX opisanego w sekcji 4.1.1., zaimplementowanego w SWI Prologu v. 6.4.0 (patrz Przykład 1, poniżej). Program po modyfikacji został wykorzystany do wyszukania i otagowania form przymiotnikowych z pominięciem wszystkich pozostałych kategorii gramatycznych. W szczególności tagi zawierają formy podstawowe (*lematy*) oznaczonych wyrazów. Zmodyfikowane narzędzie – *lemmatizer\_from\_file-2018-01-27* – rozpoznaje formy przymiotnikowe zgodnie z formalizmem zawartym w macierzystym lematyzatorze opartym na słowniku morfologicznym języka polskiego POLEX (Vetulani Z., Walczak, Obrębski, Vetulani G., 1998), który z kolei został opracowany na podstawie *Słownika języka polskiego* pod redakcją M. Szymczaka (1978-81).

Na podstawie dokonanej lematyzacji korpusu *poezja-all* została wygenerowana lista form przymiotnikowych w korpusie poezji Julii Hartwig jako lista referencyjna dla dalszych badań zgromadzonego korpusu. Podobnie, w wyniku lematyzacji pliku *proza-all*, została wygenerowana lista form przymiotnikowych zawartych w korpusie prozy Julii Hartwig.



Przykład 1. Podsumowanie procesu lematyzacji korpusu proza-all.

Pozyskane w procesie lematyzacji listy zawierają wykaz wszystkich form przymiotnikowych (*tokenów*) użytych w korpusie *poezja-all* oraz korpusie *proza-all*. Użyte tagi opisują każde wystąpienie form fleksyjnych przymiotnika oraz inne formy oznaczone jako nierozpoznane z powodu przynależności do innej kategorii gramatycznej.

Poniżej prezentujemy próbki:

- wyników lematyzacji z opisem morfologicznym dla wszystkich form występujących w korpusie (lista nr 1 i lista nr 2) oraz
- próbki wyników uzyskanych z lematyzatora z *nakładką filtrującą* (lista nr 3 i lista nr 4)

```

*2020.09.20o-file-all-poezja — Notatnik
Plik Edycja Format Widok Pomoc
znak niewyświetlany={koniec linii}
wiara={ wiara(N411, 1) }
znak niewyświetlany={koniec linii}
gruby={ gruby(ADJ1, 1) gruby(ADJ1, 9) gruby(ADJ1, 16) }
flamandzki={ flamandzki(ADJ3, 1) flamandzki(ADJ3, 9) flamandzki(ADJ3, 16) }
chłopak={ chłopak(N100, 1) chłopak(N102, 1) }
i={ i(CONJ, 1) i(EXCL, 1) }
bladooka={bladooki(ADJ3, 2) bladooki(ADJ3, 17) }
dziewczyna={ dziewczyna(N411, 1) }
,= {przecinek}
znak niewyświetlany={koniec linii}
która={ który(ADJPRO4, 2) który(ADJPRO4, 17) }
trzyma={ trzymać(V001, 4) }
rękę={ ręka(N416, 4) ręka(N417, 4) ręka(N418, 4) }
z={ z(P, 1) }
gałązką={ gałązka(N410, 5) }
na={ na(P, 1) }
wydęty={ wydęty(ADJPAP, 12) wydęty(ADJPAP, 14) wydęty(ADJPAP, 22) }
fartuchu={ fartuch(N303, 6) fartuch(N303, 7) }
,= {przecinek}
znak niewyświetlany={koniec linii}
czy={ czy(CONJ, 1) czy(PART, 1) }
warci={ warty(ADJ1, 19) warty(ADJ1, 27) }
są={ być(BYC, 7) }
pokus={ pokusa(N411, 9) }
świętego={ święty(ADJ1, 4) święty(ADJ1, 8) święty(N180, 2) święty(N180, 4) }
antoniego={ }

```

Tabela 4. Lista nr 1 Lematyzacja korpusu poezji – opis gramatyczny (morfologiczny) wszystkich jednostek w korpusie

```

o_file-proza-all — Notatnik
Plik Edycja Format Widok Pomoc
podróż={ podróż(N422, 1) podróż(N422, 4) podróż(N423, 1) podróż(N423, 4) }
dookoła={ dookoła(ADV0, 1) dookoła(P, 1) }
malowanego={ malowany(ADJPAP, 4) malowany(ADJPAP, 8) }
pieca={ piec(N320, 2) }
znak niewyświetlany={koniec linii}
najpierw={ najpierw(ADV0, 1) }
napisałam={ napisać(V112, 9) }
do={ do(P, 1) }
felicii={ }
curyłowej={ }
list={ list(N310, 1) list(N310, 4) lista(N411, 9) }
.= {kropka}
chciałam={ chcieć(V021, 9) }
jej={ jej(EXCL, 1) ona(PrsPRO4, 2) ona(PrsPRO4, 3) jej(PosPRO4, 1) jej(PosPRO4, 2) jej(PosPRO4, 3) jej(PosPRO4, 4) jej(PosPRO4, 5) }
pogratulować={ pogratulować(V104, 1) }
szczęśliwego={ szczęśliwy(ADJ1, 4) szczęśliwy(ADJ1, 8) }
wystąpienia={ wystąpienie(N520, 2) wystąpienie(N520, 8) wystąpienie(N520, 11) wystąpienie(N520, 14) wystąpienie(NV, 2) }
na={ na(P, 1) }
zjeździe={ zjazd(N310, 6) zjazd(N310, 7) }
rady={ rad(N311, 8) rad(N311, 11) rad(N311, 14) rada(N411, 2) rada(N411, 8) rada(N411, 11) rada(N411, 14) }
artystycznej={ artystyczny(ADJ1, 5) artystyczny(ADJ1, 7) artystyczny(ADJ1, 15) }
w={ w(P, 1) }
szczecinie={ szczecina(N411, 3) szczecina(N411, 6) }
,= {przecinek}
dzięki={ dzięki(P, 1) }
któremu={ który(ADJPRO4, 6) }

```

Tabela 5. Lista nr 2 Lematyzacja korpusu prozy – opis gramatyczny (morfologiczny) wszystkich jednostek w korpusie

```

2020-08-26 o_file_poezja_all_ok — Notatnik
Plik Edycja Format Widok Pomoc
znak niewyświetlany={koniec linii}
wiara={ }
znak niewyświetlany={koniec linii}
gruby={ gruby(ADJ1, 1) gruby(ADJ1, 9) gruby(ADJ1, 16) }
flamandzki={ flamandzki(ADJ3, 1) flamandzki(ADJ3, 9) flamandzki(ADJ3, 16) }
chłopak={ }
i={ }
bladooka={ bladooki(ADJ3, 2) bladooki(ADJ3, 17) }
dziewczyna={ }
,={ }
znak niewyświetlany={koniec linii}
która={ który(ADJPRO4, 2) który(ADJPRO4, 17) }
trzyma={ }
rękę={ }
z={ }
gałązką={ }
na={ }
wydęty={ wydęty(ADJPAP, 12) wydęty(ADJPAP, 14) wydęty(ADJPAP, 22) }
fartuchu={ }
,={ }
znak niewyświetlany={koniec linii}
czy={ }
warci={ warty(ADJ1, 19) warty(ADJ1, 27) }
są={ }
pokus={ }
świętego={ święty(ADJ1, 4) święty(ADJ1, 8) }
antoniego={ }

```

Tabela 6. Lista nr 3 Lematyzacja korpusu poezji – opis gramatyczny (morfologiczny) przymiotników

```

2020-08-30 o_file_proza_all — Notatnik
Plik Edycja Format Widok Pomoc
podróż={ }
dookoła={ }
malowanego={ malowany(ADJPAP, 4) malowany(ADJPAP, 8) }
pieca={ }
znak niewyświetlany={koniec linii}
najpierw={ }
napisałam={ }
do={ }
felicji={ }
curyłowej={ }
list={ }
,={ }
chciałam={ }
jej={ }
pogratulować={ }
szczęśliwego={ szczęśliwy(ADJ1, 4) szczęśliwy(ADJ1, 8) }
wystąpienia={ }
na={ }
zjeździe={ }
rady={ }
artystycznej={ artystyczny(ADJ1, 5) artystyczny(ADJ1, 7) artystyczny(ADJ1, 15) }
w={ }
szczecinie={ }
,={ }
dzięki={ }
któremu={ który(ADJPRO4, 6) }
jej={ }

```

Tabela 7. Lista nr 4 Lematyzacja korpusu prozy – opis gramatyczny (morfologiczny) przymiotników

Lematyzacja korpusu jest pierwszym i podstawowym krokiem w przetwarzaniu materiału językowego na potrzeby prowadzonych przez nas badań. Pozyskanie informacji gramatycznych na temat badanego materiału pozwoliło na dalsze modyfikacje zasobu językowego, jak na przykład:

- grupowanie wyrazów w listy ze względu na rozpoznane kategorie gramatyczne (przetwarzanie pliku *o-file* w arkuszu *csv*),
- analizę form nierozpoznanych (przetwarzanie pliku *failure-file* w arkuszu *csv*), wśród których możemy znaleźć m. in. Neologizmy i zapożyczenia.

#### 4.2.1.2 Tworzenie list referencyjnych

Listy z opisem poszczególnych jednostek (przyporządkowane kategorie gramatyczne w nawiasach, patrz listy przykładów w 4.2.1.1) po przeniesieniu do arkusza Excel stanowiły podstawę sortowania wyników. W ten sposób oczyszczone z wyrazów nierozpoznanych zostały pliki wyjściowe etapu lematyzacji celem wyekscerpowania przymiotników stanowiących listę referencyjną. Obie prezentowane powyżej (we fragmentach) listy (nr 3 i 4) zostały poddane następującym operacjom<sup>68</sup>:

- usunięto wyrażenia alfanumeryczne i specjalne (daty, wyrazy obcego pochodzenia zapisywane z apostrofem),
- usunięto znaki końca linii oraz formy nierozpoznane, tj. wszystkie formy, które nie były przymiotnikami,
- przesortowano listę według kolumny zawierającej opis gramatyczny,
- usunięto inne formy przymiotnikowe, jak imiesłowy (ADJPAP) czy zaimki przymiotne (ADJPRO),
- wydzielono listę zawierającą tylko przymiotniki wraz z powtórzeniami,
- usunięto powtórzenia form przymiotnika,
- przesortowano listy ze względu na formę podstawową przymiotnika.

Dalsze analizy korpusu prozy i poezji prowadzone były z wykorzystaniem list, odpowiednio dla prozy i poezji, zawierających formy przymiotników bez powtórzeń,

---

<sup>68</sup> Po przeniesieniu do arkusza Excel.

otrzymane na wyjściu z etapu v6 i te dwie listy stanowiły dla nas podstawowe listy referencyjne:

- na podstawie lematyzacji korpusu prozy przygotowano listę referencyjną form przymiotnikowych: v.6\_13522<sup>69</sup>
- na podstawie lematyzacji korpusu poezji przygotowano listę referencyjną form przymiotnikowych: v.6\_5136<sup>70</sup>

Poniższa Tabela 8 prezentuje dane liczbowe dotyczące list form przymiotnikowych korpusu prozy i korpusu poezji Julii Hartwig na każdym etapie przetwarzania wyników z lematyzatora.

<b>Etap przetwarzania listy przymiotników</b>	<b>Korpus proza-all</b>	<b>Korpus poezja-all</b>
v.0 – Stan początkowy	<b>668 119</b>	<b>177 414</b>
v.1 – Usunięcie wyrażen alfanumerycznych i specjalnych	3 211	897
v.2 – Usunięcie znaków końca linii	83 119	19 530
v.3 – Usunięcie form nierozpoznanych	449 555	132 191
v.4 – Ilość form przymiotnikowych (sortowanie wg opisu gramatycznego)	102 114	16 798
v.5 – Ilość form przymiotnika z powtórzeniami	59 349	7 990
<b>v.6 – Wydzielenie form przymiotnika bez powtórzeń</b>	<b>13 522</b>	<b>5 136</b>
v.7 – Wydzielenie form podstawowych przymiotnika (identyfikatory leksemów)	3 760	1 966

*Tabela 8 Etapy przetwarzania list przymiotników wygenerowanych przez lematyzator – podsumowanie.*

Otrzymane ostatecznie zestawienia pozwalają na korzystanie z listy form przymiotnikowych bądź listy z opisem gramatycznym każdej formy.

#### **4.2.2. Badania ilościowe przymiotników w korpusie prozy i poezji Julii Hartwig**

Aby móc powiedzieć na temat języka Julii Hartwig więcej niż to wynika z opisanych w literaturze przedmiotu cech i obiegowych opinii o jej warsztacie artystycznym<sup>71</sup>, obserwacji poddane zostały przymiotniki najliczniej reprezentowane w korpusie. W tym celu, przy wykorzystaniu biblioteki NLTK w środowisku Python 3.4., wygenerowane

<sup>69</sup> Zapisaną w pliku: 2021-01-13 13\_522proza\_all\_token.txt

<sup>70</sup> Zapisaną w pliku: 2021-01-09 5\_136adj\_all\_poezja\_all.txt

<sup>71</sup> Zob. Rozdział 1 niniejszej pracy.

zostały listy frekwencyjne.<sup>72</sup> (Pod pojęciem *listy frekwencyjnej* rozumiemy ponumerowaną listę słów wraz z częstotliwością ich wystąpień, w kolejności liczby wystąpień w teście (od najczęściej do najrzadziej obserwowanych). Omawiane w rozprawie listy frekwencyjne form przymiotnikowych oraz odpowiadających im leksemów zostały opracowane z wykorzystaniem odpowiednich list referencyjnych dla korpusu proza-all i poezja-all (patrz 4.2.2.1., poniżej).

#### 4.2.2.1. Formy przymiotnikowe o najwyższej frekwencji

Na najwyższych pozycjach listy frekwencyjnej form przymiotnikowych w korpusie proza-all znalazły się następujące formy:

1. była	682
2. były	438
3. innych	325
4. całą	213
5. cały	207
6. młodych	187
7. nowe	172
8. francuskiej	159
9. stanowi	158
10. nowych	150
11. byli	146
12. różnych	145
13. własnej	140
14. całe	136
15. ostatnich	135
16. drogi	132
17. innym	131
18. wielkich	124
19. pełne	121
20. musi	120
21. wielkie	120
22. inne	119
23. nowego	118
24. inni	117
25. piękne	116
26. przyjaźni	115
27. całej	113
28. młody	110
29. poetyckich	110
30. wielki	110

---

<sup>72</sup> Zestawione z wykorzystaniem skryptów udostępnionych grzecznościowo przez dr Jolantę Bachan/UAM).

Już na pierwszym miejscu listy (forma *była, były*) zauważamy homonimię form wyrazowych (*tożsamość formalna*) pochodzących od różnych leksemów (w tym przypadku od leksemu przymiotnikowego *były*, ale także od leksemu czasownikowego *być*). Zastosowany program wyliczający listę frekwencyjną zinterpretował i zestawiał interpretację przymiotnikową z czasownikową w jednej pozycji, ale nie tyle ze względów semantycznych, ile morfologicznych. Chodzi o formę *była*, która jest zarazem formą czasu przeszłego czasownika *być* oraz formą przymiotnikową występującą we frazach typu *była praca, były mąż*. W ten sposób formy homonimiczne jak *była, były* znajdujemy wysoko na liście frekwencyjnej. Z podobną sytuacją „błędu” systemu mamy do czynienia w przypadku *przyjaźni* (ranga<sup>73</sup> 26) o frekwencji 115, gdzie na jednej pozycji policzone zostały rzeczownik *przyjaźń* oraz forma przymiotnikowa w l.mn. jak np. w kontekście *przyjaźni ludzie*. Rozwiązanie problemu zawyżania pozycji rankingowej jest możliwe jedynie przez uwzględnienie kontekstu. Zwraca uwagę ponadto wysoka pozycja form *młodych* (ranga 6, frekwencja 187), *nowe* (ranga 7, frekwencja 172) i *francuskiej* (ranga 8, frekwencja 159), które stanowią wskazówkę w kierunku tematyki i cech przypisywanych opisywanych przez Hartwig obiektów.

Przyjrzyjmy się z kolei frekwencji przymiotników w korpusie poezji. Poniżej pokazujemy formy przymiotnikowe na pierwszych 30 pozycjach listy rangowej. Są to:

1. była	162
2. można	113
3. były	66
4. innych	65
5. światła	63
6. stary	62
7. cały	48
8. musi	40
9. piękne	40
10. pełne	39
11. białe	38
12. całe	37
13. piękna	37
14. byli	36
15. cała	35
16. inni	33
17. innym	33
18. nowe	32
19. młody	29
20. dobry	28

---

<sup>73</sup> Ranga wskazuje na pozycję na liście frekwencyjnej.

21. drogi	28
22. inny	28
23. mały	28
24. stara	28
25. starego	28
26. nowego	27
27. własnej	27
28. droga	25
29. pewnego	25
30. świętego	25

Na początku listy rangowej, poza kłopotliwymi formami czasownika *być*, odnotowujemy również leksem *można*, który zależnie od kontekstu daje się zinterpretować jako czasownik bezosobowy lub przymiotnik, jak np. we frazie *można pani*. Podobnie forma *światła* (ranga 5, frekwencja 63), jako rzeczownik (w liczbie mnogiej), niesie sens związany ze zjawiskiem świecenia (np. *światła samochodu*), natomiast ta sama forma, jako przymiotnik, opisuje osoby obdarzone mądrością, inteligencją, jak np. *światła kobieta*. Powtórzmy, że rozwiązanie kwestii wysokiej pozycji tych form w liście rankingowej wymaga zbadania kontekstów.

Analiza list frekwencyjnych obu korpusów wskazuje na wyraźną obecność (duża liczba wystąpień) przymiotników, które zaliczylibyśmy do *afirmatywnego pola leksykalnego* (przez *afirmatywne pole leksykalne* rozumiemy klasę określenia pozytywne dotyczące zjawisk, sytuacji i osób, takie jak *piękny/piękno, nowy/nowość, dobry/dobroć*). Szerzej zagadnienie to omówimy przy okazji obserwacji list frekwencyjnych leksemów przymiotnikowych w poszczególnych tomach analizowanego korpusu.

#### 4.2.2.2. Frekwencja leksemów przymiotnikowych w korpusie prozy

Wysoka pozycja w listach frekwencyjnych może mieć różne źródła, które weryfikuje kontekst (tłumaczy to wysoką frekwencję leksemu *były*, który po sprawdzeniu kontekstów nie występuje aż tak licznie w funkcji przymiotnika)<sup>74</sup>. W poniższym zestawieniu (Tabela 9) w kolumnach B, D, F, H, J, L, N, P, R, T zapisane zostały tytuły utworów, w kolumnach przyległych do nich (C, E, G, I, K, M, O, Q, S, U) – ilości wystąpień. Wyróżnione zostały leksemy, które występują we wszystkich listach frekwencyjnych korpusu prozy. Zauważymy wysokie frekwencje przymiotników takich jak *cały, inny, nowy*, które są leksemami częstymi w systemie języka polskiego.

<sup>74</sup> Omówiliśmy przy okazji list frekwencyjnych form przymiotnikowych przypadek leksemu czasownikowego *być*, ze względu na homonimię wyrazową, traktowany na listach jako forma przymiotnika *były*, który w funkcji przymiotnika ma niższą frekwencję. Uwaga ta dotyczy również zestawień prezentowanych dla korpusu poezji.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
Z niedalekich podróży		Apollinaire		Gerard de Nerval		Dziennik amerykański		Zawsze powroty		Pisane przy oknie		Wybrancy losu		Podziękowanie za goście		Dziennik		Dziennik tom 2		
1	były	131	były	448	były	220	były	342	były	311	były	178	były	104	były	141	były	543	były	607
2	ludowy	68	młody	219	inny	170	nowy	198	cały	151	polski	113	polski	41	nowy	129	polski	205	polski	184
3	cały	66	nowy	203	cały	160	wielki	149	nowy	147	wielki	74	inny	38	francuski	116	nowy	137	inny	138
4	nowy	62	cały	191	własny	141	młody	145	piękny	139	inny	70	poetycki	32	poetycki	116	inny	129	nowy	124
5	polski	62	pewien	188	nowy	131	inny	142	inny	133	literacki	62	duży	27	inny	86	dobry	126	wielki	116
6	inny	61	inny	171	wielki	118	amerykański	136	polski	129	nowy	58	częsty	26	własny	70	cały	122	możny	113
7	wielki	57	dobry	156	romantyczny	111	stary	125	mały	109	cały	55	możny	26	cały	63	wielki	121	dobry	112
8	piękny	51	poetycki	152	pewien	99	dobry	120	wielki	107	możny	51	własny	25	częsty	59	duży	108	duży	112
9	możny	47	stary	140	literacki	96	cały	119	stary	105	własny	48	cały	24	możny	56	częsty	96	piękny	111
10	stary	47	mały	137	młody	91	możny	117	młody	99	częsty	46	wielki	24	wielki	54	możny	92	cały	104
11	mały	41	wielki	136	pełen	88	pewien	108	duży	95	bliski	38	pewien	23	młody	52	ostatni	91	ostatni	104
12	duży	40	możny	127	dobry	87	częsty	101	ostatni	94	pełen	36	nowy	21	pełen	45	własny	86	częsty	101
13	częsty	37	piękny	118	długi	77	duży	101	pełen	88	mały	34	bliski	20	duży	38	literacki	83	trudny	99
14	młody	37	pełen	110	piękny	76	mały	85	możny	85	ostatni	34	piękny	20	trudny	38	pewien	83	bliski	98
15	długi	36	częsty	109	stary	71	pełen	84	francuski	80	dobry	33	trudny	18	literacki	36	trudny	82	literacki	80
16	daleki	35	długi	107	częsty	69	czarny	78	częsty	72	trudny	32	ostatni	17	światowy	36	piękny	81	pełen	79
17	dawny	35	własny	102	ostatni	68	własny	72	własny	72	codzienny	31	mały	16	amerykański	31	francuski	80	długi	72
18	drogi	34	bliski	91	możny	65	wysoki	69	długi	70	pewien	30	wysoki	16	piękny	31	amerykański	78	dawny	70
19	dobry	33	duży	91	dawny	61	biały	68	święty	69	francuski	29	dobry	15	współczesny	31	bliski	77	własny	70
20	pewien	30	dawny	88	mały	61	piękny	67	dobry	67	piękny	29	osobisty	15	bliski	29	młody	66	młody	68

Tabela 9. Zestawienie frekwencji leksemów przymiotnikowych dla poszczególnych dzieł w korpusie prozy

Tym niemniej możemy wskazać na semantyczne powiązanie między tytułem i tematyką poszczególnych utworów a pierwszymi na liście frekwencyjnej przymiotnikami. Są to odpowiednio:

- *ludowy* (kolumna B, ranga 2, frekwencja 68) dla książki dotyczącej odbudowy i rozwoju Polski ludowej po II wojnie światowej;
- *młody* (kolumna D, ranga 2, frekwencja 219) dla książki, która charakteryzuje życie młodego Apollinaire'a i jemu współczesnych młodych artystów francuskich;
- *inny* (kolumna F, ranga 2, frekwencja 170) to przymiotnik oddający usposobienie i styl życia Gerarda de Nerval, który zmagał się z chorobą psychiczną i był inny niż otaczające go środowisko;
- *nowy* (kolumna H, ranga 2, frekwencja 198) odnosi się do reportażu, który opowiada o pobycie Poetki w Ameryce, co było dla niej nowym doświadczeniem w wielu aspektach, odnotowanych również w sposobie pisania;
- *cały* (kolumna J, ranga 2, frekwencja 151) z kolei odnosi się do dziennika paryskiego i trudno tutaj doszukiwać się semantycznej relacji, niemniej zauważmy, że wysoko na liście frekwencyjnej (4 pozycja) znajduje się przymiotnik *piękny* (kolumna J, ranga 4, frekwencja 139), który lokuje się w polu konotacji związanych z Paryżem. Dodajmy, że jest to najwyższa ranga dla tego przymiotnika w stosunku do list dla pozostałych tekstów zawartych w korpusie.
- *polski* jest przymiotnikiem, który znalazł się na najwyższych miejscach czterech list frekwencyjnych (kolumny L, N, R, T, ranga 2), co łatwo wytłumaczyć również tematyką książek to jest: problematyka felietonów pisanych do „Więzi”, eseje o polskich artystach, dwa dzienniki poświęcone sprawom kultury oraz życia codziennego i społecznego w Polsce;
- *nowy* (kolumna P, ranga 2, frekwencja 129) natomiast jest najwyżej na liście dla książki o Francji i francuskich inspiracjach Hartwig, dlatego niewątpliwie powiązanie semantyczne zauważamy z drugim na liście przymiotnikiem *francuski* (kolumna P, ranga 3, frekwencja 116).

#### 4.2.2.3. Frekwencja leksemów przymiotnikowych w korpusie poezji

Zestawienie frekwencji leksemów przymiotnikowych w korpusie poezji zawiera 22 fragmenty list frekwencyjnych prezentowane poniżej w dwu częściach i uporządkowane według tytułów oraz chronologii wydań (tabele nr 10 i 11). Kolorami oznaczono leksemy występujące na pierwszych 20 pozycjach list frekwencyjnych (tj. posiadające rangę od 1 do 20) dla wybranych przymiotników.

Wyjątek stanowi leksem *zielony*, który występuje tylko w 4 kolumnach, ale jest ważną przesłanką do zwrócenia uwagi na wyraźną obecność leksemów związanych z barwą w listach frekwencyjnych (tabele 10 i 11), w których dominuje klasyczna opozycja *biały* (15 list) i *czarny* (11 list).

Tomy poezji stanowią zbiory tekstów bardziej otwartych, jeśli chodzi o tematykę poszczególnych utworów. Dlatego nie zauważamy tak jednoznacznych powiązań między obserwowaną frekwencją leksemów przymiotnikowych a poruszaną w dziełach problematyką, jak w przypadku książek w korpusie prozy. Interesująca wydaje się obecność leksemu *biały* oraz *stary*, dystrybucję pierwszego z nich sprawdzimy w zgromadzonych korpusach.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	U
	Pozegnania		Wolne ręce		Dwoistość	Czuwanie	Chwila postoiu	Obcowanie	Czułość	Nowe wiersze	Zobaczone	Nie ma odpowiedzi	Blyski									
1	wielki	24	były	23	inny	12	były	23	stary	17	były	36	były	35	biały	3	były	22	były	28	inny	28
2	biały	21	biały	17	były	10	czarny	13	biały	14	cały	25	nowy	23	pełen	3	inny	18	cały	15	były	23
3	mały	21	mały	10	mały	10	cały	9	były	12	mlody	19	inny	22	święty	3	biały	16	stary	12	cały	15
4	piękny	17	światły	10	stary	10	nowy	9	inny	10	inny	18	piękny	20	własny	3	cały	14	mały	11	własny	13
5	były	15	inny	8	ciemny	9	pełen	9	mlody	9	dobry	14	lekki	16	drogi	3	stary	14	mlody	10	dobry	12
6	mlody	14	piękny	8	lekki	9	biały	8	lekki	8	mały	13	mlody	16	były	2	możny	12	czarny	9	duży	12
7	nowy	13	stary	8	własny	9	stary	8	pełen	8	wysoki	13	biały	14	ciemny	2	nowy	10	święty	9	możny	12
8	cichy	11	ciemny	7	piękny	8	drewniany	7	własny	8	pełen	12	dawny	14	jesienny	2	pełen	10	nowy	8	święty	12
9	pełen	11	święty	7	cały	6	inny	7	ciemny	7	stary	12	pełen	14	kwietny	2	mały	8	biały	7	pełen	11
10	czarny	9	wielki	7	czarny	6	łagodny	7	czarny	7	wielki	11	bliski	13	niewdzialny	2	podobny	8	inny	7	daleki	10
11	duży	9	ciepły	6	duży	6	własny	7	nowy	7	własny	11	cały	13	pogodny	2	wielki	8	pełen	7	ludzki	10
12	drogi	9	duży	6	nagi	6	mlody	6	żywy	7	żywy	11	czarny	12	puasty	2	ciemny	7	duży	6	piękny	10
13	lekki	9	dobry	6	święty	6	światły	6	czernony	6	czarny	10	własny	12	suchy	2	jasny	7	ostatni	6	duży	9
14	stary	9	gesty	6	noenyy	5	wielki	6	dobry	6	lekki	10	stary	11	duży	2	lekki	7	polski	6	drogi	9
15	światły	9	mocny	6	światły	5	ciemny	5	mały	6	piękny	10	święty	11	ptasi	2	obcy	7	światły	6	nowy	9
16	wysoki	9	prawy	6	chudy	4	drogi	5	nocny	6	prawy	10	daleki	10	włoski	2	piękny	7	zły	6	stary	9
17	cały	8	wspaniały	6	ciepły	4	duży	5	piękny	6	daleki	9	duży	10	stary	2	światły	7	żywy	6	religijny	8
18	ludzki	8	zielony	6	dawny	4	ostatni	5	złoty	6	pełen	9	wielki	10	biedny	1	własny	7	ciemny	5	zły	8
19	twardy	8	chłodny	5	drogi	4	pełen	5	zły	6	suchy	9	ciemny	9	cały	1	bliski	6	dawny	5	mały	7
20	zielony	8	dawny	5	mlody	4	prawy	5	miękki	5	biały	8	drogi	9	cichy	1	daleki	6	otwarty	5	mlody	7

Tabela 10. Zestawienie frekwencji leksemów przymiotnikowych dla poszczególnych dzieł 001-011 w korpusie poezji

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
	Wiersze amerykańskie		Mówiąc nie tylko od siebie	Bez pożegnania	Zwierzzenia i błyski	To wróci	Trzecie błyski	Jasne niejasne	Gorzkie zale	Zapisane	Nowe błyski	Spojrzenie										
1	były	29	były	24	były	38	były	52	były	11	były	52	były	16	były	18	były	22	estetyczny	2	były	10
2	stary	25	inny	20	inny	11	inny	24	wielki	6	inny	25	dobry	7	inny	6	dobry	7	prawdziwy	2	dobry	6
3	nowy	20	stary	20	wielki	11	wielki	18	stary	6	wielki	24	stary	7	sześciu	6	ciemny	6	bliski	2	długi	4
4	czarny	19	cały	19	święty	10	dobry	16	piękny	5	piękny	21	mały	5	własny	6	czysty	4	hiszpański	2	wielki	4
5	cały	18	biały	18	możny	9	możny	16	inny	4	dobry	21	trudny	5	nowy	5	inny	3	bolesny	1	nieczerkiwany	3
6	inny	14	własny	16	pevien	9	pevien	15	nowy	4	nowy	18	bliski	5	pusty	5	ważny	3	były	1	wspólny	3
7	biały	13	młody	14	trudny	8	trudny	15	pelen	4	pevien	18	pevien	4	zielony	5	cichy	2	ciepły	1	daleki	3
8	amerykański	11	ciemny	13	inny	7	własny	14	rodzimy	4	możny	17	piękny	4	czysty	5	ciężki	2	gęsty	1	cały	2
9	młody	8	czarny	13	nowy	7	nowy	13	biały	3	cały	15	wysoki	4	stary	5	daleki	2	jasny	1	gęsty	2
10	możny	8	mały	12	możny	7	cały	12	czarny	3	podobny	13	bezpieczny	3	biały	4	gotowy	2	ludowy	1	mały	2
11	własny	8	ostatni	11	wielki	7	dobry	12	pevien	3	własny	13	biały	3	czarny	4	lekki	2	młody	1	naturalny	2
12	długi	7	nowy	10	cały	6	piękny	11	święty	3	zły	13	młody	3	żywy	4	mały	2	nieczerkiwany	1	piękny	2
13	duży	7	piękny	9	pelen	6	podobny	11	złoty	3	młody	12	pożny	3	bezbrotny	3	młody	2	niewidoczny	1	sprawiedliwy	2
14	mały	7	światły	9	święty	6	duży	11	angielski	3	pow szechny	11	święty	3	cały	3	możny	2	nierzozumiały	1	domowy	2
15	suchy	7	jasny	8	ciemny	5	mały	10	wysoki	3	polski	11	duży	3	dawny	3	następny	2	pelen	1	duży	2
16	wspomniały	7	pelen	8	ostani	5	ostani	10	ciemny	2	prawdziwy	10	próżny	3	jasny	3	prawdziwy	2	pevien	1	prawy	2
17	ostatni	6	zielony	8	duży	5	wysoki	9	mały	2	trudny	10	nagi	3	mały	3	prosty	2	podobny	1	lekki	2
18	otwarty	6	złoty	8	lekki	5	stary	9	młody	2	czysty	10	boży	3	monotonny	3	światły	2	pusty	1	staroswiecki	2
19	pevien	6	żywy	8	nagi	5	młody	8	niewiadomy	2	stary	10	bogaty	2	osobny	3	święty	2	rodzimy	1	uniwersytecki	2
20	światły	6	ciężki	8	polski	5	czysty	8	przejazny	2	bliski	9	cały	2	otwarty	3	wielki	2	sześciu	1	biały	1

Tabela 11. Zestawienie frekwencji leksemów przymiotnikowych dla poszczególnych dzieł 012-022 w korpusie poezji. Kolumna U zawiera listę frekwencyjną dla części tomu Błyski zebrane obejmującej tylko utwory nowe, uprzednio nie publikowane, i z tego powodu jest bardzo krótka (część ta nosi nazwę Błyski nowe). Przytoczony w U fragment tej listy nie notuje wielu wystąpień które dla tekstów obszerniejszych zajmują wysokie pozycje rankingowe.

### 4.2.3. Zastosowanie Korpusomatu – pozyskanie konkordancji dla form przymiotnikowych

Z uwagi na homonimie wyrazow i wysokie miejsca w listach frekwencyjnych leksemw homonimicznych (jak np. *były*, *światły*) naley brać pod uwag kontekst dla form odpowiadajcych tym leksemom, w celu ustalenia ich znaczenia i funkcji. Konkordancje sluz nam sprawdzeniu kontekstw uycia leksemw w korpusie, w celu uwiarygodnienia hipotez dotyczcych warsztatu pisarskiego, formuowanych na podstawie wyników ilociowych. Jednoczenie obserwacja konkordancji wybranych leksemw wskazuje na treci, do ktrych odnosz si najczstsze przymiotniki w korpusie.<sup>75</sup>

Korpusomat okaza si być wygodnym narzdciem wyszukiwania konkordancji<sup>76</sup> z powodu szeregu zalet, midzy innymi umoliwienia pracy on-line na danych jzykowych uytkownika w warunkach bezpiecznych (gwarancja bezpieczestwa danych), zgodnoc formatu danych wyjciowych z wymaganiami innych, wykorzystywanych przez nas narzdzi (patrz. Rozdzia 4.1.1.), zapewnienie obsugi serwisowej ze strony obsugi systemu.

#### 4.2.3.1. Etapy analizy

Pozyskanie kontekstw dla przymiotnikw rejestrowanych w korpusie odbywa si przez wyekstrahowanie z korpusu konkordancji. W tym celu posużyliśmy si Korpusomatem (patrz Roz. 4.1.4). Do Korpusomatu wczytane zostły oddzielne pliki wchodzce w skad poszczeglnych korpusw, a take pliki *proza-all* i *poezja-all*, w sumie 34 pliki. Dodatkowo, procz autora i tytuu, wprowadzono informacje o gatunku i dacie wydania poszczeglnych utworw.

---

<sup>75</sup> Analiza znaczeniowa wykracza jednak poza ramy, ktre wyznaczyliśmy sobie w niniejszej pracy.

<sup>76</sup> Konkordancj nazywamy uoony w list zbir wszystkich wystpie form wyrazowych wraz z kontekstem uycia w analizowanym tekcie lub korpusie jzykowym.

**KORPUSOMAT**    NOWY KORPUS   MOJE KORPUSY   INFORMACJE -   KONTAKT   PREFERENCJE   WYLOGUJ

## KORPUS: KORPUS PROZY

STAN: GOTOWY  
DATA UTWORZENIA: 2020-09-07

PARSOWANE ZALEŹNOŚCIOWE: ✓   ROZPOZNAWANIE JEDNOSTEK NAZEWNICZYCH: ✓

korpus zawiera poszczególne książki wydzielenie chronologicznie i gatunkowo

Pokaż: 10 pozycji   Szukaj:

Nazwa tekstu	Autor	Liczba segmentów	Udział	Stan	Data dodania	Operacja
01 z-wiedalekich-podrozy.txt	Hartwig	40196	4.4%	●	2020-08-09 10:58	Edytuj metadane   Usuń
02 apollinaire.txt	Hartwig	148046	16.4%	●	2020-08-07 08:54	Edytuj metadane   Usuń
03 gerard-de-nerval.txt	Hartwig	102153	11.3%	●	2020-08-07 08:55	Edytuj metadane   Usuń
04 dziennik-amerykanski.txt	Hartwig	99285	11.0%	●	2020-08-07 08:58	Edytuj metadane   Usuń
05 zawize-powroty.txt	Hartwig	119374	13.2%	●	2020-08-07 08:59	Edytuj metadane   Usuń
06 pisane-przy-oknie.txt	Hartwig	96218	10.6%	●	2020-08-07 08:59	Edytuj metadane   Usuń
07 wybrancy-losu.txt	Hartwig	24010	2.7%	●	2020-08-07 09:02	Edytuj metadane   Usuń
08 podziakowanie-za-goscine.txt	Hartwig	48594	5.4%	●	2020-08-07 09:03	Edytuj metadane   Usuń
09 dziennik-1.txt	Hartwig	109789	12.1%	●	2020-08-07 09:04	Edytuj metadane   Usuń
10 dziennik-2.txt	Hartwig	116220	12.9%	●	2020-08-07 09:05	Edytuj metadane   Usuń

Pozycje od 1 do 10 z 10 łącznie   Poprzednia 1 Następna

Zestawienie 4. Widok okna KorpSomatu prezentującego zawartość korpusu prozy Julii Hartwig

Finalnie, przetwarzaniu poddane zostały zestawy plików zebrane w dwa korpusy. Wczytanie do aplikacji osobnych dokumentów tekstowych dla każdego tomu dało od razu pogląd na pewne cechy ilościowe omawianego materiału (można je z łatwością prześledzić otwierając odpowiedni korpus, w którym zostały umieszczone). Prezentacja korpusów w takim kształcie pozwala na ogólny ogląd zebranego materiału, gdzie udział procentowy poszczególnych części odpowiada odpowiedniej liczbie słów. (Dodajmy tytułem wyjaśnienia, w że KorpSomacie słowa nazywane są segmentami; zob. w zestawieniach nr 2 i 3, gdzie kolumna trzecia jest zatytułowana *Liczba segmentów*). Zauważamy, że najliczniejszy zbiór stanowi poz. 02 „Apollinaire” (16,4%), natomiast najmniejszy udział w korpusie ma poz. 07 „wybrańcy losu” (2,7%).

KORPUSOMAT								NOWY KORPUS MOJE KORPUSY INFORMACJE - KONTAKT PREFERENCJE WYLOGUJ							
STAN: GOTOWY								DATA UTWORZENIA: 2020-09-07							
PARSOWANIE ZALÉŻNOŚCIOWE: ✓ ROZPOZNAWANIE JEDNOSTEK NAZEWNICZYCH: ✓								korpus zawiera tomy poetyckie ułożone chronologicznie i gatunkowo							
Pokaz: 25 pozycji								Szukaj: <input type="text"/>							
Nazwa tekstu		Autor		Liczba segmentów		Udział		Stan		Data dodania		Operacja			
001 pozegnania.txt		Hartwig		10329		6.0%		●		2020-08-07 09:09		Edytuj metadane   Usuń			
002 wolne rece.txt		Hartwig		7709		4.5%		●		2020-08-07 09:10		Edytuj metadane   Usuń			
003 dwoistosc.txt		Hartwig		6538		3.8%		●		2020-08-07 09:11		Edytuj metadane   Usuń			
004 czuwanie.txt		Hartwig		6741		3.9%		●		2020-08-07 09:12		Edytuj metadane   Usuń			
005 chwila postojtu.txt		Hartwig		6792		3.9%		●		2020-08-07 09:13		Edytuj metadane   Usuń			
006 obcowanie.txt		Hartwig		14655		8.5%		●		2020-08-07 09:14		Edytuj metadane   Usuń			
007 czulosc.txt		Hartwig		14178		8.2%		●		2020-08-07 09:15		Edytuj metadane   Usuń			
008 nim opatrzy sie zielen.txt		Hartwig		1236		0.7%		●		2020-08-07 09:16		Edytuj metadane   Usuń			
009 zobaczone.txt		Hartwig		9569		5.5%		●		2020-08-07 09:17		Edytuj metadane   Usuń			
010 nie ma odpowiedzi.txt		Hartwig		7534		4.4%		●		2020-08-07 09:17		Edytuj metadane   Usuń			
011 blyski.txt		Hartwig		11976		6.9%		●		2020-08-07 09:18		Edytuj metadane   Usuń			
012 wiersze amerykanske.txt		Hartwig		8785		5.1%		●		2020-08-07 09:19		Edytuj metadane   Usuń			
013 mowiac nie tylko od siebie.txt		Hartwig		12849		7.4%		●		2021-01-15 11:49		Edytuj metadane   Usuń			
014 bez pozegnania.txt		Hartwig		7277		4.2%		●		2020-08-07 09:21		Edytuj metadane   Usuń			
015 zwierzenia i blyski.txt		Hartwig		14800		8.6%		●		2020-08-07 09:22		Edytuj metadane   Usuń			
016 to wroci.txt		Hartwig		2798		1.6%		●		2020-08-07 09:23		Edytuj metadane   Usuń			
017 trzecie blyski.txt		Hartwig		17104		9.9%		●		2020-08-07 09:24		Edytuj metadane   Usuń			
018 jasne niejasne.txt		Hartwig		3347		1.9%		●		2020-08-07 09:24		Edytuj metadane   Usuń			
019 gorzkie zale.txt		Hartwig		3682		2.1%		●		2020-08-07 09:25		Edytuj metadane   Usuń			
020 zapísane.txt		Hartwig		2333		1.4%		●		2020-08-07 09:26		Edytuj metadane   Usuń			
021 blyski zebrane.txt		Hartwig		627		0.4%		●		2020-08-07 09:27		Edytuj metadane   Usuń			
022 spojrzienie.txt		Hartwig		1626		0.9%		●		2020-08-07 09:28		Edytuj metadane   Usuń			

Pozycje od 1 do 22 z 22 łącznie

Poprzednia 1 Następna

Zestawienie 5. Widok okna Korpusomatu prezentującego zawartość korpusu poezji Julii Hartwig

Dla korpusu poezji udział poszczególnych tomów w całości materiału waha się między 9,9% (poz. 017 „trzecie błyski”), a 0,7% i 0,4% dla dwóch najmniejszych (odpowiednio dla poz. 008 „nim opatrzy się zieleń” i poz. 021 „błyski zebrane”). Pozycje o najmniejszym udziale procentowym w całym korpusie stanowią wydzielone części z nowymi utworami w większych wyborach utworów Julii Hartwig. Warto również zauważyć, że pięć tomów wierszy napisanych już u schyłku twórczości, czyli poz. 016 „to wróci” (1,6%), poz. 018 „jasne niejasne” (1,9%), poz. 019 „gorzkie żale” (2,1%), poz. 020 „zapísane” (1,4%) oraz poz. 022 „spojrzienie” (0,9%) są znacznie krótsze od wcześniejszych ksiązek poetyckich. Ich udział w korpusie poezji zmniejsza się do wartości mniejszej niż 1% (ostatni tomik).

## **Pozyskanie konkordancji dla przymiotników**

Dla pozyskania ze wstępnie przygotowanego korpusu Julii Hartwig, konkordancji przymiotnikowych wykorzystane zostało narzędzie Poliqarp, będące częścią Korpusomatu, służące do generowania konkordancji. Poliqarp w odpowiedzi na zapytanie o konkordancje zwraca listę przymiotników wraz z obustronnymi kontekstami o zadanej długości. Otrzymane wyniki zawierały, oprócz opisu gramatycznego, dodatkowo metadane dotyczące tytułu utworu, roku wydania i gatunku. Poniższy przykład przedstawia początek listy konkordancji dla korpusu prozy Julii Hartwig wygenerowanej przez Korpusomat dla zapytania [pos=(adj)], (zob. linia nr 1).

Poniższy przykład przedstawia początek listy konkordancji dla korpusu prozy Julii Hartwig wygenerowanej przez Korpusomat.

A	B	C
<b>Results of query:</b> [pos=" (adj)"]		
Metadata filter:		
Korpus:	korpus prozy	
Lewy kontekst	Wynik	Prawy kontekst
Początki Niemal wszystko, co zadecydowało o	moich	związkach z kulturą francuską, miało od początku charakter przypadkowy
Niemal wszystko, co zadecydowało o moich związkach z kulturą	francuską	, miało od początku charakter przypadkowy. Najpierw znajomość języka
moich związkach z kulturą francuską, miało od początku charakter	przypadkowy	. Najpierw znajomość języka. Podczas okupacji brałam lekcje
Najpierw znajomość języka. Podczas okupacji brałam lekcje u	starej	emerytowanej nauczycielki, a nauka francuskiego była wówczas rodzajem ekstrawagancji
brałam lekcje u starej emerytowanej nauczycielki, a nauka	francuskiego	była wówczas rodzajem ekstrawagancji, inteligenci uczyli się najczęściej języka
była wówczas rodzajem ekstrawagancji, inteligenci uczyli się najczęściej języka	angielskiego	, licząc, że jego znajomość przyda im się po
licząc, że jego znajomość przyda im się po	pomyślnym	dla nas zakończeniu wojny. Ale mnie pociągało wówczas to
pomyślnym dla nas zakończeniu wojny. Ale mnie pociągało wówczas	to	wszystko, co mogłoby mnie zbliżyć do Francji.
wszystko, co mogłoby mnie zbliżyć do Francji.	Pierwsze	w moim życiu tłumaczenia poetów francuskich powstały na zamówienie autora
co mogłoby mnie zbliżyć do Francji. Pierwsze w	moim	życiu tłumaczenia poetów francuskich powstały na zamówienie autora. Antologii współczesnej
zbliżyć do Francji. Pierwsze w moim życiu tłumaczenia poetów	francuskich	powstały na zamówienie autora Antologii współczesnej poezji francuskiej, Adama
moim życiu tłumaczenia poetów francuskich powstały na zamówienie autora Antologii	współczesnej	poezji francuskiej, Adama Wazyka, który - choć nie
tłumaczenia poetów francuskich powstały na zamówienie autora Antologii współczesnej	francuskiej	, Adama Wazyka, który - choć nie miałam
zamówienie autora Antologii współczesnej poezji francuskiej, Adama Wazyka,	który	- choć nie miałam jeszcze wówczas żadnego doświadczenia w
Wazyka, który - choć nie miałam jeszcze wówczas	żadnego	doświadczenia w tej dziedzinie - powierzył mi, właściwie na
- choć nie miałam jeszcze wówczas żadnego doświadczenia w	tej	dziedzinie - powierzył mi, właściwie na los szczęścia,
Blaise'a Cendrarsa. Próba okazała się na tyle	zachęcająca	, że przekłady poezji stały się jedną z uprawianych do
się na tyle zachęcająca, że przekłady poezji stały się	jedną	z uprawianych do dziś dziedzin mojej pracy. Przypadkiem też
przekłady poezji stały się jedną z uprawianych do dziś dziedzin	mojej	pracy. Przypadkiem też, a właściwie już faktem na
na torze, było przyznane w roku 1947 stypendium rządu	francuskiego	, które umożliwiło mi trzyletni pobyt w Paryżu. Taki

Przykład 2. Początek listy konkordancji dla korpusu prozy zachowanych w pliku csv (kolumny D – M, patrz Przykład 3)

Analiza wyników zwróconych przez system pokazała, że wyjściowy zbiór konkordancji był nadmiarowy w stosunku do oczekiwanego, ponieważ zawierał konteksty dla leksemów nieodpowiadających liście referencyjnej przymiotników stosowanej przez lematyzator POLEX i TSCC. Różnica ta wynikała z odmiennych definicji kategorii przymiotnika przyjętych w słowniku POLEX i Kopusomacie opartym na słowniku Morfeusz. Dane dotyczące słów Morfeusz przejmuję ze *Słownika gramatycznego języka polskiego (SGJP)* pod redakcją Zygmunta Saloniego (Saloni, Woliński, Wołosz, Gruszczyński, Skowrońska, 2012), gdzie za leksemy przymiotnikowe uznaje się leksemy mające fleksyjną kategorię przypadku, liczby i rodzaju, czyli, oprócz przymiotników, imiesłowy przymiotnikowe, zaimki przymiotnikowe oraz liczebniki (o definicji otwartej przymiotnika pisaliśmy w rozdziale 2.2.2); zob. powyżej, w Przykładzie 2, zaimki *moich, to, tej, który* w liniach 8, 15, 17, 21,23<sup>77</sup>

Należy zwrócić uwagę na fakt, że rozumienie pojęcia *części mowy* nie jest przedmiotem konsensusu w środowisku lingwistycznym. Z tego powodu mają miejsce rozbieżności w wynikach generowanych przez narzędzia informatyczne oparte na różnych podejściach teoretycznych i źródłach<sup>78</sup>.

W rezultacie otrzymano dla korpusu prozy 107 564 konteksty wystąpień różnych form przymiotnikowych. Z kolei dla korpusu poezji liczba znalezionych kontekstów wynosi 20 178.

Przedstawione powyżej rozbieżności, będące pochodną wyżej omówionych różnic, zostały wzięte pod uwagę w interpretacji danych empirycznych dla poparcia tez niniejszej pracy.

---

<sup>77</sup> Podejście to jest przyjmowane przez niektórych autorów słowników automatycznych jako użyteczne z punktu widzenia kodowania informacji fleksyjnych (Saloni i in. 2012), jednak okazuje się nieoptymalne z punktu widzenia analizy użycia przymiotników zaliczanych do *afirmatywnego pola leksykalnego* (patrz. 4.2.2.1a.)

<sup>78</sup> Przytaczamy sposób klasyfikacji przymiotników w Korpusomacie, który jest taki sam jak w NKJP, to znaczy wyróżniono 4 klasy leksemów przymiotnikowych i oznaczono skrótami adj, adja, adjc i adjp. „Przymiotniki (adj) odmieniają się przez liczbę, przypadek, rodzaj i (nie wszystkie) przez stopień. Przymiotniki przyprzymiotnikowe (adja) i przymiotniki predykatywne (adjc) to klasy fleksymów nieodmiennych i niesamodzielných, występujące wyłącznie w specyficznych konstrukcjach. Przymiotniki przyprzymiotnikowe to fleksymy zawierające formy takie jak *polsko* w polsko-niemiecki?”. (zob. Szalkiewicz, Przepiórkowski 2012:73)

## Ujednolicenie wyników Korpusomatu z paradygmatem POLEX

Sprowadzenie wyników otrzymanych z użyciem systemu Poliqarp do założonych subkategoryzacji przymiotników zgodnej z paradygmatem POLEX polegało na wykonaniu następującej sekwencji operacji:

1. Lematyzacja listy form przymiotnikowych, stanowiących wyniki zwrócone przez system Polikarp osobno dla każdego z korpusów (zob. kolumna B w prezentowanym powyżej Przykładzie nr 4)<sup>79</sup>.
2. Użycie tych list jako wejście do lematyzatora z /lemmatizer\_from\_file-2018-01-27 (wyjście o\_file\_adj) - 2020-08-26/ celem otrzymania, dla kolumny B, interpretacji gramatycznej zgodnej z przyjętym w pracy rozumieniem kategorii przymiotnika wg słownika POLEX<sup>80</sup>.
3. Oczyszczenie ze znaków końca linii<sup>81</sup> list dla obu korpusów otrzymanych na wyjściu z lematyzatora i wstawienie ich do kolumny K w pliku csv (zob. Przykład nr 3 poniżej).
4. Sprawdzenie czy wszystkie formy z kolumny K znalazły się we właściwych dla siebie wierszach i są zgodne z kolumną B i D.
5. Przesortowanie ogółu danych zebranych w pliku csv według kolumny K zawierającej opis gramatyczny form według POLEX-u i do dalszej analizy wykorzystano wyłącznie przykłady użycia przymiotników zgodne z tym paradygmatem.

---

<sup>79</sup> Zgromadzone w dwu listach: dla korpusu prozy /2021-01-11 107 564 lista korpusomat adj proza/ oraz dla korpusu poezji /2021-01-09 lista f.adj korpusomat\_poezja/.

<sup>80</sup> Zawartym w pliku: /2021-01-09 o\_file\_poezja\_korpusomat/ oraz analogicznie dla korpusu prozy /2020-01-11 o\_file korpusomat proza-adj/

<sup>81</sup> Zapisane jako /2021-01-09 lista przymiotników korpusomat/ a następnie połączona w całość z pierwotnymi danymi z korpusomatu: /2021-01-09 results korpusomat poezja\_all\_adj/.

D	E	F	G	H	I	J	K	L
Interpretacja	Nazwa własna	Autor	Tytuł	Rok wydania	Gatunek	LEMATYZATOR POLEX	Interpretacja	
[mój:adj:pl:loc:m3:pos]		Hartwig	08podziekowanie_za_goscine	2006	esej	moich={ }	ADI3, 10) francuski	ADI3, 13) }
[francuski:adj:sg:inst:f:pos]	placeName.country	Hartwig	08podziekowanie_za_goscine	2006	esej	francuska={ francuski	ADJ1, 1) przypadkowy	ADJ1, 9) przypadkowy
[przypadkowy:adj:sg:acc:m3:pos]		Hartwig	08podziekowanie_za_goscine	2006	esej	przypadkowy={ przypadkowy	ADJ4, 5) stary	ADJ4, 7) stary
[stary:adj:sg:gen:f:pos]		Hartwig	08podziekowanie_za_goscine	2006	esej	starej={ stary	ADJ3, 4) francuski	ADI3, 8) }
[francuski:adj:sg:gen:m1:pos]	placeName.country	Hartwig	08podziekowanie_za_goscine	2006	esej	francuskiego={ francuski	ADJ3, 4) angielski	ADI3, 8) }
[angielski:adj:sg:gen:m3:pos]	placeName.country	Hartwig	08podziekowanie_za_goscine	2006	esej	angielskiego={ angielski	ADJ1, 12) pomysłny	ADJ1, 14) pomysłny
[pomysłny:adj:sg:loc:n:pos]		Hartwig	08podziekowanie_za_goscine	2006	esej	pomysłnym={ pomysłny	ADIPRO5, 3) ten	ADIPRO5, 11) ten
[ten:adj:sg:acc:n:pos]		Hartwig	08podziekowanie_za_goscine	2006	esej	to={ ten		
[pierwszy:adj:sg:nom:n:pos]		Hartwig	08podziekowanie_za_goscine	2006	esej	pierwsze={ }		
[mój:adj:sg:loc:n:pos]		Hartwig	08podziekowanie_za_goscine	2006	esej	moim={ }	ADI3, 21) francuski	ADI3, 23) francuski
[francuski:adj:pl:gen:m2:pos]	placeName.country	Hartwig	08podziekowanie_za_goscine	2006	esej	francuskich={ francuski	ADJ1, 5) współczesny	ADJ1, 7) współczesny
[współczesny:adj:sg:gen:f:pos]		Hartwig	08podziekowanie_za_goscine	2006	esej	współczesnej={ współczesny	ADJ3, 5) francuski	ADI3, 7) francuski
[francuski:adj:sg:gen:f:pos]	placeName.country	Hartwig	08podziekowanie_za_goscine	2006	esej	francuskiej={ francuski	ADIPRO4, 1) który	ADIPRO4, 9) który
[który:adj:sg:nom:m1:pos]		Hartwig	08podziekowanie_za_goscine	2006	esej	który={ który	ADIPRO2, 4) żaden	ADIPRO2, 8) }
[żaden:adj:sg:gen:n:pos]		Hartwig	08podziekowanie_za_goscine	2006	esej	żadnego={ żaden	ADIPRO5, 5) ten	ADIPRO5, 7) ten
[ten:adj:sg:loc:f:pos]		Hartwig	08podziekowanie_za_goscine	2006	esej	tej={ ten	ADIPRP, 2) zachęcający	ADIPRP, 17) }
[zachęcający:adj:sg:nom:f:pos]		Hartwig	08podziekowanie_za_goscine	2006	esej	zachęcająca={ zachęcający	ADIPRO2, 10) jeden	ADIPRO2, 13) }
[jeden:adj:sg:acc:f:pos]		Hartwig	08podziekowanie_za_goscine	2006	esej	jedną={ jeden		
[mój:adj:sg:gen:f:pos]		Hartwig	08podziekowanie_za_goscine	2006	esej	mojej={ }	ADJ3, 4) francuski	ADI3, 8) }
[francuski:adj:sg:gen:m3:pos]	placeName.country	Hartwig	08podziekowanie_za_goscine	2006	esej	francuskiego={ francuski	ADIPRO4, 3) który	ADIPRO4, 11) który
[który:adj:pl:nom:m3:pos]		Hartwig	08podziekowanie_za_goscine	2006	esej	które={ który	ADI3, 10) francuski	ADI3, 13) }

Przykład 3. Uzgodnienia interpretacji przymiotników wg paradygmatu POLEX w korpusie prozy (kolumny A, B, C – patrz Przykład 2)

Opisane kroki pozwoliły znacznie ograniczyć dane wyjściowe z Korpusomatu, a także ujednoczyć je pod względem formalnym z całością otrzymanych w badaniach wyników. W rezultacie uzyskaliśmy 11 788 form przymiotnika w kontekstach dla korpusu poezji<sup>82</sup> i 64 205 kontekstów dla przymiotnika w korpusie prozy<sup>83</sup>. Opisane powyżej etapy, a także zastosowane schematy metod sprawdzających (ujednolicanie wyników, oczyszczanie wyników z elementów nadmiarowych), służyły pozyskaniu konkordancji dla form przymiotnikowych.

#### **4.2.3.2. Charakterystyka wybranych predykatów przymiotnikowych – przykład wyników z analizy konkordancji**

Przedstawione przykłady konkordancji w centralnym punkcie zawierają przymiotnik predykatywny, który stanowi kluczowy wyraz w kontekście. Poniższa tabela różnych przykładów konkordancji jest zbudowana z pięciu kolumn: pierwsza zawiera lewy kontekst, następna – słowo kluczowe, kolejna prawy kontekst, trzy ostatnie zawierają metadane tytuł dzieła, data wydania i gatunek.

Ogół pozyskanych wyników zawiera ponad 64 000 tysięcy konkordancji dla korpusu prozy i ponad 11 000 konkordancji dla korpusu poezji. Spośród zgromadzonych rezultatów, na podstawie frekwencji przymiotników wybrano próbę dla najliczniejszych form występujących w korpusach. Z tego powodu prezentowane poniżej przykłady dotyczą przymiotników sytuujących się wysoko, w pierwszej dwudziestce, na listach frekwencyjnych dla obu korpusów.

#### **Przykłady konkordancji dla najczęstszych przymiotników w korpusie prozy i korpusie poezji**

Poniższe tabele ilustrują predykatywne użycia przymiotników m.in. wraz z odpowiednimi czasownikami występującymi w kontekstach.

Pogrubieniem oznaczono współwystępujące czasowniki, które w związku wyrazowym z przymiotnikiem regularnie się powtarzają, lub są pomijane. Jest to odmienna sytuacja w stosunku do roli czasownika w związku wyrazowym z rzeczownikiem (por. Vetulani G. 2000).

---

<sup>82</sup> Rezultaty dla korpusu poezji zawarte zostały w pliku: /2021-01-10 results korpusomat poezja\_all\_adj/.

<sup>83</sup> Rezultaty dla korpusu prozy zapisano w pliku: 2021-01-11 o\_file korpusomat proza\_all.

A	B	C	D	E
pod opieką mądrego węża, który strzeże świętej relikwii.	* Piękni	<b>jestem</b> śmy my wszyscy, kiedy, brodząc po kolana w	003dwoistosc	1971
się wybraniec fortepianu Świątosław Richter Jego oddanie muzyce <b>było</b> tak	*pełne	że zdawał się przechodzić przez życie obojętnie Trzykrotnie usuwano go	006obcowanie	1987
To co nas otacza <b>wydaje</b> się kojąco piękne i trwałe	* piękni	są również ci młodziutcy kelnerzy którzy zginiąc wdzięcznie torsy podają	006obcowanie	1987
tej nocy wietrznej i grudniowej której pośpejna godzina <b>wydaje</b> się	* bliższa	cierpieniom zwierząt Hardziej niewinnych i skazanych na milczenie	006obcowanie	1987
podziwiają i śmieją się z jego historyjek Więc pokąd <b>jestes</b>	* młoda	sprawuj rządy wygrywaj wszystko co możesz wygrać póki czas dręcz	006obcowanie	1987
wypalanej głiny sposobem wyznaczonym od wielu pokoleń Dachówkwi są <b>czernione</b>	sloneczną	<b>czernieni</b> ą dacharze noszą czapki wehłiane jaskrawe i tak malutcy są	006obcowanie	1987
W pamiętnikach swoich zanotował: te Chardiny w których <b>jestem</b>	* cały	nie żądające nic w zamian te gaje mroku Poussina doliny	007czulosc	1992
, od ciebie serce się poszerza, ciało <b>czuje</b> się	* dobre	, szczęśliwe i lekkie, z tym błękitem łagodnym,	007czulosc	1992
osłoniętym liśćmi ani frankami. O, to życie <b>jest</b>	* dobre	. To życie jest godne. I jakże godność ta	007czulosc	1992
oczywista Wyjmowano spośród nich aredyziela jak kasztany z ognia <b>JEST</b>	* INNY	I JEST Z TEGO DUMNY Pada gęsty deszcz, ale	007czulosc	1992
deszczu. Pod oknem defiluje kot, <b>czarny</b> , z	* <b>białymi</b>	łapami i <b>blazeńską</b> kropką nad nosem. Wczoraj rano plakał	009zobaczono	1999
obecne przechodziły tędy armie Śpiew strzelisty skowronka mnie goni jakże	* piękny	<b>bywa</b> trwogi głos kiedy trwoga się w pieśń przemienia Krowy	009zobaczono	1999
bliskim końcu nie odmienia swobody życia. Przeszłość <b>jest</b> zbyt	* pełna	, by podziwiać ją słowami. Przeszłość zawsze była słowem	010nie ma odpowiedzi	2001
wieczach uchwałył pogodę kot idzie samotnie drogą drzewa są takie	* młode	Sosna zieloną chorągwią milcząco mówi o śmierci dym idzie prosto	010nie ma odpowiedzi	2001
co najcenniejsze. * Był poetą. Ale nie dosć	* dobrym	, żeby pisać o głodujących dzieciach, o okrutnym zabijaniu	010nie ma odpowiedzi	2001
i tak samo mnie wzrusza kiedy jeszcze oszronionym rankiem <b>staje</b>	* cała	w swojej piękności Więc choć mieszka na północy ma prawo	011blyski	2002
wszystko, co jej się ostatnio przydarzyło. Opowiesć <b>jest</b>	* długa	i zapewne bardzo szczegółowa. Kiedy wraca do chatupy,	011blyski	2002
że zna Jamesa Porttera od siedmiu lat <b>Jest</b> to człowiek	* łagodny	i spokojny powiedział Pracuje w kuchni przy seminarium duchownym	012wiersze amerykanskie	2002
nauczanie? * Cieszę się przyrodą, jakbym <b>był</b>	* młody	. Jestem zdrowszy, a to niemal to samo -	015zwierzenia i blyski	2004
zatoczę stawu Nie było jej tam ubiegłego lata PO OMACKU	* Najpiękniejsze	<b>jest</b> to co jeszcze nieskończone Niebo pełne gwiazd niezapisane jeszcze	014bez pozegnania	2004
pod opieką mądrego węża, który strzeże świętej relikwii.	* Piękni	<b>jestem</b> śmy my wszyscy, kiedy, brodząc po kolana w	003dwoistosc	1971
się wybraniec fortepianu Świątosław Richter Jego oddanie muzyce <b>było</b> tak	*pełne	że zdawał się przechodzić przez życie obojętnie Trzykrotnie usuwano go	006obcowanie	1987
To co nas otacza <b>wydaje</b> się kojąco piękne i trwałe	* piękni	są również ci młodziutcy kelnerzy którzy zginiąc wdzięcznie torsy podają	006obcowanie	1987
tej nocy wietrznej i grudniowej której pośpejna godzina <b>wydaje</b> się	* bliższa	cierpieniom zwierząt Hardziej niewinnych i skazanych na milczenie	006obcowanie	1987

Przykład 4. Konkordancje dla przymiotnika predykatywnego w korpusie poezji



### **Przykład konkordancji dla leksemów *piękny i biały***

Obserwacja frekwencji przymiotników w utworach Hartwig, może prowadzić do wniosków, dotyczących wyborów leksykalnych dokonywanych przez Poetkę. pozwala zwrócić uwagę na niektóre wykładniki jej warsztatu pisarskiego, jak na przykład celowe operowanie powtórzeniami. W tabeli 15 (poezja) możemy zjawisko to obserwować dla leksemu *biały*.

W tabelach 12 – 15 zestawiliśmy przykładowe konkordancje dla form leksemów *piękny* i *biały* z korpusów prozy i poezji. Zajmują one wysokie pozycje na odpowiednich listach frekwencyjnych i tym samym są obficie reprezentowane w zestawieniu konkordancji. W szczególności dla form leksemu *piękny* notujemy 624 konteksty w korpusie prozy (fragment listy konkordancji patrz: Tabela 12) i 123 konteksty (fragment listy konkordancji patrz: Tabela 13) w korpusie poezji.

A	B	C	D	E
jest piętroowy biały dom z kolumnadą, ani	piękny	, ani brzydki, ot, takie sobie dworzyszcz w	01z_niedalekich_podrozy	1954
czasem uda się znaleźć wśród form pospolite przjętych jakiś kształt	piękniejszy	niz inne, świadczący o dotknięciu artysty i przypominający o	01z_niedalekich_podrozy	1954
którym chłapią się wiejszy chłopey, i ma na zewnątrz	piękną	, ludową stacją Męki Pańskiej z drzewa. Do pałacyku	01z_niedalekich_podrozy	1954
Maria pełni nadal obowiązki gospodyni podczas przyjęć u poety w	pięknym	nowym mieszkaniu pełnym książek, fetyszów i dziwacznych przedmiotów,	02apollinaire	1962
, dziś stanowiąca w znacznej mierze tylko curiosum, wzór	pięknej	, lecz i naiwnej grafiki poetyckiej, w poezji nie	02apollinaire	1962
nowej sytuacji: wymowny uwodziciel, zakłęcia, prezenty,	piękne	słowa, wszystko sprzyścięgało się przeciw jej dobrej sławie.	02apollinaire	1962
nie chce wspominać. To skrucha pełna, doskonała,	piękną	, jaka czasem zdarza się wrażliwym dzieciom, czasem dorosłym	02apollinaire	1962
okazję do zarysowania własnego światopoglądu poetyckiego. Był autorem	pięknych	korrespondencji z Niemiec, choć nie ukazują one głębszych związków	03gerard de nerval	1972
tą rolę, nie starać się sprowadzić z piedestału tego	pięknego	bóstwa, które dotąd uwielbiałem tylko z daleka.	03gerard de nerval	1972
zas szaleni są na zewnątrz. Gdyby przychodzili tu części	piękne	damy, nie miałbym na co się skarżyć	03gerard de nerval	1972
Anglika, Żyda. Pokazywała nam jego zdjęcie: bardzo	piękny	mężczyzna, blondyn o jasnym spojrzeniu, w mundurze oficera	04dziennik amerykanski	1980
człowiekiem. Inponująca jest jej umiętność techniczna, kultura,	piękny	głos, ale bez śladu znikła dawna taj emniczość zamyślenie	04dziennik amerykanski	1980
nim kobiety w chustkach i szalach, młode i dziwne	piękne	mimo znęczenia, patrzące kokieteryjnie w aparat; wokoło nich	04dziennik amerykanski	1980
piękna i pięknie ubrana: piękna twarz, postawa,	piękne	włosy, wymowne oczy. Zajęcie oglądane u Wojtki:	05zawsze powroty	2001
mnie kiedyś zachwyliło, tak jej twarz wydała mi się	piękną	i wyrazista, wydała kilka tomów wspomnień ze swojego życia	05zawsze powroty	2001
monologu przerywanego częstymi, zdrowym śmiechem. Głos miał wyjątkowo	piękny	i doskonałą dykcję, sposób mówienia i tonację pełną kultury	05zawsze powroty	2001
podczas pierwszego spaceru plażą. Ale morze było spokojne,	piękne	w inny sposób, poprzeciane poziomo łachami plaży odkrytymi przez	05zawsze powroty	2001
. Wiele wydarzeń niezamotowanych z ubiegłych dni. W sobotę	piękny	spacer z Ewą Rymkiewiczową. Po niepogodzie trwającej cały dzień	05zawsze powroty	2001
brzydoty, z brudu i niedostatku - do tej krainy	pięknego	życia, pięknych ulic i pozornie niefrasobliwych ludzi. Nazajutrz	05zawsze powroty	2001
na których rozłożone były partytury u urzekającej czystości zapisu,	piękne	jak dzieło sztuki, jak tajemniczy autograf zapisany niedostępnym profanowi	06pisane przy oknie	2004
Idea ekumenizmu i wspólnoty pisarskiej znalazła w tym projekcie swoje	piękne	urzekawistnienie. Nie było łatwo czytać wiersze w kościele.	06pisane przy oknie	2004
Przed najwzszym trybunałem. Pamięć o nim wspierać będzie jego	pięknych	dzielniceach Paryża, piszą w mieszczkańskich gabinetach swoje podstarzałe już	08podziekowanie za goscine	2006
w poźrodze Holokaustu. Żyją czy nie żyją? Bywają	piękną	biografia, jego przeszłość partyzancka i działania w opozycji demokratycznej	07wybrancy losu	2006
niezwykle mi bliskiego, co rodzi tęsknotę za czymś tak	piękni	, bywają na wprost oblakami, pograżeni w jakimś śnie	09dziennik1	2011
wewnętrznie, niecierpliwym, niejednokrotnie zaprzeczający sam sobie. Miał	pięknym	, wzniostym, a zarazem swojskim jak Chopin. Do	09dziennik1	2011
bardzo blisko potem, było sensualne wprost poczucie otaczającego mnie	piękny	umysł i inteligencję, patrzył na świat z rozległej perspektywy	09dziennik1	2011
odwiedzałam. Była to ostatnia godzina przed zmierzchem,	piękną	, architektura tego zachowanego bez uszczerbku miasta, urzekający wdźwięk	09dziennik1	2011
protestu ". Richard Lourie — wspaniały starzec o twarzy	piękne	niebo, mroczna, ale polyskliwa powierzchnia stawu, aleje	10dziennik2	2016
	pięknej	, wyzłobionej zmarszczkami, które wyglądają, jakby były wykonane	10dziennik2	2016

Tabela 12. Fragment listy konkordancji dla form leksemu piękny w korpusie prozy

A	B	C	D	E
stulicy raczył Za tło pamiętne? A jednak — stróżko	piękna	, pięknem tamtych dni Zgody z Prowansją, zgody z	001pozeznania	1956
, na głębiach welonów i uwiecznie niedbałym statkiem efemeryd;	piękny	jest dzień dzisiejszy, który szybko ginie. Na mnie	001pozeznania	1956
. Pod tą wyspą jest inna wyspa, może jeszcze	piękniejsza	. Ku niej szybuje usmiechnięta pływaczka, łącząc byskawicznym lukiem	002wolne rece	1969
bagnach na nowo używających ztemię, wydajemy okrzyki w naszym	pięknym	języku. Powiewamy kubrakami w barwach króla i królowej i	003dwoistosc	1971
przez monotonne miasteczka New Jersey i przemysłowy krajobraz tego stanu	Piękne	są tam, jedynie nowe konstrukcje mostów Wznoszą się nad niebieską	004czuwanie	1978
nawet wiatr nie zawyje. *** Bywają wiersze	piękne	i pięknie odziane. Znamy poetów, którzy urzekali mową	005chwila_postoju	1980
i zarzuciły na plecy. Ich przewodniczka o hardym i	pięknym	spojrzeniu przyciska swój kapelusz, wypełniony wspaniałymi wianami obręcz do piersi	005chwila_postoju	1980
najlepszego oddajemy na pożarcie smokowi gwałtu i przemocy Młodych chłopców	piękne	dziewczęta najlepsze umysły najbardziej obiecujące talenty damne kwiatów krzyzy słów	006obcowanie	1987
tarasem w nocy To co nas otacza wydaje się kojąco	piękne	i trwałe piękni są również ci młodzicy kelnerzy którzy zginając	006obcowanie	1987
sztuka nie stanęła jeszcze do góry nogami kochano świat i	piękne	przedmioty a przynajmniej postępowano jakby tak było miłości były szalone	007czulosc	1992
nas uważnie z głębi wód serca bijące z dala od	pięknym	niegdys ciał głosy oczekujących na dzień harmonii śpiewne sopranu echa	007czulosc	1992
Wszystkie panny są tu jasnowłose chłopcy mocno zbudowani opaleni i	piękni	jak one Jesteśmy w Visby starym szwedzkim dziś miście i	009zobaczone	1999
obecne przechodziły tędy armie Śpiew strzelisty skowronka mnie goni jakże	piękny	bywa trwogi głos kiedy trwoga się w pieśń przemienia Krowy	009zobaczone	1999
Jeżeli wiedzieli czego chcą od sztuki wystrzegali się zawsze pozorów	piękna	a kiedy ich skusiły - kajali się za odstępstwo Ale	010nie ma odpowiedzi	2001
władcą. Jakiego królestwa? Uprzytomnij to sobie. *	Piękna	to rzecz: poczucie dumy z cudzej zasługi, wywyższenia	011blyski	2002
architekturę, ale wygodę życia. * Porządek jest rzeczą	piękną	, ale wiele mam też sympatii do maniaka, który	011blyski	2002
w odświętym sukienkach w których wyglądają jak angielskie królewianka Moi	piękni	Żydzi! Patrzą na was poprzez szmat własnego życia w	012wiersze_ amerykanskie	2002
i zarzuciły na plecy. Ich przewodniczka o hardym i	pięknym	spojrzeniu przyciska swój kapelusz, wypełniony wspaniałymi wianami obręcz do piersi	013mowiac_nie_tylko_od_siebie	2003
* Krzew azalii wydaje się niepokonany w swojej trwałości	piękna	: kwitnie całymi tygodniami. Ale nadechodzi dzień, kiedy	015zwierzenia_i_blyski	2004
mniej więcej wiedzieli czego chcą Tylko mój najstarszy brat bardzico	piękny	artysta z temperamentu i playboy w dawnym stylu nie mógł	014bez_pozeznania	2004
OBIETNICA Piękna twarz Jakże	piękna	kiedy powraca zdumiewając w portrecie z Fajum Twarz bliska a	016to_wroci	2007
l'age d'emeratude, wiek szmaragdowy...	Piękna	to nazwa dla wieku, który w starych cywilizacjach otoczony	017trzecie_blyski	2008
Monge jest wyspą kołyszącą się z wiatrem w czujnej pamięci	piękne	ŁZY Ciężko ranny w wojnie Braque uciekając z domu przed	018jasne_miejasne	2009
powierzyć Z NOTATNIKA Julia usmiecha się do mnie jak do	pięknego	wspomnienia pisze w pozostawionym notatniku Ładne zapiski mówiące o ostatecznym	019gorzkie_zale	2011
dla dzieci Na moje urodziny Nodie uszyła mi w prezencie	piękną	czterwoną suknię domową Do Polski pisała do nas listy swoim	022spojrzenie	2016
pas Wenecji niebo pochmurne i różowy koral Palazzo Ducale Dzień	piękny	i samotny Wierzę że obdarowano mnie jak trzeba kiedy mdziej	022spojrzenie	2016

Tabela 13. Fragment listy konkordancji dla form leksemu piękny w korpusie poezji

A	B	C	D	E
w pozostałych częściach środkowej Polski. Kobiety noszą z targu	białe	maselnicze pachnące świeżym drzewem, odstawiają przywieszone ze sobą kosze	01z_niedalekich_podrozy	1954
. Domki są zgrabne, otoczone ogródkami, w oknach	białe	, wesole franki. Kiedy wejść na sam szczyt kolonii	01z_niedalekich_podrozy	1954
Uwięziony płomień biał równieź szupem ku górze, huczący,	biały	i groźny. Piekło, przez które przechodzi talerz przy	01z_niedalekich_podrozy	1954
gęstszej mgły, aż po horyzont, ledwo widoczny pod	białym	prawie niebem. Wzdłuż drogi zalewiska deszczowe i nagie krzewy	01z_niedalekich_podrozy	1954
szosie zmieniają wartość na przemian sosny, wiśnie i osypane	białym	, jędnym jak strąk kwiatem akacje. Łąki szerokie,	01z_niedalekich_podrozy	1954
się na obiedzie. Panna miała rude włosy i cudownie	białą	skórę rudowłosych, wysoko umieszone cieniulkie brwi, mocno pomalowane	02apolinaire	1962
mieszkałam, włożyła do przegródki opatrzonej nuncjuszem mego pokoju	białą	kopertę zaadresowaną wysokim samodzielnym piśmem, znaczek miał stempl prowincjonalny	02apolinaire	1962
czekoladek, trudno dopatrzeć się do kładnie. Spod marynarki wygląda	biały	kolnier koszuli nie skrepowany krawatem, oko wpatrzone w aparat	02apolinaire	1962
domy, z okien szpitala wychylały się siostry miłosterdzia w	białych	welonach, zaciękwione i poważne, gdzieś gdzieś przemysłki się chyłkiem	02apolinaire	1962
eleganckiej prowincjonalności, cedującej je mimo stożkowej rangi. Oparzy	białego	wina, unoszące się z licznych winiami, gospód i	03gerard_de_nerval	1972
do tego typu kobiet, niewysokich, pulchnych, o	białej	skórze i rudawych włosach, które stanowią błąd dla Nerval	03gerard_de_nerval	1972
moja wielka przyjaciółka zajęła miejsce u mego boku na swej	białej	Klacz, w srebrnym czapraku. Rzekła mi: —	03gerard_de_nerval	1972
dzieje się coś dziwnego. Gangster ubrany był w jedwabną	białą	koszulę, białe buty i flanelowe szare spodnie. Po	04dziennik_amerykanski	1980
w S. B. niewiele, przeważają schludne,	białe	, z czarnymi drzwiami i okiennicami, o śnieżnobiałych frankach	04dziennik_amerykanski	1980
dziwnego. Gangster ubrany był w jedwabną białą koszulę,	białe	buty i flanelowe szare spodnie. Po wyjściu z kina	04dziennik_amerykanski	1980
on barwy: czarnobiałe, tyle że na łapkach dłuższe	białe	rękawiczki. Wieczorowe. I śmieszny, różowy, dziecienny	04dziennik_amerykanski	1980
plaza, paki, drzewa, drapacze chmur, pojedyncze	białe	domki stojące na płaskich działkach świeżo obsianych trawą, dziwnie	04dziennik_amerykanski	1980
drogą, posuwające się ostrożnie, reflektory z trudem przebijają	białe	, mroźne kłęby powietrza. W nagrodę za szczęśliwe dowieszenie	04dziennik_amerykanski	1980
w słońcu, a ku rzece mkną stada mew jak	białe	spadające rącz. Ulicy tej przydają elegancji stare stylowe rezydencje	04dziennik_amerykanski	1980
. Fale gęste, jedna za drugą, na piasku	biała	śmietana piany. Po raz pierwszy widziałam, jak	05zawsze_powroty	2001
, długa, rzadka blond broda, twarz miękka,	biała	, oczy mało wyraziste, niebieskawe, wygląd apatyczny i	05zawsze_powroty	2001
barwy wody. Od czasu do czasu tylko przedziera się	białawe	światło skrytego pod chmurami słońca, które przez szparę wysyła	05zawsze_powroty	2001
i młodzieńce, włącznie z ochronnym kaskiem, tworzą nieskalanie	białą	całość. Kilka tylko akcentów zielonych na ubraniu ożywia tę	05zawsze_powroty	2001
Pięciu czy sześciu. Pierwsze wrażenie z mieszkania - obdrapano	białe	ściany, ciasny pokój z półkami na książki i ciasna	05zawsze_powroty	2001
, jasną perspektywę szczytów Paryża z rysującym się na tle	białego	niemał nieba profilem wieży Montparnasse'u, szczytu Pantheonu, NotreDame	05zawsze_powroty	2001
pracują dwaj wieśniacy chłopcy, doprowadzając ścianę do pięknej,	białej	gładkości. Pudełkowatość tego otwartego wnętrza, sztuczność ludzkiego pomieszczenia	05zawsze_powroty	2001
, przykrytą grubą warstwą muszli o kształtach wypukłych wachlarzy -	białych	brzoź i ciemnych sosen. W lesie jest mała mogiła	05zawsze_powroty	2001
wiatrem wysyła kilka oddziałów fal, toczących się do brzegu	białych	, żółtawych, szarawych i prawie czarnych. Wyszukując co	05zawsze_powroty	2001
spotkań krakowskich. Młde były te wieczory z Amią przy	białymi	spiętrzonymi. Ludzie chodzący i stojący na brzegu, czarni	05zawsze_powroty	2001
	białym	winie i oglądaniu dokumentów dotyczących przyjaźni. Film o Julku	10dziennik2	2016

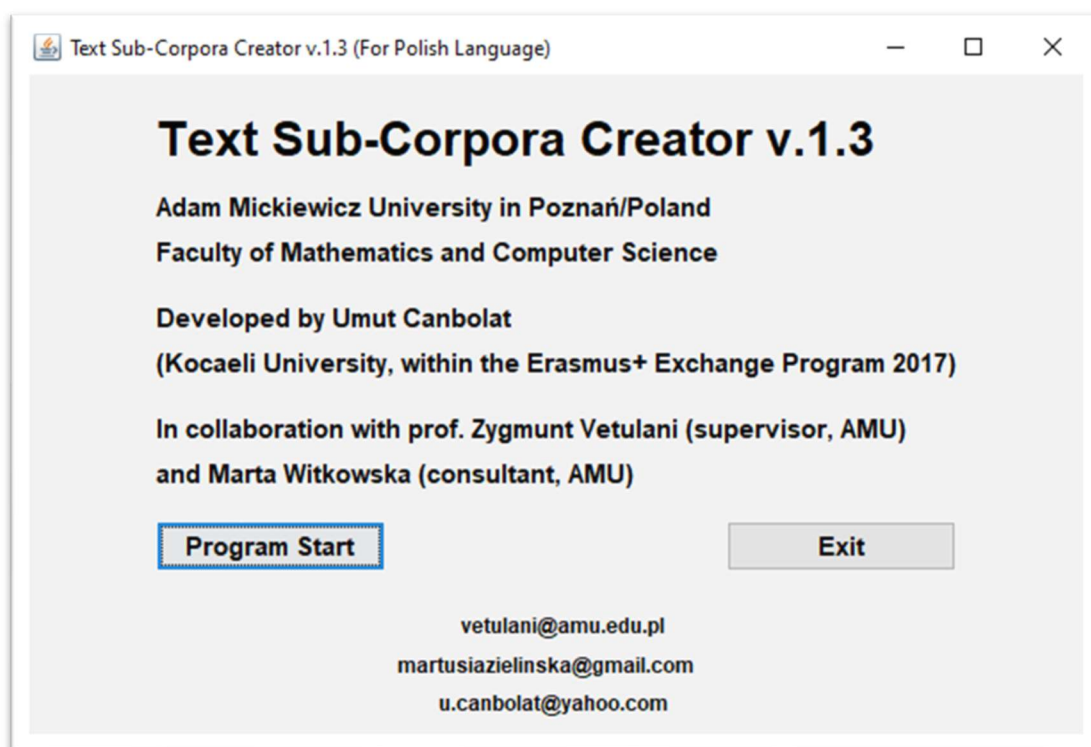
*Tabela 14. Fragment listy konkordancji dla form leksemu biały w korpusie prozy. Przykładowe wystąpienia form przymiotnikowych, które spełniają rolę predykatu zazwyczaj przy rzeczownikach.*

A	B	C	D	E
. A w sklepach senne i <b>białe</b> dziewczyny Mdleję nad	białą	nugatową skatką Z toporkiem srebrnym, jak śpiące królowy.	001pozeznania	1956
jak łagodna łania Ocieka deszczem i posłuszna leży Pod miastem	białym	, najpiękniejszym w świecie, Żywa lecz cicha i miękkiem	001pozeznania	1956
jego grobowcu pod drzewem. Na dole morze <b>bieleje</b> pod	białym	miasteczkiem rybackim. Żywi nie chodzą na wzgórze, ich	001pozeznania	1956
w tej zimowej muzyce. Przy kruczowej siostrze Marion jest	biała	Kirsten, siostry mocne. Gwałtowną pieśń murzyńska muza przekazuje	002wolne_rece	1969
chłodna rodzi trawy. Poszczepiane w męce suną żuki.	Białe	golań od miłości ślepnie. Chaos taki jak w pierwszym	003dwoistosc	1971
od tych których kochała najbardziej Nosiła jeszcze wówczas suknie z	białą	ryszką i włosy zaplecione w wianuszki na uszach Dymity w	004czuwanie	1978
wyprowadzono z domu był w spodniach płóciennych w pasy i	białym	podkoszulku Ręce zakute w kajdanki miał wykręcone do tyłu i	004czuwanie	1978
, które obala we mnie jak kručze alabastrowe naczynia:	białe	katedry, obłupane kilofem twarze świętych. Zabiera małą farmę	005chwila_postoju	1980
ona w sukni <b>białej</b> biegnie przez mokre krzewy porzeczek,	białe	pończochy w strzępach, pogubione rękawiczki, wachlarz, wsiązki	005chwila_postoju	1980
wysoka i ozdobna, poprzetykana złocieniem, dzwonkami fioletoвыми,	białym	koptrem. Przez trawę patrzy wszechwidzące oczy tumianku. Za	005chwila_postoju	1980
romkańską i wierzowice z czarnego szkła w którym przegląda się	biała	głowa migdałowego drzewa strzelisty luk mostu rozpięty nad rzeką unoszącą	006obcowanie	1987
bani Przygrywka Szumę sudyńmi liśćmi zwiabło je między złożone skrzydła	białych	krzesel i włoska muzyka pogody zamilkła w pół taktu Zahuczalo	008nim_opatrzy_sie_zielen	1995
wróciłam nieoczekiwanie na Long Island. Wasz dom w	białym	dymie krążących nad dachem mew i wypełniająca pokoje śmiertelna choroba	007czulosc	1992
ORŁOWO POD LEIDEN <b>Białe</b> niebo sierpniowe, <b>białe</b> chmury,	białe	. Jak weramon kojąca ta <b>białość</b> sierpniowa. Po szumnej	009zobaczone	1999
zupelnie pustej plaży odciętej ciepłą barwą piasku od morza niemal	białego	zlewającego się z równie <b>białym</b> niebem siedzi na szermiulej wyrzuconej	010nie_ma_odpowiedzi	2001
parują i nad sadzawką unoszą się miękkie kłęby mgły W	białej	sukni mający nad wodą melancholijna cesarzowa Sisi pod <b>białą</b> parasolką	010nie_ma_odpowiedzi	2001
szkła karetką z jękiem szybko wciąga nosze pod burym kocem	biały	niemuchomy kształt Z wysoka patrzy na nas szydzące wierzowce to	012wiersze_amerykanskie	2002
już w palacowej salce grają, a ona w sukni	białej	biegnie przez mokre krzewy porzeczek, białe pończochy w strzępach	013mowiac_nie_tylko_od_siebie	2003
owce wtórują im płaskim pobekiwaniem? I lub zanurzyć stopy w	białym	pyłę dróg biegnących wzdłuż oliwkowych sadów? Oddać się pod	013mowiac_nie_tylko_od_siebie	2003
ani mniej parasol i diamentowa spinka półmisek i metalowa sprzączka	biały	kruk i kanarek w klatce starodruk i rachunek z pralni	016to_wroci	2007
na wyspie Faro, gdzie często przemieszkował. Wyspa	biała	, jak wyrzeźbiona w kości, wietrzna, pełna niepokoju	017trzeci_blyski	2008
go w parku Poźną jestenią Spod kohnierza jestonki wyglądała	biała	Koloratka Nic wówczas nie było nam jeszcze wiadome Chwilowo Trudno	018jasne_niejasne	2009
plyniesz tam gdzie na zielonym szczycie uwieczona jest twoja pamięć	Białe	ramiona fal miosą statek sterowany tą jedną myślą odnależ ten	019gorzkie_zale	2011

Tabela 15. Fragment listy konkordancji dla form leksemu biały w korpusie poezji Poza różnymi formami przymiotnika biały pogrubione zostały też takie formy pokrewne, czasownikowe czy rzeczownikowe jak: bieleje, białość, białość. Formy dla leksemu biały w korpusie prozy występują w 250 kontekstach.

#### 4.2.4. Zastosowanie TSCC 1.3. do badań nasycenia korpusów prozy i poezji Julii Hartwig formami przymiotnika

Text Subcorpora Creator 1.3. (TSCC) umożliwił pozyskanie danych na temat rozkładu występowania form przymiotników w poszczególnych korpusach (Vetulani Z. Witkowska, Canbolat, 2018)<sup>84</sup>. Poniżej przedstawiamy ekran startowy aplikacji.



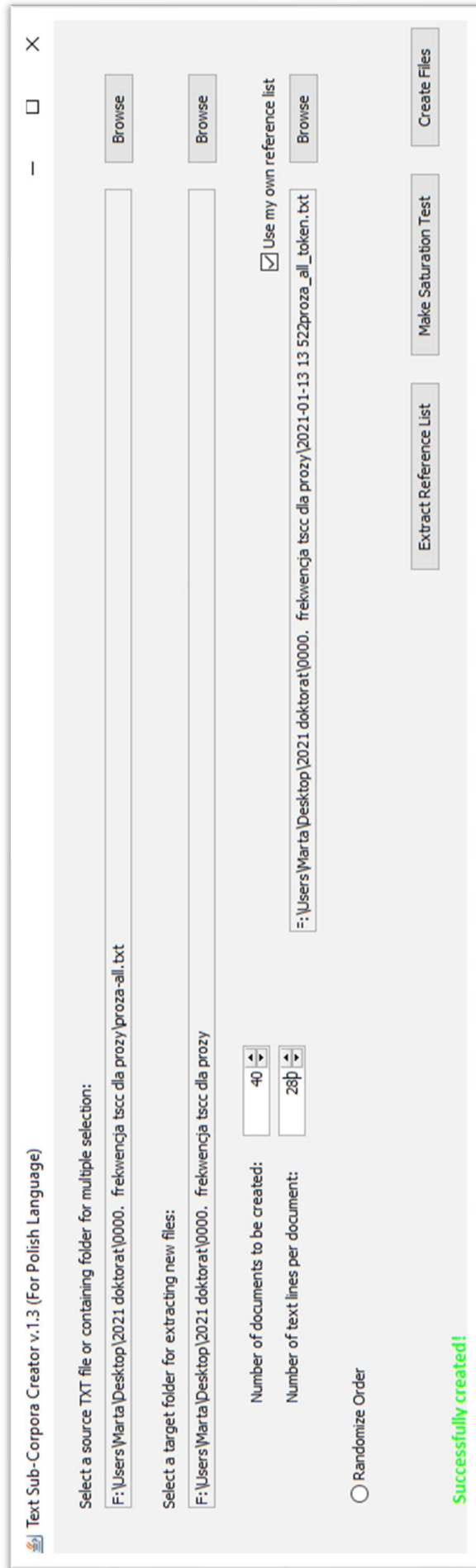
Przykład nr 6. Ekran startowy TSCC 1.3.

##### 4.2.4.1. TSCC 1.3. – metoda

W celu przeprowadzenia analizy w aplikacji TSCC 1.3. należało najpierw wczytać korpus, następnie wskazać miejsce przechowywania wyników, wczytać listę referencyjną oraz określić liczbę dokumentów (części stanowiących podkorpusy), a także liczbę linii tekstu, przypadających na dokument. Na końcu należało wskazać operację, która miała zostać przeprowadzona. Prezentujemy widok interfejsu z wczytany i przetworzonym korpusem prozy.

---

<sup>84</sup> Konsultantem implementacji systemu TSCC była Marta Witkowska.



Przykład nr 7. Funkcjonalności programu TSCC 1.3.

Program zwraca trzy pliki wynikowe, a mianowicie:

- *Korpus* podzielony na zadaną liczbę równych części, z których każda składa się z określonej przez użytkownika liczby linii,
- plik *Extracted Reference List* zawierający wszystkie słowa korpusu ułożone w listę alfabetyczną,
- folder *Saturation Test* składający się z plików z odpowiednio podzielonymi fragmentami korpusu oraz z listy z narastającą frekwencją wyrazów pojawiających się w kolejnych częściach wczytanego korpusu (*elementRanked*).

W celu przeprowadzenia badań korpus prozy Julii Hartwig (plik *proza-all*) został podzielony na 40 dokumentów po 280 linii. Poniżej prezentujemy przetworzoną postać korpusu *proza-all*, która wyświetla się w pliku *Saturation Test* (zestawienie 6). Z kolei korpus poezji rozłożony został na 50 dokumentów po 410 linii.

W przeprowadzonych badaniach najbardziej interesowały nas wyniki zawarte w pliku *Saturation Test*, tzn. listy frekwencyjne i plik *result* z danymi liczbowymi wskazujących przyrosty nowych jednostek leksykalnych w kolejnych partiach badanego korpusu.

02. 2021-01-07 TSCC > 04. Wyniki > 2021-01-13 TSCCproza\_all > Saturation Test

Przeszukaj: Saturation Test

Part 1 - elementsRanked	Part 9	Part 18 - elementsRanked	Part 26	Part 35 - elementsRanked
Part 1	Part 10 - elementsRanked	Part 18	Part 27 - elementsRanked	Part 35
Part 2 - elementsRanked	Part 10	Part 19 - elementsRanked	Part 27	Part 36 - elementsRanked
Part 2	Part 11 - elementsRanked	Part 19	Part 28 - elementsRanked	Part 36
Part 3 - elementsRanked	Part 11	Part 20 - elementsRanked	Part 28	Part 37 - elementsRanked
Part 3	Part 12 - elementsRanked	Part 20	Part 29 - elementsRanked	Part 37
Part 4 - elementsRanked	Part 12	Part 21 - elementsRanked	Part 29	Part 38 - elementsRanked
Part 4	Part 13 - elementsRanked	Part 21	Part 30 - elementsRanked	Part 38
Part 5 - elementsRanked	Part 13	Part 22 - elementsRanked	Part 30	Part 39 - elementsRanked
Part 5	Part 14 - elementsRanked	Part 22	Part 31 - elementsRanked	Part 39
Part 6 - elementsRanked	Part 14	Part 23 - elementsRanked	Part 31	Part 40 - elementsRanked
Part 6	Part 15 - elementsRanked	Part 23	Part 32 - elementsRanked	Part 40
Part 7 - elementsRanked	Part 15	Part 24 - elementsRanked	Part 32	result
Part 7	Part 16 - elementsRanked	Part 24	Part 33 - elementsRanked	
Part 8 - elementsRanked	Part 16	Part 25 - elementsRanked	Part 33	
Part 8	Part 17 - elementsRanked	Part 25	Part 34 - elementsRanked	
Part 9 - elementsRanked	Part 17	Part 26 - elementsRanked	Part 34	

Zestawienie 6. Zawartość folderu Saturation Test dla korpusu proza-all

#### 4.2.4.2. Stopień wysycenia korpusu formami przymiotnikowymi

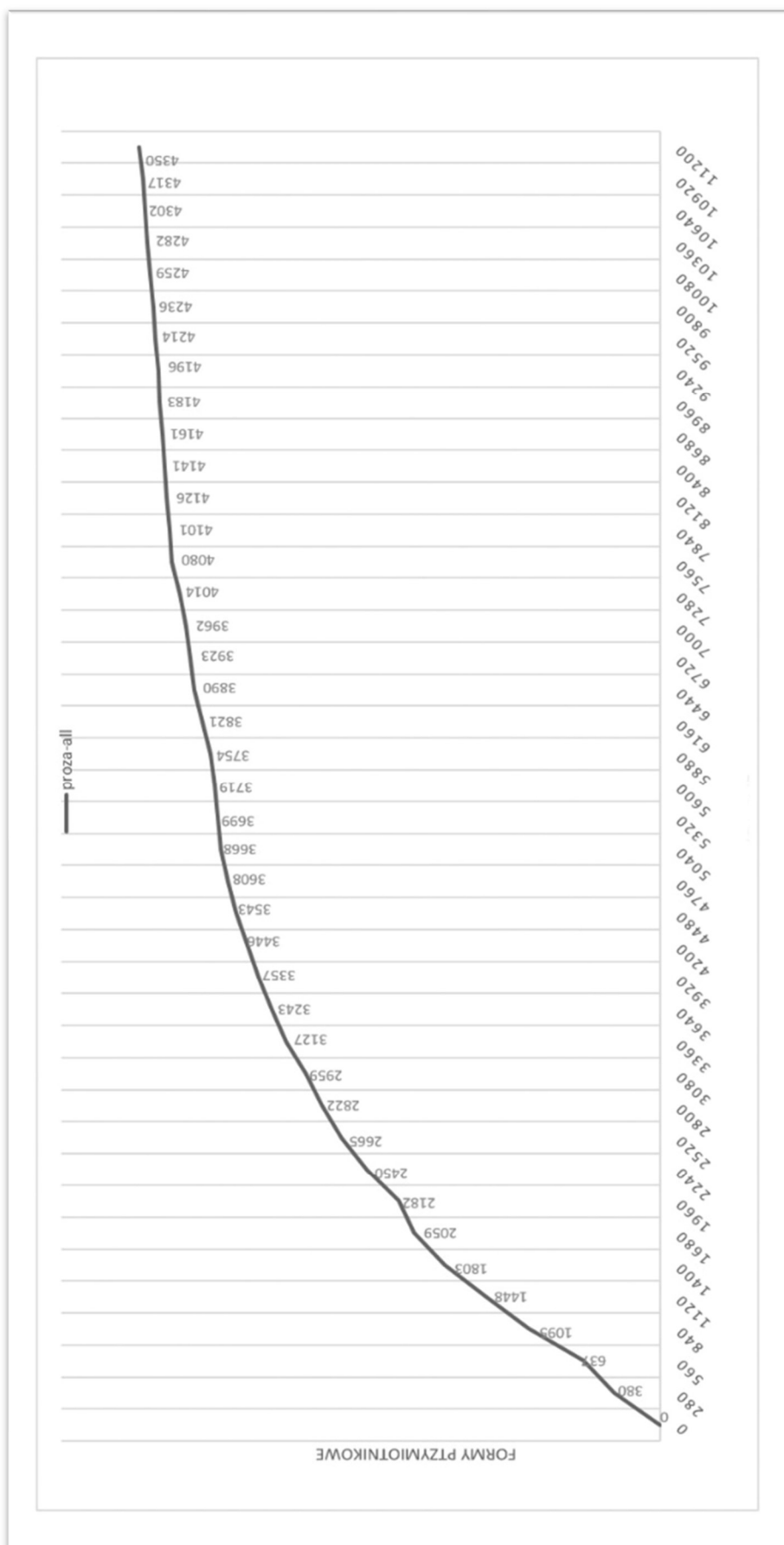
Wysycenie leksykalne jest ważnym parametrem, który dostarcza informacji dotyczącej warsztatu pisarza. W szczególności parametr ten pozwala oszacować wielkość słownictwa wirtualnego danego autora, a także wesprzeć bądź zilustrować hipotezy związane z dynamiką rozwoju warsztatu leksykalnego autora.

Poniżej prezentujemy dwa wykresy dla korpusu prozy i poezji Julii Hartwig przedstawiające przyrost nowych w form w kolejnych partiach tekstu. Przeniesienie wyników z pliku *result* na osi wykresu pozwala przedstawić stosunek przyrostu występowania nowych form przymiotnika w stosunku do linii tekstu w korpusie prozy (*krzywa nasylenia*)<sup>85</sup>. Pierwszy z nich (Wykres 1) odnosi się do przyrostu słownictwa w obrębie podkorpusu prozy, a drugi (Wykres 2) – do podkorpusu poezji.

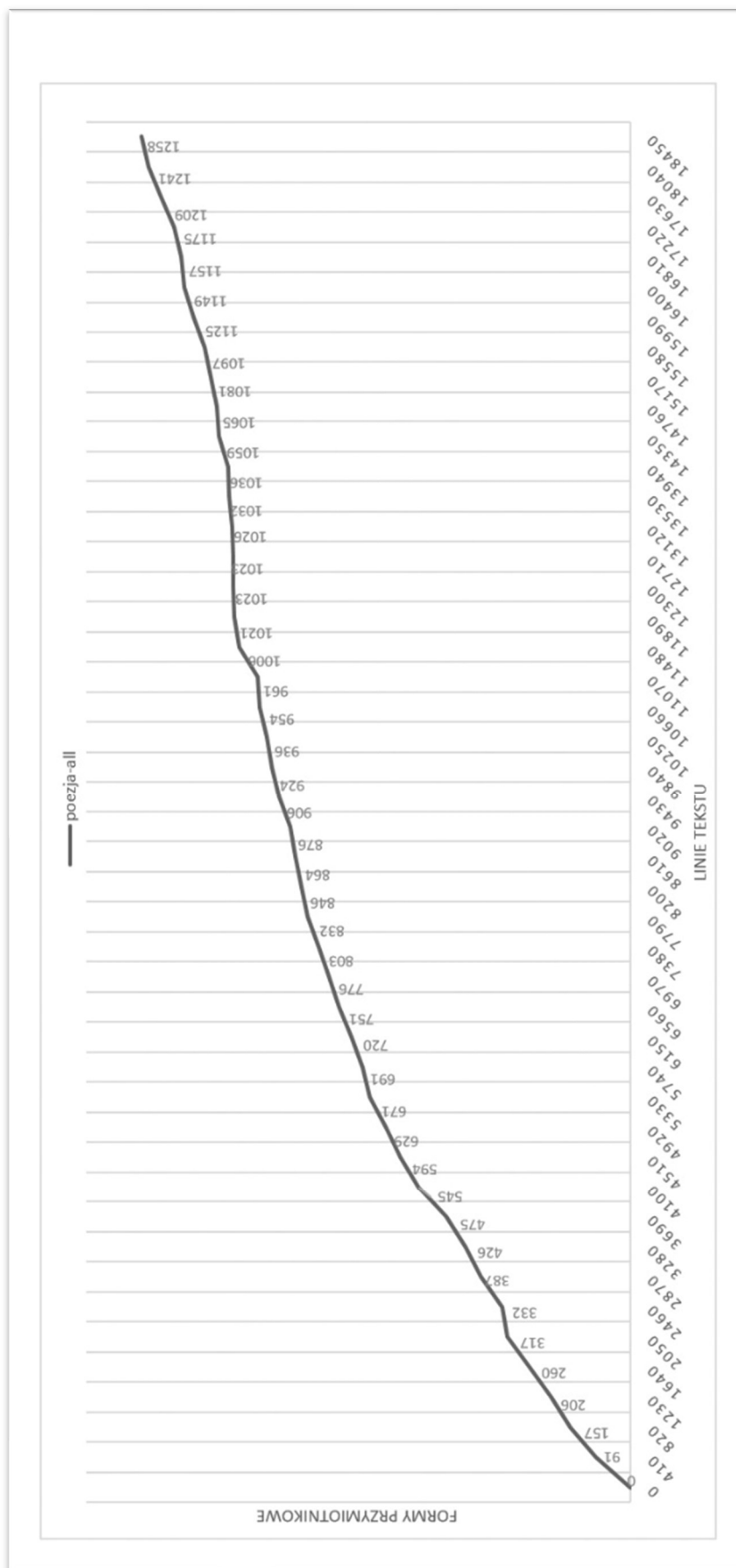
Poniższe wykresy dotyczące dwóch korpusów różnej długości wskazują, że krzywa dla obydwu wykazuje różnice. W ewentualnych dalszych badaniach należałoby powiększyć obserwowany materiał lub aktualny materiał podzielić na mniejsze części, tak aby można było lepiej zaobserwować przyrost nienotowanych wcześniej form. Tym niemniej możemy zauważyć, że dla korpusu poezji występowanie nowych form przymiotnikowych przyrasta znacząco nawet na końcu, co wynika, z jednej strony, z charakteru tego typu twórczości, a z drugiej, ze stosunkowo krótszych form wypowiedzi.

---

<sup>85</sup> Jednym z pionierów stosowania obserwacji przyrostowej w badaniach słownictwa był Charles Muller (1975) a rozpowszechnił ją Richard Kittredge (*Sublanguage – Specific Computer Aids to Translation: A Survey of the Most Promising Application Areas*; Kittredge 1983) w studiach wykonywalnością tłumaczenia automatycznego dla podjęzyków określonego języka naturalnego w ramach programu TAUM (University of Montreal). Por. też inne zastosowania z zakresu przetwarzania języka naturalnego w monografii Linguistic Problems in the Theory of Man-Machine Communication in Natural Language (Vetulani Z. 1989).



Wykres 1. Przyrost form przymiotnikowych (tokenów) dla korpusu prozy Julii Hartwig



Wykres 2. Przyrost form przymiotnikowych (tokenów) dla korpusu poezji Julii Hartwig

#### **4.2.5. Badanie rozwoju zasobu leksykalnego wykorzystywanego przez Julię Hartwig (przymiotniki)**

Badanie dotyczy ilościowych aspektów rozwoju posługiwania się zasobem leksykalnym przez Julię Hartwig w zakresie kategorii przymiotnika. Badamy metodą przyrostową powiększanie się palety leksykalnych środków wyrazu w miarę rozszerzania się kręgu tematycznego dzieł jak też wizji świata artystki. Analizie przyrostowej poddane zostały tym razem nie formy słowne (wyrazy), lecz leksemy, reprezentowane przez formy podstawowe. Obliczenia wykonane zostały przy pomocy biblioteki przetwarzania tekstu projektu GRAMLEX (Vetulani Z., Martinek, Obrębski 1997), w wersji SWI PROLOG v6) (patrz 4.1.1.) oraz standardowych narzędzi MS Word. Uzyskane diagramy przyrostowe pozwalają ocenić dynamikę rozwoju wykorzystywanego słownictwa (w zakresie przymiotnika /dla 22 tomów poezji i 10 tomów prozy/). Wykres 3 został sporządzony dla 10 tomów prozy wydanych w latach od 1954 do 2014. Tomy uporządkowane są w kolejności chronologicznej. Zawartość poszczególnych tomów została przygotowana wg zasad opisanych w Rozdziale 3. Z kolei Wykres 4 został sporządzony wg tych samych zasad dla 22 tomów prozy (1956 – 2016).

#### **4.2.6. Zastosowanie narzędzi informatycznych do analizy kompetencji językowej**

Istotną składową warsztatu pisarskiego autora jest sposób wykorzystania do zamierzeń twórczych posiadanej kompetencji językowej. Zgodnie z zapowiedzią, w badaniach zwróciliśmy szczególną uwagę na zakres operowania przez Julię Hartwig przymiotnikiem

##### **4.2.6.1. Pojęcie leksykonu wirtualnego**

W rozprawie terminu *leksykon wirtualny pisarza* używamy do określania zasobu słownictwa, którym dysponuje autor i którego używa w sposób czynny. Pojęcia pokrewne w stosunku do znaczenia, jakie nadajemy temu terminowi, są znane i używane pod różnymi nazwami przez badaczy języka i literaturoznawców od szeregu lat. Pionierem w tym zakresie był matematyk i romanista francuski Charles Muller (1909-2015), który począwszy od lat 60-tych używał pojęcia *vocabulaire virtuel* (słownictwo wirtualne)

między innymi w pracach nad warsztatem literackim pisarzy (Corneille, Racine) (patrz *Langue française et linguistique quantitative*, Muller, 1979).

W środowisku psycholingwistów polskich stosowane jest określenie *indywidualny zasób językowy*, który jest składową kompetencji językowej użytkownika mówiącego/piszącego (por. Renata Grzegorzczkowska (2012); patrz też Ida Kurcz (2005)).

#### **4.2.6.2. Zastosowanie metody przyrostowej do badania kompetencji leksykalnej pisarza**

Obserwacja danych korpusowych dla ciągłego dorobku twórczego pisarza (na określonym odcinku czasu) pozwala na oszacowanie (przy stosownych założeniach) przyrostu potencjalnego słownictwa, którym dysponuje autor (tj. jego leksykonu wirtualnego). Należy pamiętać, że zasób wirtualny nie poddaje się bezpośredniej obserwacji, lecz jego rozmiary (i dynamikę rozwoju) można szacować na podstawie obserwacji użycia, np. wykorzystując odpowiednio zorganizowany reprezentatywny korpus twórczości (omawiane w rozprawie korpusy prozy i poezji Julii Hartwig spełniają wymagane założenia). W takim wypadku wystąpienia w korpusie interesujących badacza słów stanowią zaświadczenie ich przynależności do leksykonu wirtualnego<sup>86</sup>.

Badanie przyrostowe zastosowaliśmy do ilościowego szacowania (w oparciu o korpus) rozwoju czynnej kompetencji językowej Julii Hartwig w zakresie operowania przez nią przymiotnikiem<sup>87</sup>.

Badania przeprowadzone zostały dla korpusów poezji i prozy zebranych i opracowanych w ramach przedstawionych w tej rozprawie badań. Poniżej przedstawiamy zestawienie wyników obliczeń w postaci tabel i diagramów przyrostowych. Obliczenia zostały całkowicie przeprowadzone przy pomocy opisanych w rozprawie narzędzi.

Zestawienia tabelaryczne przedstawione poniżej zostały sporządzone dla poszczególnych lat publikacji. Zarówno dla prozy jak i poezji (patrz poniżej Tabela 15 oraz Tabela 16), w latach szczególnej aktywności publikacyjnej ukazały się więcej niż jedna publikacje uwzględnione w korpusie. Były to: dla prozy rok 2006 (*Wybrańcy losu*

---

<sup>86</sup> Wygodne – i w większości przypadków dopuszczalne – jest założenie, że słowa użyte nie podlegają zapomnieniu, co oznacza, że każde użyte słowo pozostaje w leksykonie wirtualnym.

<sup>87</sup> Przypomnimy tu cytowane wcześniej w tej rozprawie wypowiedzi Julii Hartwig odnośnie roli jaką pisarka przypisuje przymiotnikowi jako środkowi wyrazu.

i *Podziękowanie*), a dla poezji lata 2002 (*Błyski i Wiersze amerykańskie*) oraz 2004 (*Bez pożegnania i Zwierzenia*).

*Dane do wykresów przyrostu słownictwa /dla prozy Julii Hartwig/*

Nr dzieła	Rok	Dzieło	Długość tekstu (ilość wyrazów tekstu <sup>88</sup> )		Liczba przym. w tekście <sup>89</sup> (L)		Słownik wirtualny przym. JH (W) (narastająco)	Przyrost sł.wirt.	
								nb. (P)	% (P/L)
1	1954	Z niedalekiej podróży	31240		1043		1043	1043	100
2	1962	Apollinaire	113725		2026		2026	1324	65
3	1972	Gerard de Nerval	78071		1521		1521	381	25
4	1980	Dziennik amerykański	78024		1729		1729	398	23
5	2001	Zawsze powroty	86491		1676		1676	307	18
6	2004	Pisane przy oknie	37013		993		993	96	10
7	2006	Wybrańcy losu	18220	54512	642	1285	3641	92	7
8		Podziękowanie	36292		1022				
9	2011	Dziennik	82618		1525		1525	180	12
10	2014	Dziennik tom 2	86854		1539		1539	150	10
<b>Łącznie:</b>			648548						

Tabela 16. Przyrosty użycia przymiotników dla prozy z lat 1954 – 2014.

Dane zestawione w Tabeli 16 posłużyły do otrzymania wykresów wizualizujących interesujące nas zjawisko powiększania się zasobu wirtualnego zasobu leksykalnego przymiotników w prozie Julii Hartwig. Kolumna 2 zawiera liczbę wyrazów tekstu zwartego (wyłączone są metadane i inne elementy pomijalne, jak np. numery stron, elementy tekstowe opisujące strukturę tekstu (tytuł, nagłówki stron, etc.). Kolumna 6 pokazuje narastająco ilości zaobserwowanych w opublikowanych dziełach przymiotników w okresie od 1954 do 2014 (łącznie 3971 różnych przymiotników). Dane z kolumny 6 należy traktować jako estymatory dolne słownictwa wirtualnego (przymiotniki). Tym samym wartości z kolumn 7 i 8 mają także charakter szacunkowy.

<sup>88</sup> Bez dat, imion własnych, etc.

<sup>89</sup> Liczba przybliżona różnych przymiotników występujących w tekście

Dane do wykresów przyrostu słownictwa /dla poezji julii hartwig/

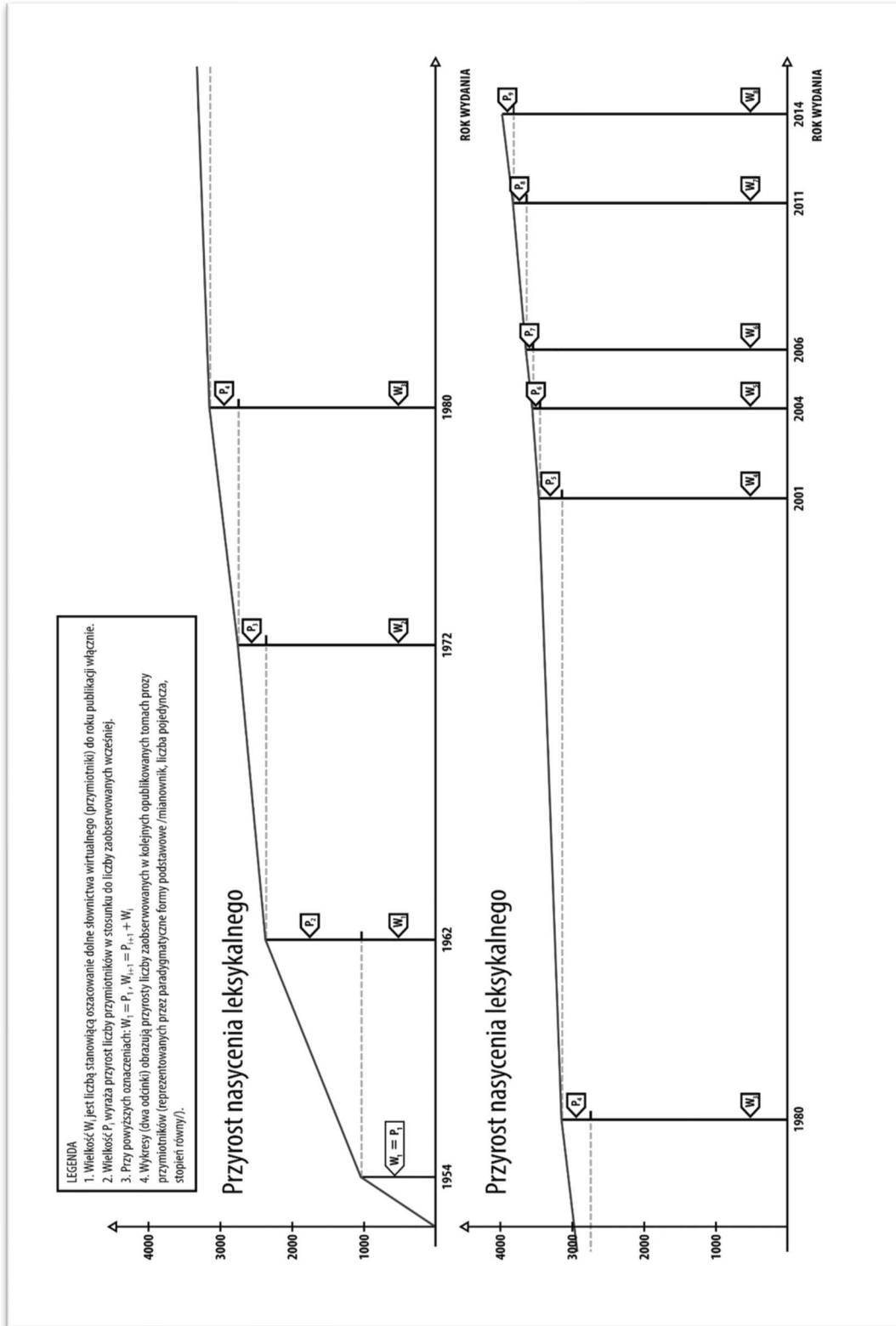
Nr dzieła	Rok	Dzieło	Długość tekstu (ilość wyrazów tekstu <sup>90</sup> )	Liczba przym. w tekście <sup>91</sup> (L)	Słownik wirtualny przym. JH (W) (narastająco)	Przyrost sł.wirt.			
						nb. (P)	% (P/L)		
1	1956	Pożegnanie	8200	496	496	496	100		
2	1969	Wolne ręce	6460	410	656	160	39		
3	1971	Dwoistość	5625	287	785	129	45		
4	1978	Czuwanie	6089	318	911	126	40		
5	1980	Chwile postoj	5294	320	1018	107	33		
6	1987	Obcowanie	13139	559	1217	199	36		
7	1992	Czułość	12674	583	1386	169	29		
8	1995	Nowe wiersze	1111	90	1401	15	17		
9	1999	Zobaczone	8761	448	1514	113	25		
10	2001	Nie ma odpowiedzi	6700	310	1563	49	16		
11	2002	Błyski	8911	16672	394	631	1647	84	13
12		Wiersze amerykańskie	7761		388				
13	2003	Mówienie tylko do siebie	10257	472	1653	6	1,3		
14	2004	Bez pożegnania	6478	17303	325	639	1816	163	
15		Zwierzenia	10825		453				
16	2007	To wróci	2547	145	1828	12	8		
17	2008	Trzecie błyski	12393	507	1945	117	23		
18	2009	Jasne niejasne	3064	168	1958	13	7		
19	2011	Gorzkie żale	3465	222	1980	22	10		
20	2013	Zapisane	2241	122	1991	11	9		
21	2014	Nowe błyski	475	47	1997	6	13		
22	2016	Spojrzenie	1539	109	2008	11	10		
<b>Łącznie:</b>			144009						

Tabela 17. Przyrosty użycia przymiotników dla poezji z lat 1956 – 2016. Patrz większy format - uzupełnienie 1

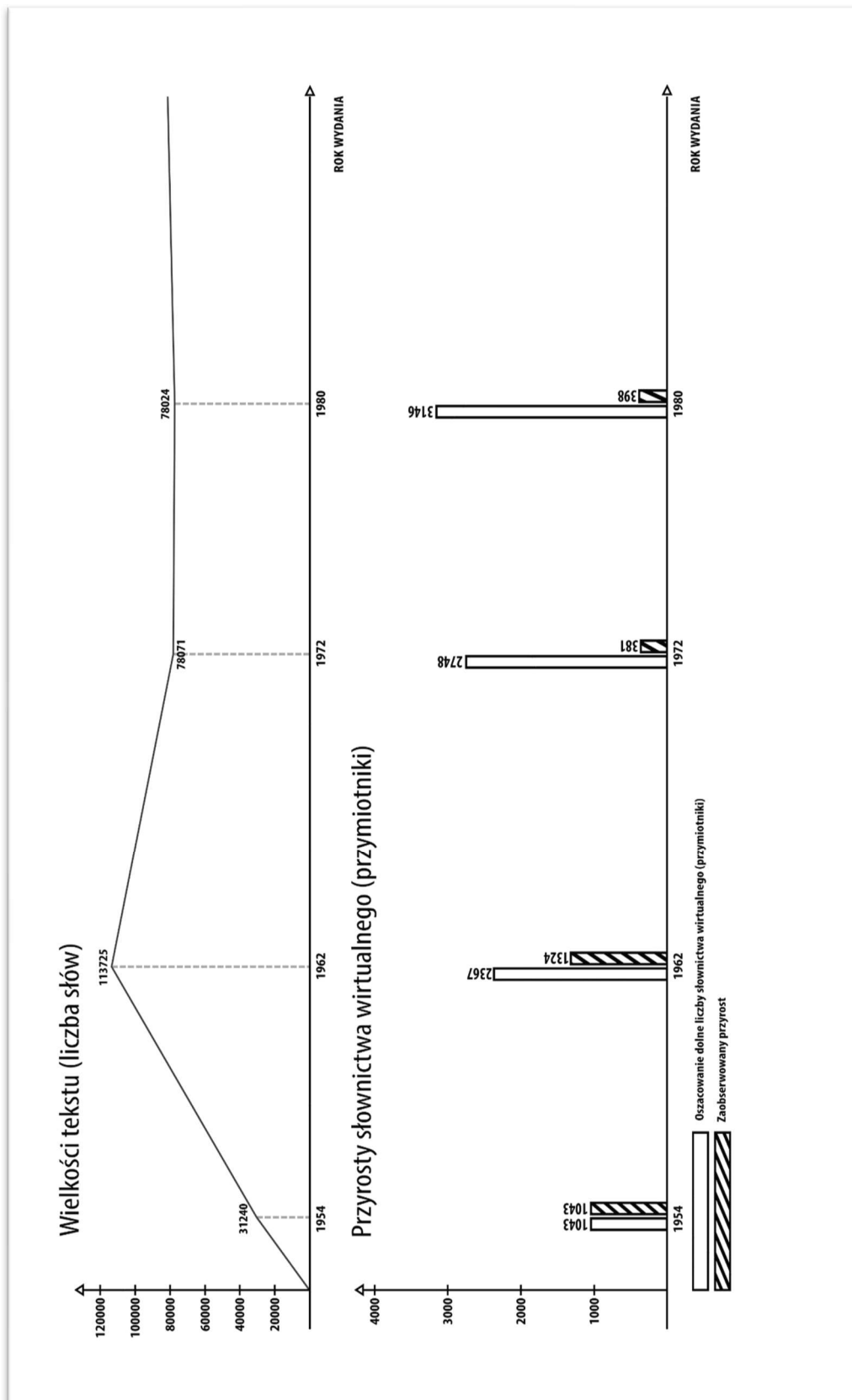
Tabela przedstawia dane dla poezji. Kolumna 6 pokazuje narastająco ilości zaobserwowanych przymiotników w tomach poezji opublikowanych w okresie od 1956 do 2016 (łącznie 2008 różnych przymiotników).

<sup>90</sup> Bez dat, imion własnych, etc.

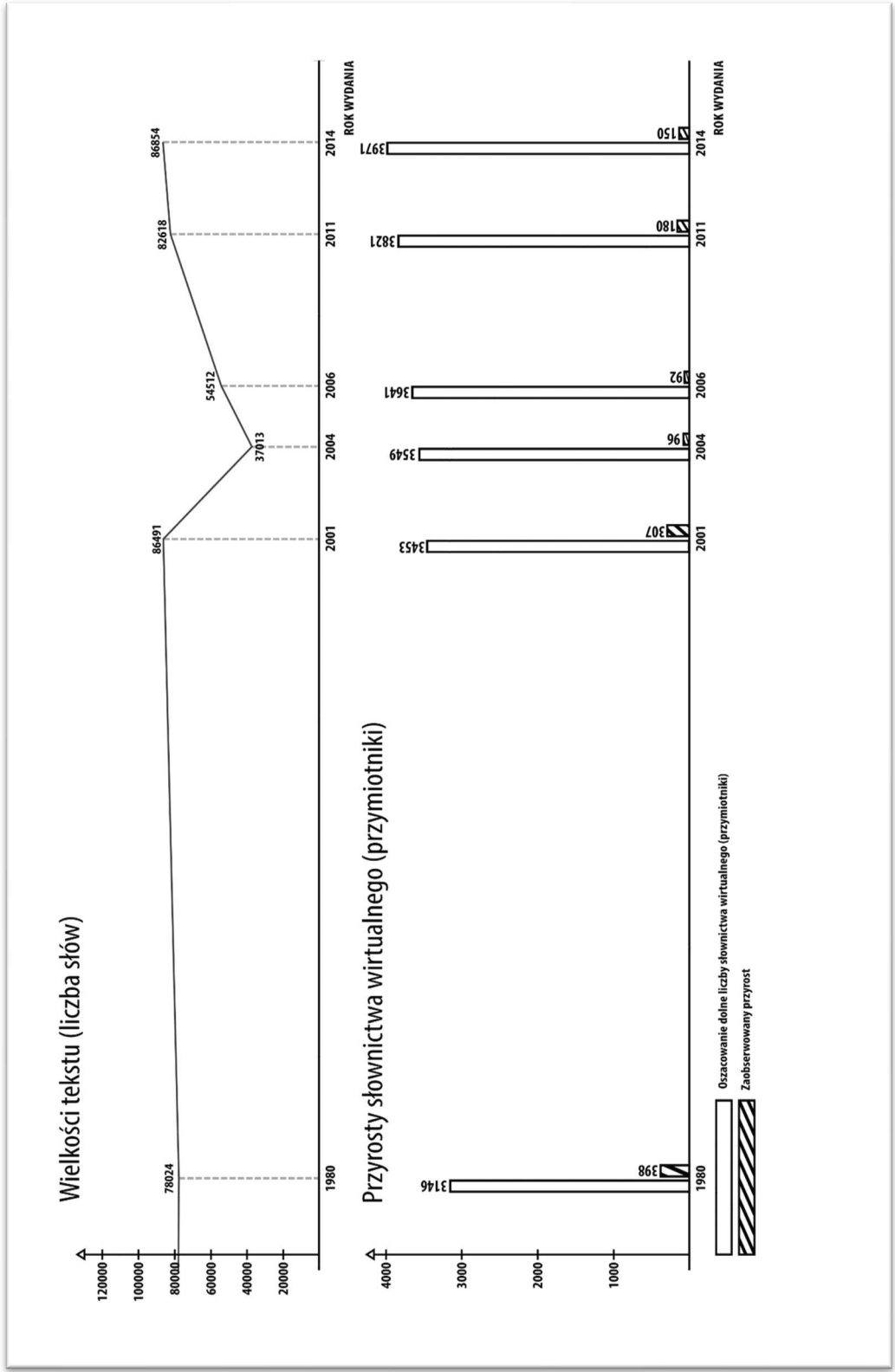
<sup>91</sup> Liczba przybliżona różnych przymiotników występujących w tekście



Wykres 3. Przyrost nasycenia leksykalnego dla korpusu prozy Julii Hartwig (lata 1954 -2016).  
 Patrz większy format - uzupełnienie 1

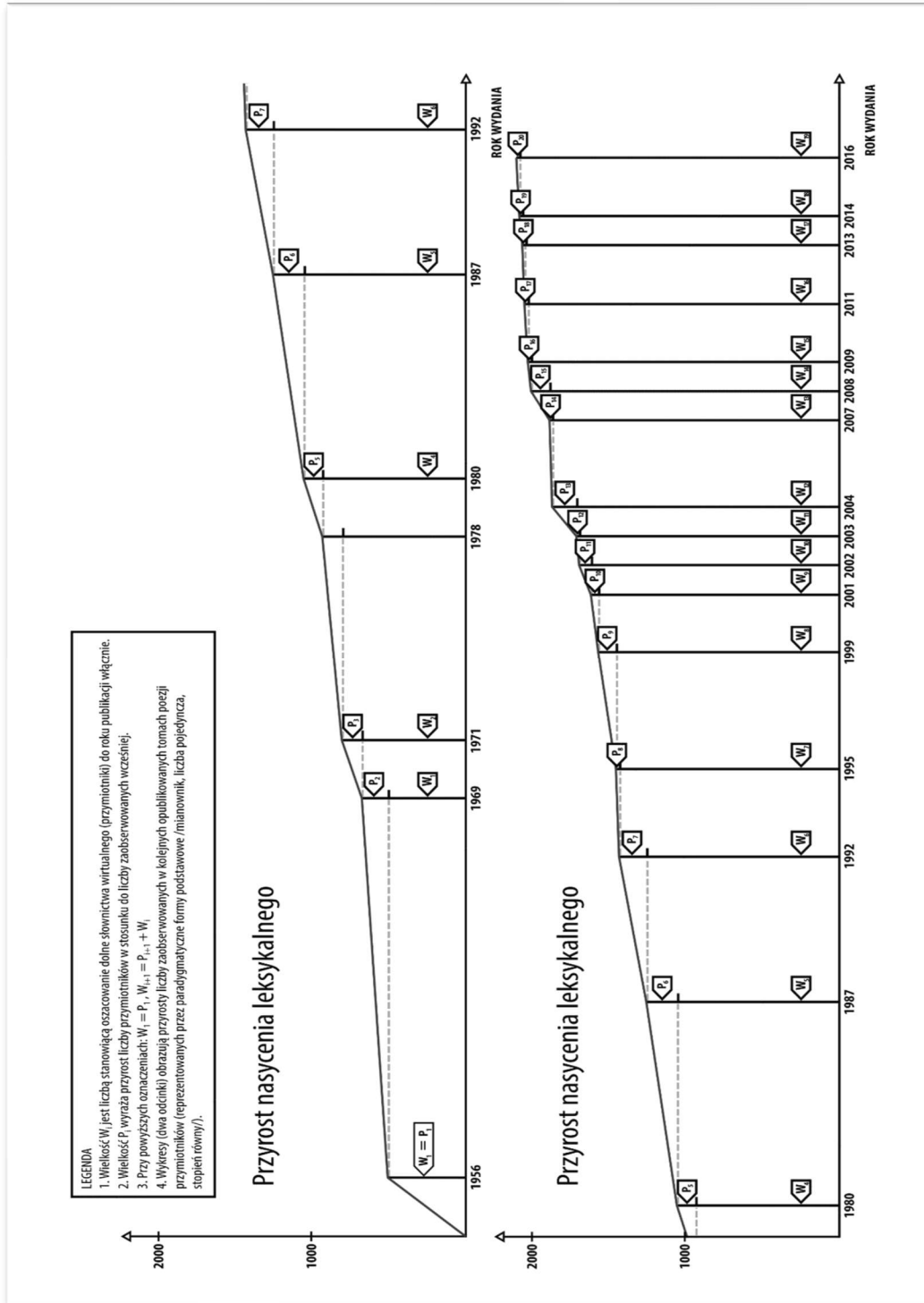


Wykres 4. Przyrosty słownictwa wirtualnego dla prozy Julii Harwig (oszacowanie dolne na lata 1954-1980). Patrz większy format - uzupełnienie 2

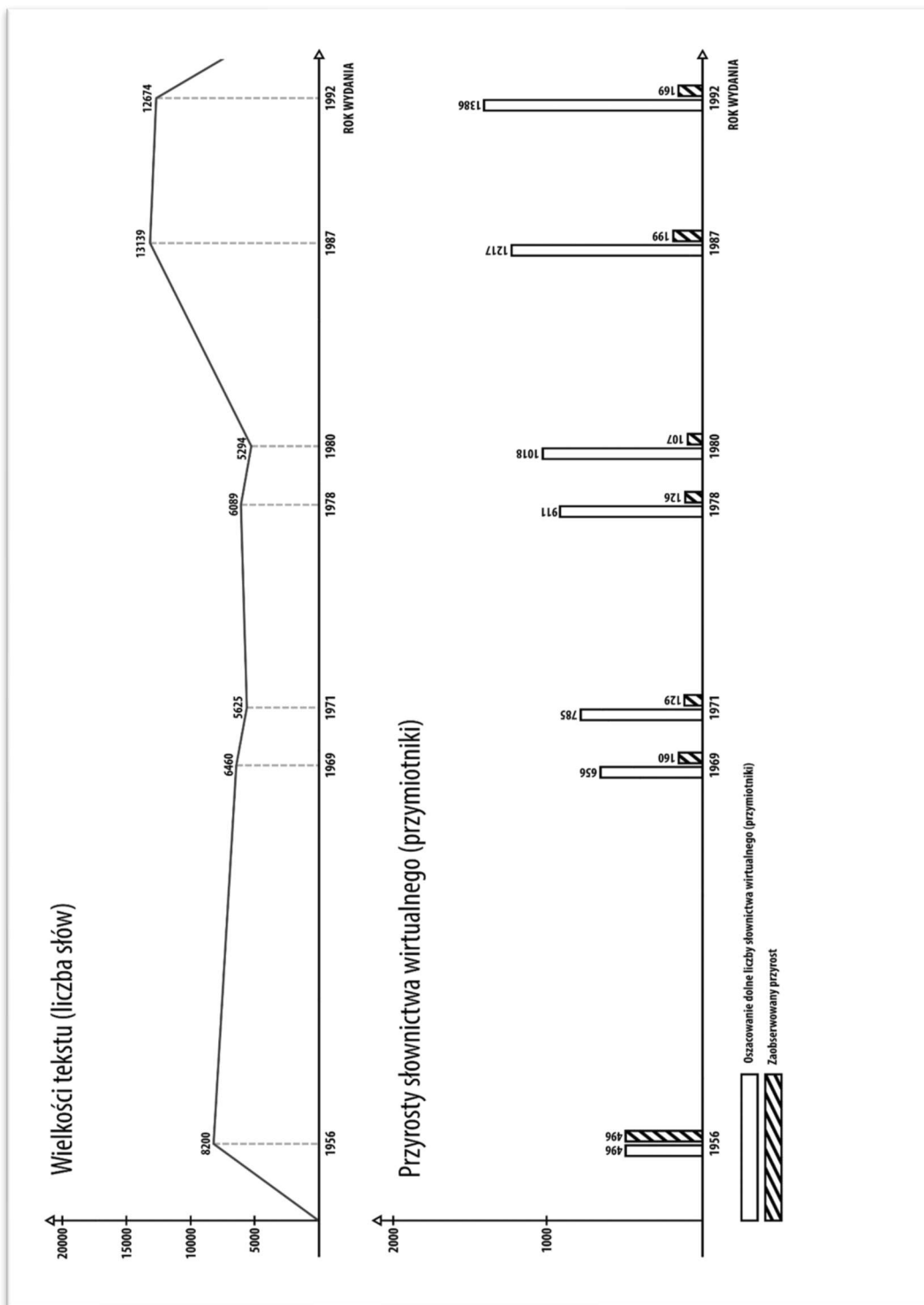


Wykres 5. Przyrosty słownictwa wirtualnego dla prozy Julii Hartwig (oszacowanie dolne na lata 1980-2014)  
 Patrz większy format - uzupełnienie 2

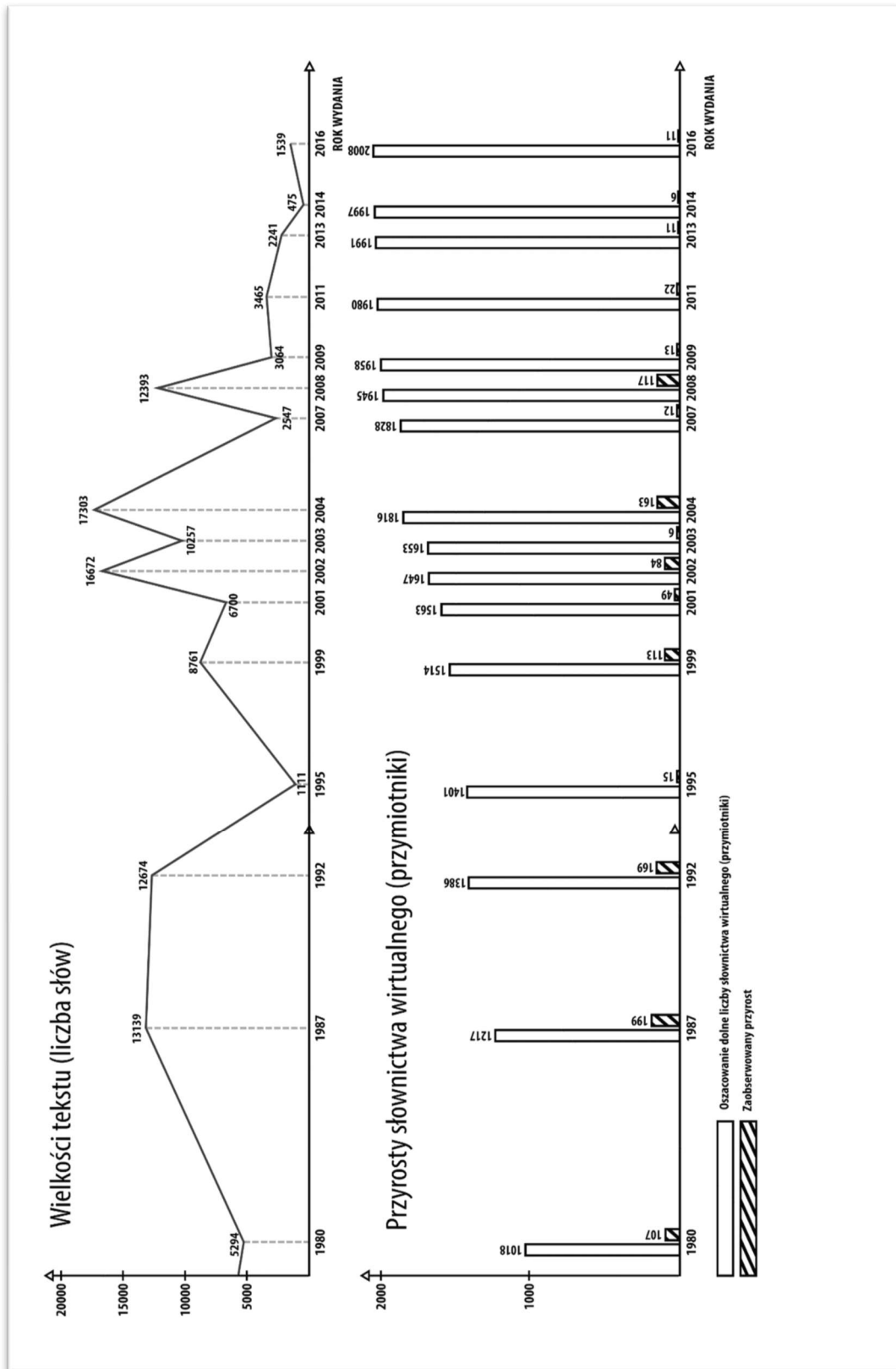
# POEZJA



Wykres 6. Przyrost nasycenia leksykalnego dla korpusu poezji Julii Hartwig (lata 1956-2016)  
 Patrz większy format - uzupełnienie 3



Wykres 7. Przyrosty słownictwa wirtualnego dla poezji Julii Hartwig (oszacowanie dolne na lata 1956-1992)  
 Patrz większy format - uzupełnienie 4



Wykres 8. Przyrosty słownictwa wirtualnego dla poezji JH (oszacowanie dolne na lata 1980-2016)  
 Patrz większy format - uzupełnienie 4

### 4.3. Problemy i korzyści wynikające z automatycznej ekstrakcji przymiotników wraz z kontekstami

Jak powszechnie wiadomo podstawowe trudności wynikające z automatycznego przetwarzania korpusu wynikają ze złożoności języka naturalnego i niedoskonałości (niedopasowania) do jego specyfiki narzędzi informatycznych. Podczas przeszukiwania tekstu, w celu wygenerowania kontekstów interesujących lingwistę, programy muszą radzić sobie z problemami związanymi z wielkością materiału, wieloznacznością językową, segmentacją zdań itd. (por. opis trudności w trakcie badań nad rzeczownikami predykatywnymi języka polskiego w (Vetulani G. 2012: 69-81).

#### 4.3.1. Trudności i ograniczenia wynikające z zastosowanej metody

W tym miejscu ograniczymy się do wymienienia najważniejszych, trudności technicznych, które zostały zidentyfikowane i pokonane w trakcie przeprowadzonych badań. Były to:

- konieczność ujednoznaczniania przyjętych definicji i założeń przy wykorzystywaniu odmiennych aplikacji (np. odnośnie kategorii przymiotnika wg formalizmu POLEX vs Morfeusz), w ten sposób, aby otrzymane wyniki mogły być ze sobą zestawiane nie tylko pod względem ilościowym, ale przede wszystkim merytorycznym,
- uwzględnienie homonimii wyrazowej,
- konieczność pominięcia konkordancji niereprezentatywnych dla analizowanych zjawisk,
- konieczność „ręcznego” sprawdzania otrzymanych postaci błyskawicznych rezultatów,
- potrzeba dostosowania do potrzeb prowadzonych badań określonych narzędzi informatycznych; w szczególności poprawiono jakość narzędzia *lemmatizer\_from\_file-2018-01-27* (wyjście o\_file\_adj) poprzez poszerzenie bazy form przymiotnika i dodanie ich do źródła "lexem".<sup>92</sup>

---

<sup>92</sup> Z założenia przy przetwarzaniu tekstów literackich możemy spodziewać się form językowych, które nie są rejestrowane przez słownik (np. neologizmy czy zapożyczenia *ad hoc*), dlatego po wykonaniu lematyzacji korpusu prozy Julii Hartwig szczególnemu oglądowi poddany został plik *failure file*. Spośród przejranych form wytypowano przymiotniki nierozpoznane przez lematyzator i zakodowano je zgodnie z formalizmem słownika POLEX. Kolejne podejście do wygenerowania danych dla korpusu prozy

### 4.3.2. Korzyści wynikające z zastosowanej metody

Mimo opisanych wyżej trudności obrana metoda badań okazała się owocna, ponieważ pozwoliła na otrzymanie korpusów przystosowanych, poprzez odpowiednie anotacje, do dalszych prac nad rolą przymiotnika w twórczości literackiej. W szczególności:

- otrzymano korpusy otagowane dla kategorii przymiotnika;
- otrzymano długie listy konkordancji dla przymiotników użytych przez Julię Hartwig (odpowiednio 11788 i 64205 jednostek dla poezji i prozy), przydatne do dalszych badań nad aspektami języka Autorki (stanowią one dokumentację źródłową charakterystycznych dla Hartwig związków wyrazowych wraz z kontekstami, ilustrującą wewnętrzne bogactwo i zróżnicowanie jej warsztatu pisarskiego),
- wykorzystane narzędzia umożliwiły sprawne przeszukiwanie całości dorobku Hartwig i wskazywanie przykładów wraz kontekstami użyc, a także ich grupowania według przypisanych metadanych,
- poprawienie funkcjonalności lematyzatora (patrz 4.3.1. powyżej) spowodowało, że zostały wykonane działania w kierunku opracowania generatora bazy form gramatycznych lematyzatora, co ma na celu półautomatyczne rozszerzanie zakresu lematyzatora POLEX przez użytkownika nieposiadającego kompetencji informatycznych.

---

z lematyzatora z nakładką przymiotnikową potwierdziło, że narzędzie właściwie rozpoznaje nowe formy przymiotnikowe dołączone do słownika w wersji: *lemmatizer\_from\_file-2018-01-27* (wyjście o *\_file\_adj*).

## ROZDZIAŁ 5

### PODSUMOWANIE

#### 5.1. Wnioski

Wielokrotnie już dowiedziono, iż twórczość Julii Hartwig, powszechnie uznawana za wybitną stanowi istotny wkład do dziedzictwa literatury i kultury polskiej, a także europejskiej.

Badania podjęte na potrzeby rozprawy miały w zamyśle poszerzenie i ugruntowanie, przy wykorzystaniu technik informatycznych i metodologii badań empirycznych, dotychczasowych ustaleń dotyczących stylu i warsztatu Julii Hartwig. Zamyśl ten został zrealizowany w sposób następujący:

1. Zasadniczym i trwałym wynikiem warsztatowym, stanowiącym dla nas podstawę aktualnych i przyszłych badań było zebranie korpusu całości publikowanej prozy i poezji Julii Hartwig z lat 1954 – 2016 i nadanie mu postaci cyfrowej, przetwarzalnej przy pomocy narzędzi informatycznych (zarówno wytworzonych na potrzeby pracy doktorskiej (TSCC), jak też istniejących niezależnie, (narzędzia pochodne od projektu POLEX czy Korpusomat). Zasób ten wraz z wytworzoną metodologią badań z zakresu humanistyki cyfrowej stanowić będzie dla nas podstawę do prowadzenia dalszych prac, z których niektóre zostały w tej rozprawie zasygnalizowane. (Na przykład wykorzystanie informacji semantyczno-leksykalnej gromadzonej w bazach typu wordnet do badania aspektów warsztatowych Julii Hartwig). W związku z powyższym stwierdzamy, że udało się sprostać trudnemu zadaniu skompletowania reprezentatywnego korpusu dzieł Julii Hartwig.<sup>93</sup>
2. Najważniejsze w tekstach Julii Hartwig, bo orzekające o świecie i odsyłające do konkretnych wrażeń, są (według niej samej) przymiotniki. Przeprowadzona analiza ilościowa pozwoliła zauważyć powiązania między tytułami i tematyką utworów prozatorskich a wysoką częstotliwością przymiotników należących

---

<sup>93</sup> Patrz Rozdziały: 3. *Korpus prozy i korpus poezji Julii Hartwig*, 4.1. *Wykorzystanie narzędzi przetwarzania języka naturalnego*.

do obszaru leksykalnego związanego z tytułem lub tematyką dzieła. Podobnej prawidłowości nie zaobserwowaliśmy dla korpusu poezji.<sup>94</sup>

3. Julia Hartwig stosuje przymiotniki zarówno w pozycji przyrzeczownikowej, jak i przyczasownikowej. W szczególności układa je w szeregi, w których przymiotniki oddzielone są przecinkami. Technika ta świadczy niewątpliwie o zwięzłości stylu i budowaniu napięcia wypowiedzi. Poetka używa przymiotników (w szczególności w określony powyżej sposób) do orzekania o elementach otaczającego świata, jak sytuacje, krajobrazy i ludzie, a także do relacjonowania swoich doświadczeń ze sztuką, podróżami i lekturami. Pisząc o ludziach często sięga po przymiotniki, charakteryzując ich fizjonomię.<sup>95</sup>
4. Znacząca przewaga reprezentacji niektórych książek nad innymi w całości badanego materiału świadczy o przewadze poruszanej w nich tematyki nad pozostałymi obszarami, co przekłada się również na reprezentację tendencji do posługiwania się konkretnymi przymiotnikami (w danym czasie, w konkretnym dziele). Obserwacja frekwencji przymiotników w korpusie Poetki pozwala wskazać na formy dominujące, które orzekają o różnych podmiotach związanych z konkretnymi zdarzeniami w życiu Julii Hartwig. Wysoko na listach frekwencyjnych w badanych korpusach plasuje się na przykład leksem *nowy*. Jest to powiązane z tym, co wiemy o Autorce, która: była wrażliwa na pojawianie się w jej otoczeniu nowych obiektów i osób, dużo podróżowała, zmieniała miejsca pobytów, była otwarta na nowiny w świecie kultury i sztuki, a także na spotkania z nowymi ludźmi.<sup>96</sup>
5. Koncentracja Julii Hartwig na wybranych zagadnieniach i obszarach otaczającego ją świata, a także na własnych doświadczeniach ma swój wyraz w strukturze częstościowej słownictwa. Stwierdzamy, na przykład, że odniesienia do zjawisk i obiektów związanych z kulturą francuską (wyrażone poprzez użycie przymiotnika *francuski*) pojawiają się częściej niż odniesienia do amerykańskiego kręgu kulturowego (jak np. przymiotnik *amerykański*). Zaobserwowane wartości frekwencji dla powyższych przymiotników możemy

---

<sup>94</sup> Patrz Rozdział 4.2.2.2. *Frekwencja leksemów przymiotnikowych w korpusie prozy*.

<sup>95</sup> Patrz Rozdziały: 2.2.2. *Formy predykcji przymiotnikowej*, 2.3. *Świat Julii Hartwig odzwierciedlony w przymiotniku*.

<sup>96</sup> Patrz Rozdział 4.2.2. *Badania ilościowe przymiotników w korpusie prozy i poezji Julii Hartwig*.

wyjaśnić poprzez odwołania autobiograficznie. (Poetka dużo pisała o swojej pracy twórczej poświęconej literaturze francuskiej).<sup>97</sup>

6. Badania frekwencyjne potwierdziły wyraźną obecność przymiotników odnoszących się do afirmatywnego pola leksykalnego, która obserwowana jest zarówno w korpusie prozy jak i w korpusie poezji (np. przymiotniki *piękny, cały, pełny*). Pozwalają one stwierdzić, że dominującą rolę w warsztacie poetyckim odgrywają przymiotniki odnoszące się do barw, w szczególności regularne wystąpienia leksemów *biały/czarny*, które notujemy w większości tomów (w przeciwieństwie do prozy). Widać również znaczący udział w wynikach frekwencyjnych przymiotników *stary/młody*, które powtarzają się na większości list frekwencyjnych leksemów dla dzieł poetyckich. Stwierdzamy, że przedstawiona dystrybucja antonimów orzekających w równie wysokim stopniu o obiektach, że są stare lub młode, białe lub czarne, potwierdza uwagi zawarte w literaturze przedmiotu, a mianowicie że w twórczości Hartwig spotykają się dwa światy (jasny i ciemny), które łączy dialog. Dialog ten jest wyrazem nadziei wybrzmiewającej z wypowiedzi poetyckich i prozatorskich Poetki.<sup>98</sup>
7. Dostrzegamy podobieństwa i różnice w dynamice przyrostu nowych leksemów przymiotnikowych w korpusach prozy i poezji Julii Hartwig. Analiza przyrostów liczby nowo obserwowanych przymiotników w korpusie prozy (patrz Tabela 16 oraz Wykres 3) oraz przyrostów liczby nowo obserwowanych przymiotników w korpusie poezji (Tabela 17 oraz Wykres 6) pokazuje, że w obu przypadkach przyrosty w ostatnich okresach twórczości są (w przybliżeniu) liniowe. Dla prozy przyrost liczby zaobserwowanych przymiotników dla ostatnich 10 lat (tj. od 2004 do 2014) wyniósł 322 (w stosunku do 3549), co oznacza 9%. Dla poezji dla ostatnich 8 lat liniowego przyrostu (od 2008 do 2016) wyniósł on 63 przymiotniki, tj. 3,2% w stosunku do stanu z roku 2008. Z porównania wynika znacznie większa dynamika przyrostu słownictwa w tekstach prozy w stosunku do tekstów poezji. Różnorodność gatunków prozy i zróżnicowanie tematyczne prozy stanowi o stosunkowo otwartym repertuarze przymiotników, natomiast

---

<sup>97</sup> Patrz Rozdział 4.2.2.1. *Formy przymiotnikowe o najwyższej frekwencji*

<sup>98</sup> Patrz Rozdziały: 2.3. *Świat Julii Hartwig odzwierciedlony w przymiotnikach*, 4.2.2.3 *Frekwencje leksemów przymiotnikowych w korpusie poezji*.

ostatnie tomiki poetyckie, charakteryzujące się krótkimi formami wypowiedzi, wykazują minimalne przyrosty liczby nowych przymiotników, co może świadczyć o tym, że Poetka w ostatnim okresie twórczości korzysta ze sprawdzonej palety środków wyrazu w zakresie przymiotników.<sup>99</sup>

## 5.2. Dalsze prace

Referowane prace zorientowane na osiągnięciu celów dalekosiężnych w zakresie analizy warsztatu pisarskiego twórcy polegały na:

- a) skompletowaniu właściwego, dobrze osadzonego w badaniach lingwistycznych i sprawdzonego w praktyce instrumentarium informatycznego,
- b) zebraniu i opracowaniu materiału empirycznego w formie korpusu przystosowanego do pracy z użyciem narzędzi informatycznych,
- c) wypracowanie metodologii badań właściwej tematu pracy i zastosowanie jej do wybranych zagadnień szczegółowych związanych z aspektami leksykalno-ilościowymi.

Wybrane wnioski i obserwacje przedstawione zostały powyżej. Inne zostały tylko zasygnalizowanych w pracy, gdyż ich uwiarygodnienie wymaga dalszych badań. W szczególności nasze plany badawcze dotyczyć będą dwóch powiązanych ze sobą nurtów, którymi są:

- a) uwzględnienie w badaniach warsztatu literackiego Julii Hartwig warstwy składniowo-semantycznej, co wymagać będzie rozszerzenia lematyzatora POLEX o moduł składniowo-semantyczny (praca w tym kierunku jest obecnie prowadzona na poziomie koncepcyjnym; przewidywane jest wykorzystanie parsera POLINT oraz leksykalnej bazy danych PolNet),
- b) przeprowadzenie badań porównawczych pomiędzy warsztatem pisarskim Julii Hartwig a aspektami warsztatowymi innych twórców o porównywalnym wpływie na rozwój literatury pięknej w języku polskim (planuje się przeprowadzenie tych badań w oparciu o wypracowane w niniejszej pracy założenia metodologiczne, z uwzględnieniem dalszych postępów w tym zakresie).

---

<sup>99</sup> Patrz Rozdział 4.2.6.2. *Zastosowanie metody przyrostowej do badania kompetencji leksykalnej pisarza*)

## **ANEKS. System kodowania fleksji polskiej POLEX**

System kodowania fleksji POLEX powstał w ramach projektu KBN POLEX (1994-1996) na potrzeby elektronicznego słownika morfologicznego języka polskiego (Vetulani Z. i inni 1978a). Został on opracowany w oparciu o dotychczasową, dobrze opisaną wiedzę z zakresu morfologii języka polskiego (patrz np. Grzegorzczkowska i in.1998). System cechuje się jednorodnym formatem opisu wszystkich kategorii morfologicznych (części mowy), tak odmiennych jak nieodmiennych, podzielonych na klasy fleksyjne (format POLEX; patrz Vetulani Z., Walczak, Obrębski, Vetulani G., 1998a). System kodowania POLEX został po raz pierwszy wykorzystany w projekcie europejskim COPERNICUS 621 GRAMLEX (Vetulani Z., Martinek, Obrębski 1998b) dla ponad 25000 leksemów (fragment słownika POLEX). Poniższe zestawienie jest cytowane za (*ibid.*).

MORPHOLOGICAL COVERAGE		
Kategoria	Liczba jednostek	Przykłady
Words	22 138	
Lexemes <sup>3</sup>	24 679	
ADJ <sup>3</sup>	3897	Biały
ADJPAP	1461	Robiony
ADJPP	35	Skostniały
ADJPRO	37	Taki
ADJPRP	775	Robiący
ADV	802	Mało
ADVANP	22	Zrobiwszy
ADVPRO	38	Gdzie
ADVPRP	379	Robiąc
APP	10	Hura
BYC <sup>4</sup>	1	Być
CONJ	89	I
EXCL	19	Dość
N	12786	Kot
NPRO	35	On
ONO	7	Miau
P	97	Do
PART	75	Nie
PPRO	3	Doń
V	4075	Jeść
VM	33	Można
VNI	2	Wiadomo
ADJ <sup>5</sup>	3897	Biały

ADJ – przymiotnik (*adjective*); ADJPAP – imiesłów bierny (*adjectival passive participle*); ADJPP – imiesłów przymiotnikowy przeszły (*adjectival past participle*); ADJPRO – zaimek przymiotnikowy (*adjectival pronoun*); ADJPRP – imiesłów przymiotnikowy teraźniejszy (*adjectival present participle*); ADV – przysłówek (*adverb*); ADVANP – imiesłów przysłówkowy uprzedni (*adverbial anterior participle*); ADVPRO – zaimek przysłówkowy (*adverbial pronoun*); ADVPRP – imiesłów przysłówkowy teraźniejszy (*adverbial present participle*); APP – wykrzyknik prymarnie apelatywny (*appellative interjection*); BYC – być (*to be*); CONJ – spójnik (*conjunction*); EXCL – wykrzyknik (*exclamation*); N – rzeczownik (*noun*); NPRO – zaimek rzeczowny (*nominal pronoun*); ONO – onomatopoea (*onomatopoeia*); P – przyimek (*preposition*); PART – partykuła (*particle*); PPRO – przyimkowy zaimek rzeczowny (*nominal prepositional pronoun*); V – czasownik (*verb*); VM – czasownik modalny (*modal verb*); VNI – czasownik nieodmienny (*non inflected verb*)

Poszczególnej kategorii gramatycznej (części mowy), spośród wymienionych w powyższej tabeli, odpowiada informacja fleksyjna, na którą składa się:

1. lista pozycji paradygmatycznych właściwych tej kategorii (dla rzeczownika będzie to lista 14 pozycyjna odpowiadająca wartościom atrybutów przypadku i liczby: 1. mianownik, l.poj., 2. dopełniacz, l.poj., ..., 7. wołacz, l.poj, 8. mianownik, l.mn., 2. dopełniacz, l.mn., ..., 7. wołacz, l.mn.)
2. zbiór zestawów (list) końcówek dla wszystkich możliwych wzorców odmiany dla tej kategorii (202 dla kategorii N /rzeczownik/).

Na potrzeby optymalnego dostosowania formalizmu opisu morfologicznego do wymogów komputerowego przetwarzania tekstu przyjęto, że:

1. zbiór końcówek jest ustalony jednoznacznie dla każdego leksemu,
2. (def.) dla danej formy wyrazowej ciąg znaków nie będących końcówką nazywamy *tematem*<sup>100</sup> (a tym samym, każda poprawna forma wyrazowa jest zbudowana z tematu i końcówki),
3. kompletny opis fleksyjny leksemu składa się z:
  - a) zestawu końcówek uporządkowanego w ciąg, którego pozycje odpowiadają pozycjom paradygmatycznym właściwym danej kategorii leksykalnej
  - b) przyporządkowania do każdego z tematów właściwej/właściwych pozycji paradygmatu (a tym samym odpowiedniej końcówki/końcówek) tak, aby utworzony ciąg znaków (wyraz) należał do tego leksemu.

Opisany formalizm POLEX został zastosowany w słowniku GRAMLEX (patrz Vetulani Z., Martinek, Obrębski, Vetulani G., 998b).

Interpretacja poszczególnych pozycji hasła słownikowego jest jednakowa dla wszystkich kategorii i zawiera cztery pozycje oddzielone średnikami. Na pozycji pierwszej rzeczownika *kot* występuje forma podstawowa /będąca identyfikatorem leksemu/ (tu *kot* /mianownik, l.poj/). Na pozycji drugiej, występuje ciąg tematów (*kot-*, *koci-*). Pozycja trzecia zawiera kod (N212) listy końcówek fleksyjnych (w przypadku hasła *kot* na ten ciąg składają się końcówki: (końcówka pusta), *a*, *u*, *a*, *em*, *e*, *e*; *y*, *ów*, *om*, *y*, *ami*, *ach*, *y* /odpowiednio dla form *kot*, *kota*, *kotu*, *kota*, *kotem*, *kocię*, *kocię*. *koty*, *kotów*, *kotom*, *koty*, *kotami*, *kotach*, *koty*/). Końcowy (czwarty) element kodu zawiera opis przyporządkowania końcówki określonego tematowi (np. tematowi *koci-* przyporządkowana jest końcówka z pozycji 6-tej (-e) co daje formę miejscownika (np. „Cała nadzieja w *kocie*”), l/poj, jak również końcówka z pozycji 7-mej(-e) co generuje formę wołacza (np. „O mój *kocie!*). Ten sam format jest stosowany do wszystkich słów języka polskiego, także nieodmiennych. Zauważmy, że w opisie słowa nieodmiennego *wiadomo* występuje tylko jeden temat, który jest tu identyczny z formą podstawową (lista końcówek dla słów nieodmiennych jest jednoelementowa i zawiera końcówkę pustą). W hasle dla rzeczownika *kot* występują dwa tematy *kot-* oraz *koci-*

---

<sup>100</sup> Uwaga. Wprowadzonego tu znaczenia terminu *temat* nie należy mylić tradycyjnym rozumieniem tego terminu w innych opracowaniach lingwistycznych z zakresu morfologii wyrazów.

„2: 6,7” informuje, że z tematem drugim mogą się łączyć końcówki z pozycji 6 i 7. W haśle przymiotnika *biały* występuje siedem tematów *biały-*; *biał-*, *bial-*, *bielsz-*, *biels-*, *najbielsz-*, *najbiels-*. Wśród obserwowanych form odmiennej przymiotnika biały nie występuje forma odpowiadająca pozycji paradygmatycznej 86, co odnotowane jest w haśle słownikowym przez „0: 86”.

Algorytm lematyzacji w oparciu o powyższe założenia został opisany w (Vetulani Z., Walczak, Obrębski, Vetulani G., 1998a). Jego implementacja w SWI Prologu była wykorzystywana w referowanych w niniejszej rozprawie pracach.

Poniżej podajemy trzy przykłady ilustrujące hasła tego słownika:

- ADJ (przymiotnik):
  - akustyczny; akustyczn; ADJ1; 0: 29-84,86; 1: 1-28,85
  - biały; biały, biał, bial, bielsz, biels, najbielsz, najbiels; ADJ1; 0: 86; 1: 1-18, 20-26, 28, 85; 2: 19, 27; 3: 29-46, 48-54, 56; 4: 47, 55; 5: 57-74, 76-82, 84; 6: 75, 83
- N (rzeczownik):
  - kot; kot, koci; N212; 1: 1-5,8-14;2: 6,7
- WNI (słowo nieodmienne):
  - wiadomo; wiadomo; VNI;1: 1

Podobnie jak dla rzeczownika (patrz przykład kot zanalizowany powyżej), do zrozumienia zawartości hasła dla leksemu wystarczająca jest znajomość 86-cio elementowego zestawu końcówek paradygmatycznych ADJ1. Są to końcówki: -y, -a, -e, -ego, -ej, -emu, -ej, -ego, -y, -ą, -e, -ym, -ą, -ym, -ej, -y, -a, -e, -i, -e, -ych, -ym, -ych, -e, -ymi, -ych, -i, -e, -y, -a, -e, -ego, -ej, -emu, -ej, -ego, -y, -ą, -e, -ym, -ą, -ym, -ej, -y, -a, -e, -i, -e, -ych, -ym, -ych, -e, -ymi, -ych, -i, -e, -y, -a, -e, -ego, -ej, -emu, -ej, -ego, -y, -ą, -e, -ym, -ą, -ym, -ej, -y, -a, -e, -i, -e, -ych, -ym, -ych, -e, -ymi, -ych, -i, -e, -o, -u.

Zestaw ten, wraz z opisem ilustruje poniższa tabela.

Poz.,	Końcówka	Opis przykładu	Przykłady
1.	-y	Stopień: równy; Liczba: pojedyncza; Przypadek: mianownik; Rodzaj: męskoosobowy, męskozwierzęcy, męskonieożywiony	<i>biały</i> człowiek, <i>biały</i> koń, <i>biały</i> dom
2.	-a	Stopień: równy; Liczba: pojedyncza; Przypadek: mianownik; Rodzaj: żeński	<i>biała</i> kobieta
3.	-e	Stopień: równy; Liczba: pojedyncza; Przypadek: mianownik; Rodzaj: nijaki	<i>białe</i> kaczątko
4.	-ego	Stopień: równy; Liczba: pojedyncza; Przypadek: dopełniacz; Rodzaj: męskoosobowy, męskozwierzęcy, męskonieożywiony, nijaki	<i>białego</i> człowieka, <i>białego</i> konia
5.	-ej	Stopień: równy; Liczba: pojedyncza; Przypadek: dopełniacz; Rodzaj: żeński	<i>białej</i> myszy
6.	-emu	Stopień: równy; Liczba: pojedyncza; Przypadek: celownik; Rodzaj: męskoosobowy, męskozwierzęcy, męskonieożywiony, nijaki	<i>białemu</i> człowiekowi, <i>białemu</i> koniowi, <i>białemu</i> domowi, <i>białemu</i> kaczątku
7.	-ej	Stopień: równy; Liczba: pojedyncza; Przypadek: celownik; Rodzaj: żeński	<i>białej</i> kobiecie
8.	-ego	Stopień: równy; Liczba: pojedyncza; Przypadek: biernik; Rodzaj: męskoosobowy, męskozwierzęcy	<i>białego</i> człowieka, <i>białego</i> konia
9.	-y	Stopień: równy; Liczba: pojedyncza; Przypadek: biernik; Rodzaj: męskonieożywiony	<i>białe</i> zbcze
10.	-ą	Stopień: równy; Liczba: pojedyncza; Przypadek: biernik; Rodzaj: żeński	<i>białą</i> kobietę
11.	-e	Stopień: równy; Liczba: pojedyncza; Przypadek: biernik; Rodzaj: nijaki	<i>białe</i> kaczątko
12.	-ym	Stopień: równy; Liczba: pojedyncza; Przypadek: narzędnik; Rodzaj: męskoosobowy, męskozwierzęcy, męskonieożywiony, nijaki	<i>białym</i> człowiekiem, <i>białym</i> koniem
13.	-ą	Stopień: równy; Liczba: pojedyncza; Przypadek: narzędnik; Rodzaj: żeński	<i>białą</i> kobietą
14.	-ym	Stopień: równy; Liczba: pojedyncza; Przypadek: miejscownik; Rodzaj: męskoosobowy, męskozwierzęcy, męskonieożywiony, nijaki	<i>białym</i> człowieku, <i>białym</i> koniu, <i>białym</i> zbczu, <i>białym</i> kaczątku
15.	-ej	Stopień: równy; Liczba: pojedyncza; Przypadek: miejscownik; Rodzaj: żeński	<i>białej</i> kobiecie
16.	-y	Stopień: równy; Liczba: pojedyncza; Przypadek: wołacz; Rodzaj: męskoosobowy, męskozwierzęcy, męskonieożywiony	o <i>biały</i> człowieku!, o <i>biały</i> koniu!, o <i>biały</i> tumanie!!
17.	-a	Stopień: równy; Liczba: pojedyncza; Przypadek: wołacz; Rodzaj: żeński	o <i>biała</i> kobieto!
18.	-e	Stopień: równy; Liczba: pojedyncza; Przypadek: wołacz; Rodzaj: nijaki	o <i>białe</i> zaśniężone zbcze!
19.	-i	Stopień: równy; Liczba: mnoga; Przypadek: mianownik; Rodzaj: męskoosobowy	<i>biali</i> mężczyźni
20.	-e	Stopień: równy; Liczba: mnoga; Przypadek: mianownik; Rodzaj: męskozwierzęcy, męskonieożywiony, żeński, nijaki	<i>białe</i> niedźwiedzie, konie, domy, kobiety, drzewa
21.	-ych	Stopień: równy; Liczba: mnoga; Przypadek: dopełniacz	<i>białych</i> ludzi, koni, domów, kobiet, dzieci, drzew
22.	-ym	Stopień: równy; Liczba: mnoga; Przypadek: celownik	<i>białym</i> ludziom, koniom, kobietom, dzieciom, drzewom
23.	-ych	Stopień: równy; Liczba: mnoga; Przypadek: biernik; Rodzaj: męskoosobowy	<i>białych</i> ludzi
24.	-e	Stopień: równy; Liczba: mnoga; Przypadek: biernik; Rodzaj: męskozwierzęcy, męskonieożywiony, żeński, nijaki	<i>białe</i> psy, domy, kobiety, dzieci, drzewa
25.	-ymi	Stopień: równy; Liczba: mnoga; Przypadek: narzędnik	<i>białymi</i> ludźmi, kotami, domami, kobietami, dziećmi, drzewami
26.	-ych	Stopień: równy; Liczba: mnoga; Przypadek: miejscownik	<i>białych</i> ludziach, kotach, domach, kobietach, dzieciach,
27.	-i	Stopień: równy; Liczba: mnoga; Przypadek: wołacz; Rodzaj: męskoosobowy	o <i>biali</i> ludzie!
28.	-e	Stopień: równy; Liczba: mnoga; Przypadek: wołacz; Rodzaj: męskozwierzęcy, męskonieożywiony, żeński, nijaki	o <i>białe</i> koty, domy, kobety, dzieci!

W tabeli pominięte zostały rzadko używane<sup>101</sup> formy stopnia wyższego i najwyższego oraz formanty przymiotnika *biały* (pozycje 27 – dla stopnia wyższego, 57 – 84 dla stopnia najwyższego, pozycje 85 i 86 dla formantów).

<sup>101</sup> Por. jednak użycie stopnia wyższego w tytule utworu zespołu Procol Harum „*Bielszy odcień bieli*”.

## SŁOWNIK TERMINÓW

**Anotowanie** – patrz znakowanie lingwistyczne.

**Argument** – wyrażenie uzupełniające opis własności lub relacji, przedmiot tego opisu.

**Błyski** – na ogół krótka forma wypowiedzi poetyckiej właściwa Hartwig, często skondensowana do jednego zdania.

**Czasownik podporowy** – czasownik tworzący ze słowem predykatywnym konstrukcję, w której traci swój podstawowy sens i pełno rolę posiłkową.

**Humanistyka cyfrowa** – w odniesieniu do badań humanistycznych zespół praktyk, które wykorzystują cyfrową (pierwotnie bądź wtórnie) formę przedmiotu badań.

**Idiolekt** – patrz słownik umysłowy człowieka.

**Język indywidualny** – patrz słownik umysłowy człowieka.

**Język osobisty** – patrz słownik umysłowy człowieka.

**Językowy obraz świata** – zespół sądów o świecie utrwalonych w języku, czyli w znaczeniach słów.

**Konkordancja** – uporządkowany zbiór wybranych wyrazów (lub innych jednostek tekstowych) wraz z lewym i prawym kontekstem

**Korpus referencyjny** – zasób danych tekstowych, które cechuje szerokie pokrycie tematyczne i jest reprezentatywny dla dużej liczby zjawisk.

**Korpus** – właściwie zorganizowany zbiór tekstów stanowiący materiał empiryczny do badań lingwistycznych.

**Krzywa nasycenia** – reprezentacja graficzna obrazująca dynamiką pojawiania się nowych form językowych w korpusie.

**Leksykon wirtualny pisarza** – zasób słownictwa, którym dysponuje pisarz.

**Lematyzacja** – procedura określona na tekście, która przypisuje wyrazom tekstu leksemy odpowiadające ich znaczeniom.

**Lingwistyka komputerowa** – dziedzina nauki zajmująca się badaniem i analizą a z wykorzystaniem narzędzi informatycznych oraz tworzeniem elektronicznych narzędzi i zasobów

**Lista frekwencyjna** – lista słów wraz z częstością ich wystąpień, uporządkowana w kolejności liczby wystąpień w teście (od najczęściej do najrzadziej obserwowanych).

- Lista referencyjna** – lista słów utworzona ze względu na określony temat lub cechę języka stanowiąca punkt odniesienia dla prowadzonych badań (w naszym przypadku np. lista form podstawowych przymiotnika).
- Nakładka filtrująca** – modyfikacja działania kodu ograniczająca analizę zbioru danych do określonych jednostek, które spełniają wyznaczone kryterium.
- Nasylenie leksykalne** – jest to zjawisko zatrzymania (lub spowolnienia) przyrostu obserwowanych nowych form językowych w badanym korpusie.
- Otwartość korpusu** – cecha korpusu, który podlega systematycznej aktualizacji.
- Podjęzyk** – odmiana specjalistyczna języka dotycząca określonej dziedziny.
- Predykat** – wyrażenie, które opisuje własność obiektu lub relację między obiektami i jest najważniejszym elementem logiczno-semantycznym w zdaniu elementarnym.
- Przemysł językowy** – przemysłowe wykorzystywanie technologii informatycznych, które generuje produkty i usługi komercyjne związane z językiem naturalnym, stosowane w różnego typu aplikacjach.
- Przetwarzanie języka naturalnego** – dziedzina lingwistyki komputerowej mająca za przedmiot analizę różnych aspektów języka naturalnego.
- Ranga** – wskazuje na pozycję na liście frekwencyjnej.
- Referencja** – odesłanie do konkretnego obiektu lub zjawiska w świecie.
- Reprezentatywność korpusu** – cecha przysługująca korpusowi, który odzwierciedla określoną rzeczywistość językową i reprezentuje język ze względu na jego właściwości gramatyczne i leksykalne.
- Słownictwo wirtualne** – patrz leksykon wirtualny pisarza
- Słownik umysłowy człowieka** – indywidualny zasób słownictwa, którym dysponuje użytkownik języka.
- Struktura predykatywno-argumentowa** – model zdania oparty na predykanie, który orzeka o własnościach danego elementu (argumentu) lub o relacji między elementami (argumentami).
- Walencja** – własność wyrażenia do otwierania pozycji składniowo-semantycznych dla innych wyrażen.
- Wartościowość syntaktyczna** – patrz walencja.
- Warsztat pisarski** – sposób posługiwania się przez pisarza językiem w twórczości literackiej.
- Zasoby językowe** – zgromadzone zbiory tekstów pisanych lub mówionych.

**Znakowanie lingwistyczne** – oznaczanie słów w korpusie znakami ze względu na formę gramatyczną, właściwości semantyczne, prozodyczne, itd.

**Zrównoważenie korpusu** – dobór tekstów w korpusie gwarantujący jego reprezentatywność ze względu na zjawiska będące przedmiotem badań.

## INDEKS

- Ajdukiewicz, K., 24  
anotowanie, 70, 71, 133  
Apollinaire, G., 9, 12, 34, 37, 38, 56, 85  
Aragon, L., 9, 12  
argument, 25, 27, 28, 29, 30, 31, 35, 133, 134  
Auderska, H., 33, 146  
Bachan, J., 81  
Bally, C., 16  
Bańko, M., 35, 54, 142, 145, 146  
Baran, M., 21, 140  
Barańczak, S., 22, 140  
Bartmiński, J., 20, 39, 140, 147  
Bauvoire, S. de, 12  
Bellert, I., 24, 140  
Bem, M., 60, 140  
Bień, J., 53, 140  
Bird, S., 73, 140  
błyski, 58, 68, 114, 115, 133, 140  
Bogacki, K., 45, 140, 148  
Brzeziński, J., 16, 140, 145, 148  
Bühler, K., L., 33  
Buttler, D., 32  
Canbolat, U., 46, 74, 105, 148  
Cankudis, A., 19, 142  
Cendras, B., 9  
Chomsky, N., 20, 24, 26, 141  
Choukri, K., 46, 147, 148  
Colmerauer, A., 24, 140  
Corneille, P., 113  
Cruden, A., 45, 140  
czasownik podporowy, 35, 133  
Czerska, T., 43, 140  
Dahl, V., 24, 141  
Dąbrowski, A., 45, 51, 54, 55  
denotacja, 40  
Dickinson, E., 10  
Doroszewski, W., 16, 34, 39, 45, 71, 141  
Dutka, E., 21, 22, 43, 59, 141  
Edelman, M., 12  
Eder, M., 17, 20, 141  
Fillmor, 24  
Flakowicz, M., 41, 141  
Flakowicz-Szczyrba, M., 21, 43, 141  
Fliciński, P., 16, 141  
Gajcy, T., 17  
Gajda, S., 16, 141  
Gałczyński, K., I., 12  
Geremek, B., 12  
Ginsberg, A., 10  
Górski, R., L., 47, 54, 142, 145, 146  
Grochowski, M., 24  
Grochowski, M., 141  
Gruszczyński, W., 32, 94, 146  
Grzegorzczkowa, R., 23, 33, 39, 40, 113, 128, 141  
Guiraud, P., 21, 141  
Habert, B., 61  
Hajnicz, E., 44, 49, 59, 142  
Hartwig, J., 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 20, 21, 22, 23, 25, 28, 31, 32, 33, 34, 35, 36, 37, 39, 40, 41, 42, 43, 44, 48, 54, 55, 57, 59, 60, 62, 63, 64, 65, 66, 70, 71, 72, 73, 74, 75, 80, 82, 85, 90, 91, 92, 100, 105, 107, 109, 112, 113, 114, 122, 123, 124, 125, 126, 127, 133, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 150, 151  
Heinz, A., 33, 142  
Herbert, Z., 12, 60  
Herbertowie, K. i Z., 11, 142  
Hojdis, B., 19, 142  
Hull, L., 21, 142  
humanistyka cyfrowa, 14, 18, 19, 124, 133, 142  
idiolekt, 133  
Iwazkiewicz, J., 12  
Jacob, M., 9  
Jacobs, A., M., 63, 142  
Jacobson, R., 20  
Jędrzejko, E., 26, 29, 30, 142  
język indywidualny, 15, 16, 17, 18, 23, 28, 133  
język osobniczy, 16, 17, 143, 145  
językowy obraz świata, 39, 133, 140, 141  
Jockers, M., 18, 142  
Kaliska, A., 26, 31, 35, 142

Karolak, S., 24, 25, 31, 33, 35, 142, 143  
 kategoryzacja, 43  
 Kennedy, G., 44, 45, 48, 52, 143  
 Kielar, M., B., 22, 143  
 Kieraś, W., 73, 143  
 Kilgariff, A., 52, 143  
 Kittredge, R., 49, 109  
 Klein, E., 73, 140  
 Klemensiewicz, Z., 16, 17, 143  
 Kobyliński, Ł., 73, 143  
 kolokacje werbo-nominalne, 45, 147  
 Kołakowski, L., 12  
 kompetencja językowa, 5, 46, 112, 113, 147  
 konkordancja, 4, 15, 45, 71, 89, 92, 94, 97,  
 100, 122, 123, 133, 150  
 konotacja, 17, 33, 85  
 korpus, 3, 4, 6, 13, 14, 15, 18, 19, 25, 37,  
 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54,  
 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65,  
 66, 67, 68, 69, 71, 72, 73, 74, 75, 76, 77,  
 79, 80, 81, 82, 83, 85, 86, 90, 91, 92, 94,  
 95, 97, 100, 105, 107, 109, 113, 122,  
 123, 124, 125, 126, 127, 133, 134, 135,  
 140, 142, 144, 145, 146, 147, 150  
 korpus monitorujący, 49, 50, 145  
 korpus referencyjny, 45, 46, 50, 52, 70, 133  
 Kosowska, I., 21, 22, 40, 143  
 Kozłowska, A., 16, 143  
 Kreisberg, A., 32  
 krzywa nasycenia, 109, 133  
 Kubacki, P., 45, 51, 54, 55  
 Kubis, M., 20, 45, 51, 54, 55, 64, 147, 148  
 Kuczyńska-Koschany, K., 21, 143  
 Kurcz, I., 20, 45, 48, 52, 53, 113, 143, 144  
 Kurkiewicz, J., 43, 142  
 Kurkowska, H., 35, 144  
 Kwaśnicka-Janowicz, A., 52, 61, 145  
 Kwiatkowski, J., 21, 41, 144  
 Legeżyńska, A., 13, 22, 42, 140, 144  
 leksykon wirtualny pisarza, 112, 133, 134  
 lematyzacja, 4, 72, 75, 76, 77, 78, 79, 80,  
 95, 131  
 Leńczuk, M., 48, 144  
 Less- Less-Resourced Languages /LRL/, 46  
 Lewandowska-Tomaszczyk, B., 54, 142,  
 145, 146  
 Lewicki, A., 45, 48, 52, 53, 143, 144  
 lingwistyka komputerowa, 1, 13, 18, 19, 20,  
 70, 133, 134  
 lista frekwencyjna, 14, 15, 45, 53, 71, 81,  
 83, 85, 86, 126, 133  
 lista referencyjna, 4, 14, 33, 71, 73, 74, 75,  
 79, 80, 81, 134  
 Looper, E., 73, 140  
 Lubaś, W., 16, 144  
 Lutosławski, W., 17, 144  
 Łaziński, M., 47, 142  
 Łempicka, Z., 33, 146  
 Łuszczkiewicz, P., 21, 144  
 Marciniak, J., 45, 51, 54, 55  
 Marek, W., 28  
 Mariani, J., 46, 147, 148  
 Martinek, J., 72, 112, 128, 130, 147  
 Maryl, M., 18, 144  
 Mazowiecki, T., 12  
 Menken, S., 59, 74, 147, 148  
 Merwin, W., S., 13  
 metody ilościowe, 3, 21  
 Mędak, K., 71, 144  
 Michaux, H., 9  
 Michnik, A., 12  
 Międzyrzecki, A., 10, 11, 12, 142  
 Miller, G., A., 20, 144  
 Miłosz, C., 10, 11, 12, 22, 60, 140, 142, 144  
 Montague, R., 24, 144  
 Moretti, F., 18, 55, 145  
 Muller, C., 109, 112, 113, 145, 147  
 Nagórko, A., 32, 145  
 nakładka filtrująca, 76, 134  
 Nasiłowska, A., 21, 145  
 nasycenie leksykalne, 134, 151  
 Nelson, F., 51  
 Nerval, G. de, 9, 12, 42, 56, 67, 85, 114,  
 139  
 Nowakowska-Kempna, I., 32  
 Obrębski, T., 13, 20, 45, 51, 54, 55, 72, 75,  
 112, 128, 130, 131, 147  
 Ogrodniczuk, M., 50, 73, 143  
 Osiński, J., 45, 51, 54, 55  
 otwartość korpusu, 47, 49, 134  
 Pajdzińska, A., 17, 20, 39, 145  
 Pałka, P., 52, 61, 145  
 Paroubek, P., 46, 148  
 Partee, B., 24, 145  
 Pawłowski, A., 21, 145

Pęzik, P., 50, 145  
 Piasecki, M., 20  
 Picasso, P., 37  
 Piotrowski, T., 61  
 podjęzyki, 49  
 Polański, K., 26, 48, 71, 145  
 predykat, 3, 4, 15, 25, 26, 27, 28, 29, 30,  
 31, 32, 33, 34, 35, 36, 37, 38, 70, 94, 97,  
 122, 133, 134, 141, 142, 146, 147, 150  
 Przepiórkowski, A., 45, 53, 54, 66, 70, 94,  
 142, 145, 146  
 przetwarzanie języka naturalnego, 4, 15, 20,  
 70, 109, 134  
 Przybylski, R., K., 23, 40, 145  
 Puzynina, J., 16, 32, 145  
 Racine, J., B., 113  
 ranga, 82, 134  
 referencja, 40, 134  
 reprezentatywność korpusu, 3, 47, 48, 49,  
 59, 134, 135, 142  
 Reverdey, P., 9  
 Rimbaud, A., 9  
 Rocławski, B., 53  
 Rudnicka, E., 20  
 Rudnicka-Fira, E., 17, 146  
 Rybicki, J., 17, 20, 48, 55, 61, 146  
 Saloni, Z., 32, 71, 94, 146  
 Sambor, J., 17, 45, 48, 52, 53, 140, 143,  
 144, 146  
 Sartre, P., 12  
 Skibski, K., 23, 38, 39, 146  
 Skorupka, S., 33, 146  
 Skowrońska, D., 32, 94, 146  
 Sławkowa, E., 17, 146  
 Słonimski, A., 12  
 słownictwo wirtualne, 112, 134  
 słownik umysłowy człowieka, 20, 133  
 Stachurski, E., 17, 146  
 Stankowska, A., 21, 146  
 struktura predykatywno-argumentowa, 3,  
 15, 24, 28, 33, 35, 36, 38, 134  
 Szafran, K., 48, 53, 143  
 Szałkiewicz, Ł., 94, 146  
 Szaruga, L., 21, 146  
 Szumska, D., 25, 26, 28, 32, 33, 36, 38, 146  
 Szupryczyńska, M., 32, 33, 146  
 Szymborska, W., 11, 12  
 Szymczak, M., 32, 45, 47, 71, 75, 146  
 Szymczkowa, M., 48  
 Śliwiński, W., 26, 32, 33, 146  
 Świdziński, M., 24, 146  
 Tarski, A., 24, 147  
 Taylor, J., R., 20, 147  
 Telicki, M., 1, 21, 22, 40, 41, 140, 147  
 Tokarski, R., 17, 20, 32, 39, 71, 145, 147  
 Topolińska, Z., 25, 142, 147  
 Turowicz, J., 12, 60  
 Turowiczowie, A. i J., 10, 142  
 Vetulani, G., 13, 25, 26, 28, 35, 36, 45, 46,  
 51, 54, 55, 59, 61, 66, 72, 75, 97, 122,  
 128, 130, 131, 147, 148  
 Vetulani, Z., 13, 20, 24, 28, 29, 44, 45, 46,  
 51, 54, 55, 59, 64, 72, 74, 75, 105, 109,  
 112, 128, 130, 131, 147, 148  
 Walczak, B., 13, 16, 45, 72, 75, 128, 131,  
 147, 148  
 walencja, 35, 134  
 Walker, D., 19  
 Walkowska, J., 45, 46, 51, 54, 55, 148  
 warsztat pisarski, 1, 3, 6, 10, 11, 13, 14, 15,  
 16, 20, 21, 23, 25, 32, 33, 36, 39, 40, 48,  
 70, 71, 74, 89, 100, 109, 112, 123, 124,  
 127, 134  
 Węgrzynek, K., 37, 148  
 Whitman, W., 10  
 Wilkoń, A., 16, 148  
 Williams, W., C., 10  
 Winiecka, E., 22, 42, 143, 144, 145, 146,  
 148  
 Wiśniewski, J., 40, 148  
 Witalewski, K., 45, 51, 54, 55  
 Witkowska, M., 1, 46, 59, 64, 74, 105, 147,  
 148  
 Witosz, B., 45  
 Wojtasiewicz, O., A., 37, 149  
 Woliński, M., 32, 53, 94, 140, 146  
 Wołosz, R., 32, 94, 146  
 Woronczak, J., 45, 48, 52, 53, 143, 144  
 Wyka, K., 17, 149  
 Zampolli, A., 19  
 zasoby językowe, 19, 52, 79, 134, 145  
 Zgólkowie, H. i T., 17, 149  
 Zipf, G., 21, 149  
 złożoność korpusu, 4, 66  
 znakowanie lingwistyczne, 70, 133, 135

# BIBLIOGRAFIA

## DZIEŁA JULII HARTWIG (pierwsze wydania)

### PROZA:

- Hartwig J., 1954: *Z niedalekich podróży*, Ludowa Spółdzielnia Wydawnicza, Warszawa.
- Hartwig J., 1962 *Apollinaire*, Państwowy Instytut Wydawniczy, Warszawa.
- Hartwig J., 1972: *Gérard de Nerval*, Państwowy Instytut Wydawniczy, Warszawa.
- Hartwig J., 1980: *Dziennik amerykański*, Zeszyty Literackie, Warszawa.
- Hartwig J., 2001: *Zawsze powroty. Z dzienników podróży*, Sic! Warszawa.
- Hartwig J., 2004: *Pisane przy oknie*, Biblioteka Więzi, Warszawa.
- Hartwig J., 2006: *Podziękowanie za gościnę. Moja Francja*, Słowo/obraz terytoria, Gdańsk.
- Hartwig J., 2006: *Wybrańcy losu*, Sic! Warszawa.
- Hartwig J., 2011: *Dziennik*, Wydawnictwo Literackie, Kraków.
- Hartwig J., 2014: *Dziennik tom 2*, Wydawnictwo Literackie, Kraków.

### TOMY POETYCKIE:

- Hartwig J., 1956: *Pożegnania*, Czytelnik, Warszawa.
- Hartwig J., 1969: *Wolne ręce*, Państwowy Instytut Wydawniczy, Warszawa.
- Hartwig J., 1971: *Dwoistość*, Czytelnik, Warszawa.
- Hartwig J., 1978: *Czuwanie*, Czytelnik, Warszawa.
- Hartwig J., 1980: *Chwila postoju*, Wydawnictwo Literackie, Kraków.
- Hartwig J., 1987: *Obcowanie*, Czytelnik, Warszawa.
- Hartwig J., 1992: *Czułość*, Znak, Kraków.
- Hartwig J., 1995: *Nim opatrzy się zieleń. Wybór wierszy*, Znak, Kraków.
- Hartwig J., 1999: *Zobaczone a5*, Kraków.
- Hartwig J., 2001: *Nie ma odpowiedzi*, Sic! Warszawa.
- Hartwig J., 2002: *Wiersze amerykańskie*, Sic! Warszawa.
- Hartwig J., 2002: *Błyski*, Sic! Warszawa.
- Hartwig J., 2003: *Mówiąc nie tylko do siebie. Poematy prozą*, Sic! Warszawa.
- Hartwig J., 2004: *Bez pożegnania*, Sic! Warszawa.
- Hartwig J., 2004: *Zwierzenia i błyski*, Sic! Warszawa.
- Hartwig J., 2007: *To wróci*, Sic! Warszawa.
- Hartwig J., 2008: *Trzecie błyski*, Sic! Warszawa.
- Hartwig J., 2009: *Jasne niejasne*, a5, Kraków.
- Hartwig J., 2011: *Gorzkie żale*, a5, Kraków.
- Hartwig J., 2013: *Zapisane*, a5, Kraków.
- Hartwig J., 2014: *Błyski zebrane*, Zeszyty Literackie, Warszawa.
- Hartwig J., 2016: *Spojrzenie*, a5, Kraków.

## REFERENCJE

- Baran, M. (2017): Wobec genologii. Gatunek indywidualny Julii Hartwig. [w:] *Annales Universitatis Paedagogicae Cracoviensis, Studia Poetica*, 5, s. 203–219, <https://doi.org/10.24917/23534583.5.16> [dostęp 15.07.2021].
- Barańczak, S. (2000): O tomie „Zobaczone” Julii Hartwig. [w:] *Zeszyty Literackie*, 1(69). Warszawa – Paryż – Mediolan, s. 107–110.
- Bartmiński, J. (red.) (1990): *Językowy obraz świata*. Wyd. UMCS, Lublin.
- Bartmiński, J. (2010): Pojęcie językowego obrazu świata i sposoby jego operacjonalizacji. [w:] *Jaka antropologia literatury jest dzisiaj możliwa*, Czapliński, P., Legeżyńska, A., Telicki, M. (red.). Wyd. „Poznańskie Studia Polonistyczne”, Poznań, s. 155–178.
- Bellert, I. (1989): *Feature system for Quantification structures in natural language*. Foris Publication, Dordrecht.
- Bem, P. (2017): *Dynamika wariantu. Miłosz tekstologicznie*. Wyd. IBL PAN, Warszawa.
- Bień, J.S., Woliński, M. (2003): Wzbogacony korpus „Słownika frekwencyjnego polszczyzny współczesnej”. [w:] *Prace językoznawcze dedykowane Profesor Jadwidze Sambor*, Linde-Usiekiewicz, J., Huszcza, R. (red.). Wydział Polonistyki Uniwersytetu Warszawskiego, Warszawa, s. 6-10.
- Bird S., Klein E., Looper E. (2009): *Natural Language Processing with Python*. O’Reilly Media, Inc., Beijing, Cambridge, Farnham, Köln, Sebastopol, Taipei, Tokyo.
- Brzeziński, J. (1975): *Język Franciszka Dionizego Książnika*. Wyd. Uczelniane WSP, Zielona Góra.
- Bogacki, K. (1997): POLLEX - un dictionnaire électronique du polonaise. [w:] *BULAG, Actes du Colloque International FRACTAL*. Grenoble, s. 55–64.
- Colmerauer, A. (1978): Metamorphosis Grammars [w:] *Natural Language Communication with Computers, Lecture Notes in Computer Science* 63, Bolc L. (red.). Springer, New York.
- Cruden, A. (1769): *Complete Concordance to the Old and New Testaments*. Lutterworth Press, Londyn.
- Czerska, T. (2014): Starość (nie)wypowiedziana w późnej twórczości Julii Hartwig. [w:] *Autobiografia* 1(2), s. 63–174.
- Chomsky, N. (1957): *Syntactic Structures*. The Hague: Mouton.
- Dahl, V. (1981): Translating from Spanish into Logic through Logic. [w:] *American Journal of Computational Linguistics*, 3(7), s. 149–164, <https://aclanthology.org/J81-3002> [20.07.2021].
- Doroszewski, W. (1949): *Język Teodora Tomasza Jeża (Zygmunta Miłkowskiego). Studium z dziejów języka polskiego XIX wieku*. Towarzystwo Naukowe Warszawskie, Warszawa.
- Doroszewski, W. (red.), (1958-1969): *Słownik języka polskiego*, t.1-11. Wyd. Naukowe PWN, Warszawa.
- Doroszewski, W. (1972): O atrybutywnych i predykatywnych funkcjach przymiotników. [w:] *Poradnik Językowy*, Zeszyt 7(301), s. 399-401.

- Dutka, E. (2016): *Centra, prowincje, zaułki. Twórczość Julii Hartwig jako auto/bio/geo/grafia*. Wyd. Universitas, Kraków.
- Eder, M. (2014): Metody ściśle w literaturoznawstwie i pułapki pozornego obiektywizmu – przykład stylometrii. [w:] *Teksty Drugie*, 2, s. 90-105.
- Eder, M. (2016): Słowa znaczące, słowa kluczowe, słowozbiory – o statystycznych metodach wyszukiwania wyrazów istotnych. [w:] *Przegląd Humanistyczny*, 3 (454), s. 31-44.
- Flakowicz, M. (2009): Graniczność jako wybór gatunkowy i światopoglądowy – poematy prozą Julii Hartwig, [w:] *Literaturoznawstwo: historia, teoria, metodologia, krytyka*, 1(3). Łódź, s. 201-215.
- Flakowicz-Szczyrba, M. (2012): Poezja Julii Hartwig wobec egzystencji i metafizyki [w:] *Pamiętnik Literacki*, 2(103), s. 59-100.
- Flakowicz-Szczyrba, M. (2014): *Dowód na istnienie. Poezja Julii Hartwig wobec egzystencji i metafizyki*. Wyd. IBL PAN, Warszawa.
- Fliciński, P. (2004): Idiostyl pisarza jako problem stylistyki. [w:] *Studia nad polszczyzną współczesną i historyczną, Prace dedykowane Profesorowi Stanisławowi Bąbie w 65-lecie urodzin*, Liberek J. (red.). Wyd. „Poznańskie Studia Polonistyczne”, Poznań, s. 95-108.
- Gajda, S. (2013): Teoria stylu i stylistyka [w:] *Przewodnik po stylistyce polskiej. Style współczesnej polszczyzny*, Malinowska E., Nicoń J., Żydek-Bednarczyk U. (red.). Wyd. Universitas, Kraków, s. 15-33.
- Guiraud, P. (1954): *Les caractères statistiques du vocabulaire*. PUF, Paris.
- Grochowski, M. (1997): *Wyrażenia funkcyjne. Studium leksykograficzne*. Prace IJP PAN, Kraków.
- Grzegorzczkova, R. (2001): *Wprowadzenie do semantyki językoznawczej*. Wyd. Naukowe PWN, Warszawa.
- Grzegorzczkova, R. (2012): *Wstęp do językoznawstwa*. Wyd. Naukowe PWN, Warszawa.
- Grzegorzczkova, R. (2015): Co to jest językowy obraz świata i dlaczego warto go badać? [w:] *Poradnik Językowy*, 1, s. 7-13.
- Grzegorzczkova, R., Laskowski, R., Wróbel, H.(red.) (1998): *Gramatyka współczesnego języka polskiego. Morfologia*. Wyd. Naukowe PWN, Warszawa.
- Górski, R. L., Łaziński, M. (2012): Reprezentatywność i zrównoważenie korpusu.[w:] *Narodowy Korpus Języka Polskiego*, Przepiórkowski, A., Bańko, M., Górski, R. L., Lewandowska-Tomaszczyk, B. (red.).Warszawa, s. 25-36.
- Hajnicz, E. (2011): *Najbardziej znane korpusy tekstów. Opracowanie przeglądowe*. Prace IPI PAN, 1021, Warszawa.
- Hartwig, J. (2010): Starość nie jest doklejoną do życia wyspą. Rozmowa J. Kurkiewicza z J. Hartwig. [w:] *Gazeta wyborcza*, 169  
[http://wyborcza.pl/1,106967,8163295,Julia\\_Hartwig\\_Starosc\\_nie\\_jest\\_doklejona\\_do\\_zycia.html#ixzz15LbRUXLW](http://wyborcza.pl/1,106967,8163295,Julia_Hartwig_Starosc_nie_jest_doklejona_do_zycia.html#ixzz15LbRUXLW) [dostęp 30.04.2021].
- Hartwig, J. (2012): *Wspólna obecność. Korespondencja Julii Hartwig i Artura Międzyrzeckiego z Anną i Jerzym Turowiczami*. Wyd. a5, Kraków.

- Hartwig, J. (2012): *Julia Hartwig, Artur Międzyrzecki, Czesław Miłosz. Korespondencja*. Wydawnictwo Literackie, Kraków
- Hartwig, J. (2013): *Julia Hartwig, Artur Międzyrzecki – Katarzyna i Zbigniew Herbertowie*. Fundacja Zeszytów Literackich, Warszawa.
- Hartwig, J. (2015): Doctoris Honoris Causa Lectio. [w:] *Julia Hartwig. Doctor Honoris Causa Universitatis Studiorum Mickiewicziana Poznańensis*. Wyd. Naukowe UAM, Poznań 2015, s. 31-40.
- Heinz, A. (1965): *System przypadkowy języka polskiego*. Nakład Uniwersytetu Jagiellońskiego, Kraków.
- Hojdis, B., Cankudis, A. (2020): Cyfrowe edycje literackie i naukowe jako element cyfrowej humanistyki. [w:] *Sztuka Edycji*, 1, s. 7–15, DOI: <http://dx.doi.org/10.12775/SE.2020.0001> [dostęp 18.02.2021].
- Hull, L. (2010): „Bądź wdzięczna, byłaś hojnie obdarzona” – Julia Hartwig. [w:] Hull, L. *Obecność nieprzynależna. Pokoleniowe wydziedziczenie w poezji po 1989 roku*. Uniwersytet Warmińsko-Mazurski, Olsztyn.
- Jacobs, A.M. (2018): The Gutenberg English Poetry Corpus: Exemplary Quantitative Narrative Analyses. [w:] *Frontiers in Digital Humanities*, 5:5, s. 1-14. DOI: 10.3389/fdigh.2018.00005, [10.07.2018].
- Jędrzejko, E. (1988): *Słownik polskich zwrotów werbo-nominalnych. Zeszyt próbny*. Energeia, Warszawa.
- Jędrzejko, E. (1993): *Nominalizacje w systemie i w tekstach współczesnej polszczyzny*. Uniwersytet Śląski, Katowice.
- Jockers, M. (2013): *Macroanalysis. Digital Method and Literary History*. University of Illinois Press, Champaign.
- Kaliska, A. K. (2016): *Wykładowi predykcji w języku polskim i francuskim*. Poznań, [https://repozytorium.amu.edu.pl/bitstream/10593/17415/1/Kaliska\\_9788394601713.pdf](https://repozytorium.amu.edu.pl/bitstream/10593/17415/1/Kaliska_9788394601713.pdf) [dostęp 10.07.2021]
- Kaliska, A. K. (2018): *Wykładowi instrumentalno-sposobowe a struktura semantyczna zdania. Polsko-francuska perspektywa porównawcza*. Wyd. Naukowe UAM, Poznań.
- Karolak, S. (1984): Składnia wyrażen predykatywnych [w:] *Gramatyka współczesnego języka polskiego. Składnia*, Topolińska Z. (red). Wyd. Naukowe PWN, Warszawa, s. 11-210.
- Karolak, S. (2001): *Od gramatyki do semantyki*. Sławistyczny Ośrodek Wydawniczy, Warszawa-Kraków.
- Kennedy, G. (1998): *An Introduction to Corpus Linguistics*. Addison Wesley Lingman Ltd, Londyn, Nowy Jork.
- Kielar, M. B. (2000): O tomie „Zobaczone” Julii Hartwig. [w:] *Zeszyty Literackie*, 1(69). Warszawa – Paryż – Mediolan, s. 111-112.
- Kieraś, W., Kobylński, Ł., Ogrodniczuk, M. (2018): Korpusomat — a tool for creating searchable morphosyntactically tagged corpora. [w:] *Computational Methods in Science and Technology*, 24(1), s. 21–27.
- Kilgariff, A. (2005): Language is never, ever, ever random. [w:] *Corpus Linguistics and Linguistic Theory*, 1/2, s. 263–276.

- Klemensiewicz, Z. (1961): Jak charakteryzować język osobniczy? [w:] Klemensiewicz, Z., *W kręgu języka artystycznego i literackiego*, Warszawa, s. 204–214 [artykuł z 1946 roku].
- Kosowska, I. (2014): Próba zadomowienia. Francja w twórczości Julii Hartwig [w:] *Ruch Literacki*, R. LV, Z. 2 (323), s. 217-234.  
<https://journals.pan.pl/dlibra/publication/96259/edition/83006/content>  
[dostęp 20.07.2021]
- Kozłowska, A. (2015). Miejsce badań nad idiolektem w obrębie językoznawstwa. [w:] *Poznańskie Spotkania Językoznawcze*, 30, s. 71–83.  
<https://doi.org/10.14746/psj.2015.30.4> [dostęp 20.07.2021]
- Kučera, H., Francis, W. N. (1967): *Computational Analysis of Present-day American English*, Brown University Press, Providence, RI.
- Kuczyńska-Koschany, K. (2015): Francja Julii Hartwig: inspiracja i esprit. [w:] *Pochwała istnienia. Studia o twórczości Julii Hartwig*, Kulesza-Gulczyńska B., Winiecka E. (red.). Bogucki Wyd. Naukowe, Poznań, s. 63-74.
- Kulesza-Gulczyńska, B., Winiecka, E. (red.) (2015): *Pochwała istnienia. Studia o twórczości Julii Hartwig*. Bogucki Wyd. Naukowe, Poznań.
- Kurcz, I., Lewicki, A., Sambor, J. i Woronczak, J. (1974a): *Słownictwo współczesnego języka polskiego. Listy frekwencyjne. Tom I. Teksty popularnonaukowe*. IJP PAN, Warszawa.
- Kurcz, I., Lewicki, A., Sambor, J. i Woronczak, J. (1974b): *Słownictwo współczesnego języka polskiego. Listy frekwencyjne. Tom II. Drobne wiadomości prasowe*. IJP PAN, Warszawa.
- Kurcz, I., Lewicki, A., Sambor, J. i Woronczak, J. (1976): *Słownictwo współczesnego języka polskiego. Listy frekwencyjne. Tom IV. Proza artystyczna*. IJP PAN, Warszawa.
- Kurcz, I. (1976): *Psycholingwistyka. Przegląd problemów badawczych*. IJP PAN, Warszawa.
- Kurcz, I., Lewicki, A., Sambor, J. i Woronczak J. (1977): *Słownictwo współczesnego języka polskiego. Listy frekwencyjne. Tom V. Dramat artystyczny*. IJP PAN, Warszawa.
- Kurcz, I., Lewicki, A., Sambor, J., Szafran, K. i Woronczak J. (1990): *Słownik frekwencyjny polszczyzny współczesnej*, t. I, t. II. IJP PAN Kraków.
- Kurcz, I. (1976): *Psycholingwistyka. Przegląd problemów badawczych*. Wyd. Naukowe PWN, Warszawa.
- Kurcz, I. (2005): *Psychologia języka i komunikacji*. Wyd. Naukowe „Scholar” Warszawa.
- Kurkowska, H. (1953): *Budowa słowotwórcza przymiotników polskich*. Zakład im. Ossolińskich, Wrocław.
- Kwiatkowski, J. (1973): Felieton poetycki [w:] *Twórczość*, 9, s. 124.
- Kwiatkowski, J. (1995). Dwie poezje Julii Hartwig. [w:] Kwiatkowski J., *Magia poezji*. Wydawnictwo Literackie, Kraków, s. 369-373.
- Legeżyńska, A. (2015): Gdyby Czesław Miłosz był kobietą... [w:] *Pochwała istnienia. Studia o twórczości Julii Hartwig*, Kulesza-Gulczyńska, B., Winiecka, E. (red.). Bogucki Wyd. Naukowe, Poznań, s. 19-31.

- Legeżyńska, A. (2017): *Julia Hartwig. Wdzięczność*. Wyd. Uniwersytetu Łódzkiego, Łódź.
- Leńczuk, M. (2020): Projektowanie metadanych w korpusie tekstów polskich do 1500 roku – wielopoziomowa struktura informacji. [w:] *Forum Lingwistyczne*, 7, s. 59-69, DOI: <http://doi.org/10.31261/FL.2020.07.05> [dostęp 20.07.2021]
- Lewicki, A. M., Masłowski, W., Sambor, J. i Woronczak, J. (1975). *Słownictwo współczesnego języka polskiego. Listy frekwencyjne. Tom III. Publicystyka*. IJP PAN, Warszawa.
- Lubaś, W. (1992): *Osobliwości językowe poezji Ignacego Krasickiego*. IJP PAP, Kraków.
- Lutosławski, W. (1898): Principes de stylométrie appliqués à la chronologie des oeuvres de Platon. [w:] *Revue des Études Grecques*, 11(41). Paris, s. 61-81.
- Łuszczkiewicz, P. (1998): Być poetą osobnym. [w:] *Nowe Książki*, 11, s. 4.
- Marciszewski, W. (red.) (1970): *Mała Encyklopedia Logiki*. Zakład im. Ossolińskich, Wrocław, s. 229.
- Maryl, M. (2016): Bardziej racjonalna historia literatury. [w:] *Nowe Książki*, 7-8, s. 20-21.
- Maryl, M. (2017): Kim są współcześni humaniści cyfrowi? [w:] *Teksty Drugie*, 1 (163), s. 286-300.
- Mędak, S. (1997): *Słownik form koniugacyjnych czasowników polskich*. Wyd. Universitas, Kraków.
- Maziarz, M., Piasecki, M., Rudnicka, E. (2014): *Słowosieć - polski wordnet. Proces tworzenia tezaury*. [w:] *Polonica*, 34, s. 79-98.
- Miłosz, Cz. (2000): O tomie „Zobaczone” Julii Hartwig. [w:] *Zeszyty Literackie*, 1(69). Warszawa – Paryż – Mediolan, s. 111-112.
- Miller G.A., Beckwith R., Fellbaum Ch., Gross D., Miller K. J. (1990): Introduction to WordNet: An On-line Lexical Database [w:] *International Journal of Lexicography*, 3. Oxford, s. 235-244.
- Montague, R. (1970): English as Formal Language. [w:] Vinsentini B. i in., *Linguaggi nella Societa e nella Technica*, przedruk Thomanson R.H. (red.), (1974): *Formal Philosophy. Selected Papers*, Yale University Press.
- Montague, R. (1970): Universal Grammar. [w:] „Theoria”, 36, s. 373-398, przedruk Thomanson R.H. (red.), (1974): *Formal Philosophy. Selected Papers*, Yale University Press.
- Moretti, F. (2016): *Wykresy, drzewa, ciągi. Abstrakcyjne modele na potrzeby historii literatury*. Wyd. Uniwersytetu Jagiellońskiego, Kraków.
- Moretti, F. (2017): Literatura – zmierzona. [w:] *Teksty Drugie*, 1 (163), s. 210-219.
- Moretti, F. (2018): Wzorce a interpretacja. [w:] *Rocznik Komparatystyczny-Comparative Yearbook*, 9, s. 9-24, <https://wnus.edu.pl/rk/pl/issue/966/article/15727/> [dostęp 20.06.2021].
- Muller, Ch. (1979): *Langue française et linguistique quantitative. Receueil d'article*, Editions Slatkine, Genève.
- Nasiłowska, A. (2011): Epoka Julii Hartwig. [w:] *Znak*, 7/8, s. 99, <https://www.miesiecznik.znak.com.pl/674-675201anna-nasilowskaepoka-julii-hartwig/> [dostęp 15.06.2021]

- Nagórko, A. (1987): *Zagadnienia derywacji przymiotników*. Wyd. Uniwersytetu Warszawskiego, Warszawa.
- Murray, J.A.H., Bradley, H., Craigie, W.A., Onions, C.T. (red.), (1928): *Oxford English Dictionary*. Oxford University Press, Oxford.
- Pajdzińska, A., Tokarski, R. (red.), (2001): *Semantyka tekstu artystycznego*. Wyd. UMCS, Lublin.
- Pajdzińska, A., Tokarski, R. (2004): *Punkty widzenia w tekście i w dyskursie*. Wyd. UMCS, Lublin.
- Pałka, P., Kwaśnicka-Janowicz, A. (2017): *Przewodnik po elektronicznych zasobach językowych dla polonistów (słowniki, kartoteki, korpusy, kompendia)*. Wyd. JAK, Kraków. [http://www.tmjp.pl/images/pdf/przewodnik\\_po\\_elektronicznych\\_zasobach.pdf](http://www.tmjp.pl/images/pdf/przewodnik_po_elektronicznych_zasobach.pdf) [dostęp 15.07.2021]
- Partee, B. H., Meulen, A., Wall, R. E. (1990): *Mathematical methods in linguistics*. Kluwer Academic Publishers, Dordrecht.
- Pawłowski, A. (2001): *Metody kwantytatywne w sekwencyjnej analizie tekstu*. Katedra Lingwistyki Formalnej Uniwersytetu Warszawskiego, Warszawa.
- Pęzik, P. (2020): Budowa i zastosowania korpusu monitorującego MoncoPL. [w:] *Forum lingwistyczne*, 7, s. 133-150, DOI: <http://doi.org/10.31261/FL>. [dostęp 11.07.2021]
- Polański, K. (red.), (1980): *Syntaktyczno-generatywny słownik czasowników polskich*. Zakład im. Ossolińskich, Wrocław-Warszawa-Kraków-Gdańsk.
- Polański, K. (red.) (1995): *Encyklopedia językoznawstwa ogólnego*. Zakład im. Ossolińskich, Wrocław-Warszawa-Kraków, s. 29-30, 290
- Przepiórkowski, A. (2004): *Korpus IPI PAN. Wersja wstępna*. IPI PAN, Warszawa.
- Przepiórkowski, A., Bańko, M., Górski, R. L., Lewandowska-Tomaszczyk, B. (red.) (2012): *Narodowy Korpus Języka Polskiego*. Wyd. Naukowe PWN, Warszawa.
- Przybylski, R. K. (2015): Akceptacja sprzeczności [w:] *Pochwała istnienia. Studia o twórczości Julii Hartwig*, Kulesza-Gulczyńska, B., Winiecka, E. (red.). Bogucki Wyd. Naukowe, Poznań, s. 33-42.
- Puzynina, J. (1988): Ironia jako element języka osobniczego. [w:] *Język osobniczy jako przedmiot badań lingwistycznych*, Brzeziński J. (red.). Wyższa Szkoła Pedagogiczna, Zielona Góra, s. 35-44.
- Rospond, S. (1972): Rola wielkich pisarzy w kształtowaniu języka ogólnonarodowego. [w:] *Symbolae polonicae in honorem Stanisłai Jodłowski*, Zaleski, J. (red.). Zakład Narodowy im. Ossolińskich, Wrocław, s. 137-144.
- Rudnicka-Fira, E. (1986): *Słownictwo „Dziadów” Adama Mickiewicza w świetle analizy statystycznej (wybór problematyki)*. Wyd. Uniwersytetu Śląskiego, Katowice.
- Rybicki, J. (2014): Pierwszy rzut oka na stylometryczną mapę literatury polskiej. [w:] *Teksty Drugie*, 2, s. 106-128.
- Rybicki, J. (2017): Drugi rzut oka na stylometryczną mapę literatury polskiej. [w:] *Forum Poetyki*, 8, s. 6-21.
- Saloni, Z., Świdziński, M. (1998): *Składnia współczesnego języka polskiego*. Wyd. Naukowe PWN, Warszawa.
- Saloni, Z., Woliński M., Wołosz R., Gruszczyński W., Skowrońska D. (2012): *Słownik gramatyczny języka polskiego*, wyd. II. Wyd. Uniwersytetu Warszawskiego, Warszawa.

- Sambor, J. (1969): *Badania statystyczne nad słownictwem (na materiale „Pana Tadeusza”)*. Zakład Narodowy im. Ossolińskich, Wrocław-Warszawa-Kraków.
- Skibski, K. (2015): Apostazja, szereg, parataksa – retardacja a składnia poetycka Julii Hartwig. [w:] *Pochwała istnienia. Studia o twórczości Julii Hartwig*, Kulesza-Gulczyńska B., Winiecka, E. (red.). Bogucki Wyd. Naukowe, Poznań, s. 135-142.
- Skibski, K. (2017): *Poezja jako iteratura*. Wyd. Naukowe UAM, Poznań.
- Skorupka, S., Auderska, H., Łempicka, Z. (1968): *Mały słownika języka polskiego*. Wyd. Naukowe PWN, Warszawa.
- Sławkowa, E. (2016): *Tekst literacki w kręgu językoznawstwa*, tom II. Wyd. Uniwersytetu Śląskiego, Katowice.
- Stankowska, A. (2015): Czułość wobec istnienia. Wokół postawy klasycznej Julii Hartwig. [w:] *Pochwała istnienia. Studia o twórczości Julii Hartwig*, Kulesza-Gulczyńska, B., Winiecka, E. (red.). Bogucki Wyd. Naukowe, Poznań, s. 43-52.
- Stachurski, E. (1989): *Słownictwo w utworach polskich naturalistów. Badania statystyczne*. Wyd. Naukowe WSP, Kraków.
- Szałkiewicz, Ł., Przepiórkowski, A. (2012): Anotacja morfoskładniowa [w:] *Narodowy Korpus Języka Polskiego*, Przepiórkowski, A., Bańko, M., Górski, R. L., Lewandowska-Tomaszczyk, B. (red.). Wyd. Naukowe PWN, Warszawa, s. 59-96.
- Szaruga, L. (2002): Głosy o tomie Julii Hartwig „Nie ma odpowiedzi”. [w:] *Zeszyty Literackie*, 1(77). Warszawa – Paryż – Mediolan, s. 131.
- Szumaska, D. (2006): *Przymiotnik jako przyłączone wyrażenie predykatywne*. Wyd. Universitas, Kraków.
- Szupryczyńska, M. (1980): *Opis składniowy przymiotnika*. Wyd. UMK, Toruń.
- Szymczak, M. (red.) (1978-81): *Słownik języka polskiego*, t. I-III. Wyd. Naukowe PWN Warszawa.
- Śliwiński, W. (1990): *Łączliwość składniowo semantyczna przymiotników z rzeczownikami we współczesnym języku polskim*. Nakład Uniwersytetu Jagiellońskiego, Kraków.
- Śliwiński, W. (1993): *Słownik semantyczno-syntaktyczny przymiotników polskich*. Nakład Uniwersytetu Jagiellońskiego, Kraków.
- Tarski, A. (1933): *Pojęcie prawdy w językach nauk dedukcyjnych*. Nakładem Towarzystwa Naukowego Warszawskiego, Warszawa.
- Taylor, J.R. (2007): *Gramatyka kognitywna*. Wyd. Universitas, Kraków.
- Telicki, M. (2005): Związki twórcze Julii Hartwig z kulturą francuską [w:] *Napis*, seria IX, s. 216-324, [http://rcin.org.pl/Content/56545/Wa248\\_68663\\_P-I-2795\\_telicki-zwiazki.pdf](http://rcin.org.pl/Content/56545/Wa248_68663_P-I-2795_telicki-zwiazki.pdf) [dostęp 15.07.2021].
- Telicki, M. (2009): *Poetycka antropologia Julii Hartwig*. Wyd. Poznańskie, Poznań.
- Tokarski, R. (2001): Słownictwo jako interpretacja świata. [w:] *Współczesny język polski*, Bartmiński, J. (red.), Lublin.
- Topolińska, Z. (red.) (1984): *Gramatyka współczesnego języka polskiego. Składnia*. Wyd. Naukowe PWN, Warszawa.
- Vetulani, G., Vetulani, Z. (1984): Norme et statistique. Remarques a propos de „Langue française et linguistique quantitative” de Charles Muller. [w:] *Lingua Posnaniensis*, 27, s. 127-137.

- Vetulani, G. (2000): *Rzeczowniki predykatywne języka polskiego. W kierunku syntaktycznego słownika rzeczowników predykatywnych*. Wyd. Naukowe UAM, Poznań.
- Vetulani, G. (2012): *Kolokacje werbo-nominalne jako samodzielne jednostki języka. Syntaktyczny słownik kolokacji werbo-nominalnych języka polskiego na potrzeby zastosowań informatycznych. Część I*. Wyd. Naukowe UAM, Poznań.
- Vetulani, G. (2013): Problemy i korzyści wynikające z automatycznego przetwarzania korpusów – na przykładzie badań z zakresu predykcji rzeczownikowej w języku polskim. [w:] *Roczniki Humanistyczne*, Tom LXI, z. 8, s. 13-24.
- Vetulani, Z. (1989): *Linguistic Problems in the Theory of Man-Machine Communication in Natural Language*. Brockmeyer, Bochum.
- Vetulani, Z., Martinek, J., Obrębski, T. (1997): Lexical resources and tools for tagging Polish texts within GRAMLEX [w:] *Linguisticae Investigationes*, Vol. 21:2, s. 401–416.
- Vetulani, Z., Walczak, B., Obrębski, T., Vetulani, G. (1998a): *Unambiguus coding of the inflection of Polish nouns and its application in electronic dictionaries – format POLEX/Jednoznaczne kodowanie fleksji rzeczownika polskiego i jego zastosowanie w słownikach elektronicznych – format POLEX*. Wyd. Naukowe UAM, Poznań.
- Vetulani, Z., Martinek, J., Obrębski, T., Vetulani, G. (1998b): *Dictionary Based Methods and Tools for Language Engineering*. Wyd. Naukowe UAM, Poznań.
- Vetulani, Z. (2004): *Komunikacja człowieka z maszyną. Komputerowe modelowanie kompetencji językowej*. Akademicka Oficyna Wydawnicza Exit, Warszawa.
- Vetulani, Z., Kubis, M., Obrębski, T. (2010): PolNet - Polish WordNet: Data and Tools. [w:] *Proceedings of the 7th International Conference on Language Resources and Evaluation LREC'10*, Calzolari N., Choukri K., Maegaard B., Mariani, J., Odijk J., Piperidis S., Rosner M., Tapias D. (red.). Valletta s. 3793–3797.
- Vetulani, Z., Witkowska, M., Menken, S. (2015): Corpus Based Studies on Language Expression of Opinions. [w:] *Human Language Technologies as a Challenge for Computer Linguistics, 7th Language and Technology Conference. Proceedings*, Vetulani, Z., Mariani, J. (red.). Wyd. Fundacja UAM, Poznań, s. 365-369.
- Vetulani, Z., Witkowska, M., Canbolat, U. (2018): TSCC: a New Tool to Create Lexically Saturated Text Subcorpora. [w:] *Proc. of LREC 2018 "1st Workshop on Computational Impact Detection from Text Data"* Diesner, J., Rehm, G., Witt, A. (red.). ELRA/ELDA, Paris, s. 22-26
- Vetulani, Z., Witkowska, M., Menken, S., Canbolat, U. (2018): Saturation Test in Application to Validation of Opinion Corpora: A Tool for Corpora Processing. [w:] *Human Language Technology. Challenges for Computer Science and Linguistics. 7th Language and Technology Conference, LTC 2015. Poznań, Poland, November 27-29, 2015*, Vetulani, Z., Mariani, J., Kubis, M. (red.). LNAI Series, nr 10930, s. 378-390.
- Vetulani, Z., Witkowska, M., Kubis, M. (2019): TSCC as a tool for lexical structure studies in application to literary output. [w:] *Proceedings of the 9th Language and Technology Conference. Human Language Technologies as a Challenge for Computer Science and Linguistics* Vetulani Z., Paroubek P. (red.). Poznań, s. 22-24.

- Vetulani, Z., Choukri, K., Mariani, J. Paroubek, P. (2019): Language Technologies for Less-Resourced-Languages. LRL Workshop Series at the Language and Technology Conferences (LTC) from 2009 to 2019. [w:] *International Conference On Language Technologies For All: Enabling Linguistic Diversity And Multilingualism Worldwide. Proceedings of LT4All*, Gilles A. (red.). ELRA, Paris, s. 123-126, <https://lt4all.elra.info/proceedings/lt4all2019/> [dostęp 10.07.2021]
- Vetulani, Z., Vetulani G. (2019): The Case of Polish on its Way to Become a Well-Resourced Language. [w:] *International Conference On Language Technologies For All: Enabling Linguistic Diversity And Multilingualism Worldwide. Proceedings of LT4All*, Gilles A. (red.). ELRA, s. 388-392, <https://lt4all.elra.info/proceedings/lt4all2019/> [dostęp 10.07.2021]
- Walczak, B. (1988): Sienkiewicz w dziejach języka polskiego [w:] *Język osobniczy jako przedmiot badań lingwistycznych* Brzeziński J. (red.). Wyższa Szkoła Pedagogiczna, Zielona Góra, s. 129-139.
- Walkowska, J. (2012): *Modelowanie kompetencji dialogowej człowieka na potrzeby jej emulacji w zarządzających wiedzą systemach informatycznych współpracujących z wieloma użytkownikami*. Rozprawa doktorska. IPI PAN, Warszawa.
- Węgrzynek, K. (1995): *Składnia przymiotnika w ujęciu generatywno-transformacyjnym*. IJP PAN, Kraków.
- Wilkoń, A. (1978): Język a styl tekstu literackiego. [w:] *Język Artystyczny*, 1, Katowice, s. 11-21, [https://bazhum.muzhp.pl/media/files/Jezyk\\_Artystyczny/Jezyk\\_Artystyczny-r1978-t1/Jezyk\\_Artystyczny-r1978-t1-s11-21/Jezyk\\_Artystyczny-r1978-t1-s11-21.pdf](https://bazhum.muzhp.pl/media/files/Jezyk_Artystyczny/Jezyk_Artystyczny-r1978-t1/Jezyk_Artystyczny-r1978-t1-s11-21/Jezyk_Artystyczny-r1978-t1-s11-21.pdf) [dostęp 10.07.2021]
- Wiśniewski, J.T. (2013): Barokowa muzyka z Wenecji w wierszach Julii Hartwig. [w:] Wiśniewski, J.T., *Ku harmonii. Poetyckie style słuchania muzyki*. Wyd. Uniwersytetu Łódzkiego, Łódź, s. 101-113.
- Winiecka, E. (2015): Zapis amerykański Julii Hartwig. [w:] *Pochwała istnienia. Studia o twórczości Julii Hartwig*, Kulesza-Gulczyńska, B., Winiecka, E. (red.). Bogacki Wyd. Naukowe, Poznań, s. 99-114.
- Witosz, B. (2009): *Dyskurs i stylistyka*. Wyd. Uniwersytetu Śląskiego, Katowice.
- Wojtasiewicz, O.A. (1972): O polskich przymiotnikach niepredykatywnych. [w:] *Poradnik Językowy*, 7(301), s. 394-399.
- Wyka, K. (1962): Słowa-klucze. [w:] *Zagadnienia Rodzajów Literackich*, t. IV, z. 2 (7); przedruk: (1973) *Stylistyka polska. Wybór tekstów*, wybór, oprac. i wstęp Miodońska-Brookes E., Kulawik A., Tatra M., Wyd. Naukowe PWN, Warszawa, s. 153-167.
- Zgółkowie, H. i T. (1992): *Słownictwo współczesnej poezji. Listy frekwencyjne*, tom I i II. Wyd. Naukowe UAM, Poznań.
- Zipf ,G.K.(1943): *Human Behavior and Principle of Least Effort*. Addison-Wesley, Cambridge.

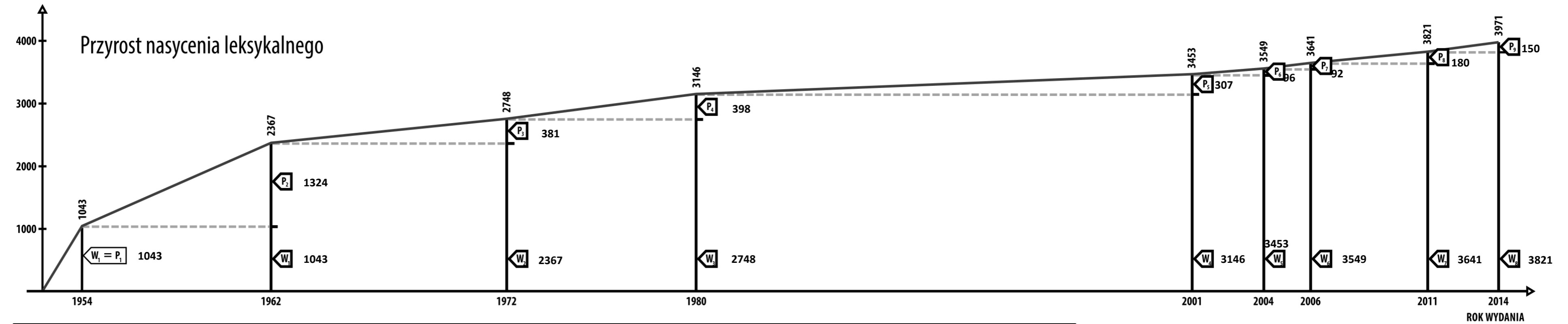
## ILUSTRACJE

• Zestawienie 1. Fragment korpusu prozy – przykład archiwizowania poszczególnych wersji korpusu prozy Julii Hartwig	63
• Zestawienie 2. Podsumowanie zawartości korpusu prozy Julii Hartwig.	64
• Zestawienie 3. Podsumowanie zawartości korpusu poezji Julii Hartwig.	65
• Tabela 1. Zestawienie ilościowe dla korpusu prozy i poezji Julii Hartwig zebranych w całości w plikach dwóch plikach nazwanych proza-all i poezja-all.	65
• Tabela 2. Zestawienie dla korpusu prozy – przedstawione dane liczbowe dotyczą wersji korpusu po normalizacji.	67
• Tabela 3. Zestawienie metadanych dla korpusu poezji – przedstawione dane liczbowe dotyczą wersji korpusu po normalizacji.	68
• Przykład 1. Podsumowanie procesu lematyzacji korpusu proza-all.	76
• Tabela 4. Lista nr 1 Lematyzacja korpusu poezji – opis gramatyczny (morfologiczny) wszystkich jednostek w korpusie	77
• Tabela 5. Lista nr 2 Lematyzacja korpusu prozy – opis gramatyczny (morfologiczny) wszystkich jednostek w korpusie	77
• Tabela 6. Lista nr 3 Lematyzacja korpusu poezji – opis gramatyczny (morfologiczny) przymiotników	78
• Tabela 7. Lista nr 4 Lematyzacja korpusu prozy – opis gramatyczny (morfologiczny) przymiotników	78
• Tabela 8 Etapy przetwarzania list przymiotników wygenerowanych przez lematyzator – podsumowanie.	80
• Tabela 9. Zestawienie frekwencji leksemów przymiotnikowych dla poszczególnych dzieł w korpusie prozy	84
• Tabela 10. Zestawienie frekwencji leksemów przymiotnikowych dla poszczególnych dzieł 001-011 w korpusie poezji	87
• Tabela 11. Zestawienie frekwencji leksemów przymiotnikowych dla poszczególnych dzieł 012-022 w korpusie poezji	88
• Zestawienie 4. Widok okna Korpusomatu prezentującego zawartość korpusu prozy Julii Hartwig	90
• Zestawienie 5. Widok okna Korpusomatu prezentującego zawartość korpusu poezji Julii Hartwig	91
• Przykład 2. Początek listy konkordancji dla korpusu prozy zachowanych w pliku csv (kolumny D – M, patrz Przykład 3)	93
• Przykład 3. Uzgodnienia interpretacji przymiotników wg paradygmatu POLEX w korpusie prozy (kolumny A, B, C – patrz Przykład 2)	96
• Przykład 4. Konkordancje dla przymiotnika predykatywnego w korpusie poezji	98
• Przykład 5. Konkordancje dla przymiotnika predykatywnego w korpusie prozy	99
• Tabela 12. Fragment listy konkordancji dla form leksemu piękny w korpusie prozy	101
• Tabela 13. Fragment listy konkordancji dla form leksemu piękny w korpusie poezji	102
• Tabela 14. Fragment listy konkordancji dla form leksemu biały w korpusie prozy	103
• Tabela 15. Fragment listy konkordancji dla form leksemu biały w korpusie poezji Poza różnymi formami przymiotnika biały pogrubione zostały też takie formy pokrewne, czasownikowe czy rzeczownikowe jak: bieleje, białość.	104
• Przykład nr 6. Ekran startowy TSCC 1.3.	105
• Przykład nr 7. Funkcjonalności programu TSCC 1.3.	106
• Zestawienie 6. Zawartość folderu Saturation Test dla korpusu proza-all	108
• Wykres 1. Przyrost form przymiotnikowych (tokenów) dla korpusu prozy Julii Hartwig	110

• Tabela 16. Przyrosty użycia przymiotników dla prozy z lat 1954 – 2014.	114
• Tabela 17. Przyrosty użycia przymiotników dla poezji z lat 1956 – 2016.	115
• Wykres 3. Przyrost nasycenia leksykalnego dla korpusu prozy Julii Hartwig (lata 1954 -2016)	116
• Wykres 4. Przyrosty słownictwa wirtualnego dla prozy Julii Hartwig (oszacowanie dolne na lata 1954-1980)	117
• Wykres 5. Przyrosty słownictwa wirtualnego dla prozy Julii Hartwig (oszacowanie dolne na lata 1980-2014)	118
• Wykres 6. Przyrost nasycenia leksykalnego dla korpusu poezji Julii Hartwig (lata 1956 -2016)	119
• Wykres 7. Przyrosty słownictwa wirtualnego dla poezji Julii Hartwig (oszacowanie dolne na lata 1956- 1992)	120
• Wykres 8. Przyrosty słownictwa wirtualnego dla poezji JH (oszacowanie dolne na lata 1980 2016)	121

## UZUPEŁNIENIE 1

### PROZA



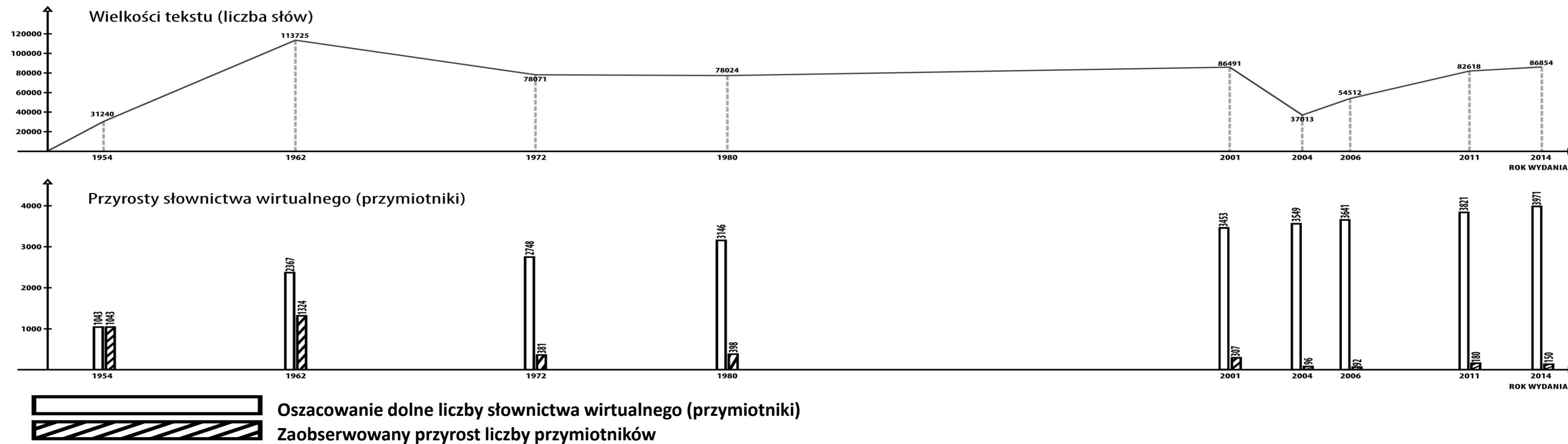
#### LEGENDA

1. Wielkość  $W_i$  jest liczbą stanowiącą oszacowanie dolne słownictwa wirtualnego (przymiotniki) do roku publikacji włącznie.
2. Wielkość  $P_i$  wyraża przyrost liczby przymiotników (w roku publikacji) w stosunku do liczby zaobserwowanych wcześniej.
3. Przy powyższych oznaczeniach:  $W_i = P_i$ ,  $W_{i+1} = P_{i+1} + W_i$ .
4. Wykres obrazuje przyrosty liczby zaobserwowanych przymiotników w kolejnych opublikowanych tomach prozy (reprezentowanych przez paradygmatyczne formy podstawowe /mianownik, liczba pojedyncza, stopień równy/).

Ad. Wykres ze str. 115. Przyrost nasycenia leksykalnego dla korpusu prozy Julii Hartwig (lata 1954 -2014).

## UZUPEŁNIENIE 2

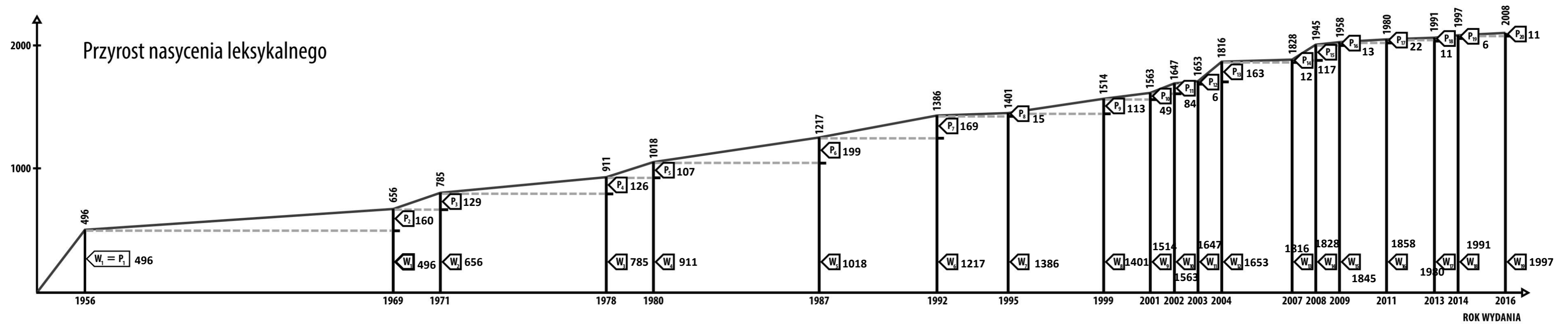
### PROZA



Ad. Wykres na str. 116 i 117. Przyrosty słownictwa wirtualnego dla prozy Julii Hartwig (oszacowanie dolne na lata 1954 – 2014).

### UZUPEŁNIENIE 3

#### POEZJA

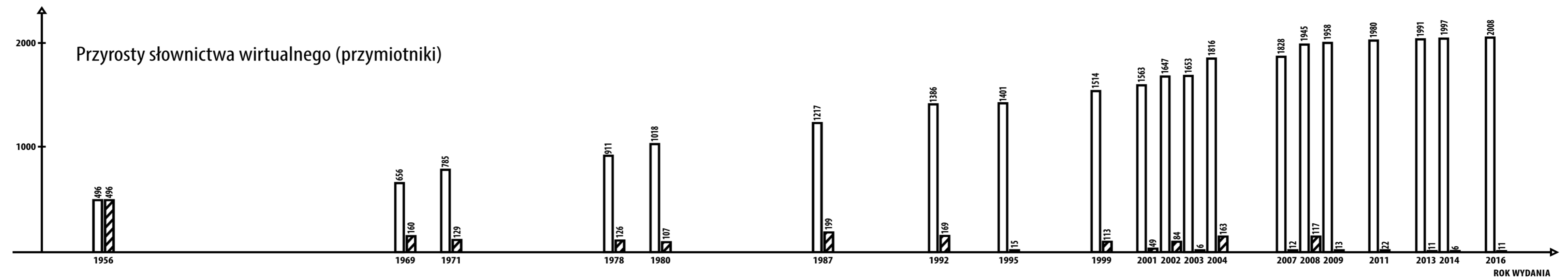
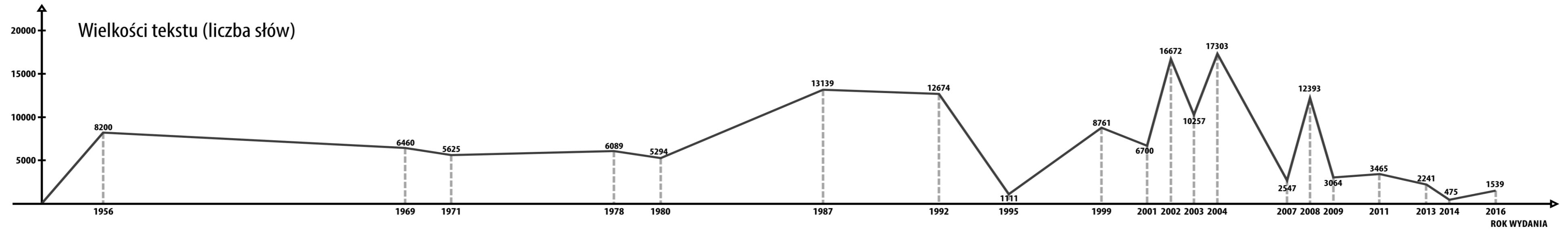




#### LEGENDA

1. Wielkość  $W_i$  jest liczbą stanowiącą oszacowanie dolne słownictwa wirtualnego (przymiotniki) do roku publikacji włącznie.
2. Wielkość  $P_i$  wyraża przyrost liczby przymiotników (w roku publikacji) w stosunku do liczby zaobserwowanych wcześniej.
3. Przy powyższych oznaczeniach:  $W_i = P_i$ ,  $W_{i+1} = P_{i+1} + W_i$ .
4. Wykres obrazuje przyrosty liczby zaobserwowanych przymiotników w kolejnych opublikowanych tomach poezji (reprezentowanych przez paradygmatyczne formy podstawowe /mianownik, liczba pojedyncza, stopień równy/).

Ad. Wykres na str. 118. Przyrost nasycenia leksykalnego dla korpusu poezji Julii Hartwig (lata 1956 – 2016).

**UZUPEŁNIENIE 4  
POEZJA**



 Oszacowanie dolne liczby słownictwa wirtualnego (przymiotniki)  
 Zaobserwowany przyrost liczby przymiotników

Ad. Wykres na str. 119 i 120. Przyrosty słownictwa wirtualnego dla poezji Julii Hartwig (oszacowanie dolne na lata 1956 – 2016).