

Aleksandra Deskur 

Uniwersytet im. Adama Mickiewicza, Poznań
aleksandra.deskur@amu.edu.pl

Ewa Nowak-Pasterska 

Uniwersytet im. Adama Mickiewicza, Poznań
ewa.n@amu.edu.pl

Dorota Rojszczak-Robińska 

Uniwersytet im. Adama Mickiewicza, Poznań
dorroj@amu.edu.pl

Wojciech Stelmach 

Uniwersytet im. Adama Mickiewicza, Poznań
ws45802@amu.edu.pl

Magdalena Wismont 

Uniwersytet im. Adama Mickiewicza, Poznań
magwis7@amu.edu.pl

Kinga Zalejarz 

Uniwersytet im. Adama Mickiewicza, Poznań
kinga.zalejarz@amu.edu.pl

Olga Ziółkowska 

Uniwersytet im. Adama Mickiewicza, Poznań
olga.stramczewska@o2.pl

ANOTACJA GRAMATYCZNA TEKSTU STAROPOLSKIEGO – PROBLEMY I WYZWANIA¹

Słowa kluczowe: anotacja gramatyczna, siatka klas i kategorii, tagset, średniowieczna polszczyzna, staropolskie apokryfy Nowego Testamentu

Keywords: grammatical annotation, class and category network, tagset, medieval Polish language, Old Polish New Testament apocrypha

Wprowadzenie

Projekt „Początki języka polskiego i kultury religijnej w świetle średniowiecznych apokryfów Nowego Testamentu. Uniwersalne narzędzie do badań polskich tekstów

¹ Artykuł realizowany w ramach grantu Narodowego Centrum Nauki (projekt nr 2017/26/E/HS2/00083) pt. „Początki języka polskiego i kultury religijnej w świetle średniowiecznych

apokryficznych” jest realizowany od marca 2018 r. w ramach grantu Narodowego Centrum Nauki na Wydziale Filologii Polskiej i Klasycznej UAM. Cel naukowy grantu stanowi wielopłaszczyznowa, multidyscyplinarna analiza ważnych źródeł staropolskiej kultury religijnej i języka religijnego, tj. dziewięciu staropolskich apokryfów Nowego Testamentu, a także rozwój badań nad polskim językiem religijnym, duchowością i kulturą dawną. Staropolskie apokryfy Nowego Testamentu to siedem tekstów zachowanych w rękopisach, tj. *Rozmyślanie przemyskie*², *Sprawa chędogo o męce Pana Chrystusowej*³, *Rozmyślania dominikańskie*⁴, *Karta Rogawskiego*⁵, *List Lentulusa*⁶, *Ewangelia Nikodema*⁷ i *Historyja Trzech Kroli*⁸, oraz dwa druki, tj. *Żywot Pana Jezu Krysta Baltazara Opeca*⁹ i *Żywot świętej Anny*¹⁰, łącznie ponad 2000 stron rękopisów i druków. Są to teksty fundamentalne dla dziejów polszczyzny.

Jednym z nadrzędnych celów projektu badawczego było stworzenie kompleksowego narzędzia internetowego, mającego służyć wszechstronnym badaniom staropolskich tekstów apokryficznych. Miarodajnym efektem naszej pracy jest strona *Staropolskie apokryfy Nowego Testamentu* (apocrypha.amu.edu.pl), a materiały dostępne na niej są stale uzupełniane. Pod jej adresem znajdują się teksty apokryfów, podane w transkrypcji i transliteracji, przygotowane przez zespół grantowy na potrzeby narzędzia. Po wyborze określonego tekstu z zakładki *Teksty* użytkownik ma możliwość wglądu do niego w zakresie wybranych przez siebie kart i wersów (nie wykluczając możliwości przejrzania całego tekstu). Następnie, po wyborze

apokryfów Nowego Testamentu. Uniwersalne narzędzie do badań polskich tekstów apokryficznych” pod kierownictwem Doroty Rojszczak-Robińskiej. Udział współautorów jest równy i wynosi po 14,28%.

- 2 *Rozmyślanie przemyskie*, dalej RP – rękopis obecnie przechowywany w Bibliotece Narodowej w Warszawie (sygn. 8024 III). Jest to papierowy kodeks liczący 852 strony.
- 3 *Sprawa chędogo o męce Pana Chrystusowej*, dalej Sch – rękopis przechowywany w Bibliotece Narodowej (sygn. 3040 III), skopiowany w 1544 r. przez Wawrzyńca z Łaska. Znajduje się w papierowym kodeksie wraz z dwoma innymi polskimi utworami (*Ewangelia Nikodema* i *Historyja Trzech Kroli*).
- 4 *Rozmyślania dominikańskie*, dalej RD – pergaminowy kodeks liczący 122 karty, bogato iluminowany, obecnie przechowywany w Bibliotece Klasztoru Karmelitanek Bosych w Krakowie (sygn. 287).
- 5 *Karta Rogawskiego*, dalej KR – pergaminowa karta, która obecnie znajduje się w dziale rękopisów Biblioteki XX. Czartoryskich w Krakowie w teczce pod sygn. 2566.
- 6 *List Lentulusa*, dalej LL – rękopis Biblioteki Jagiellońskiej nr 2151, zapisany ok. poł. XV w., na k. 302v–303r.
- 7 Zob. przypis 3.
- 8 Zob. przypis 3.
- 9 B. Opec, *Żywot wszechmocnego Syna Bożego Pana Jezu Krysta*, dalej ŻPK – druk Hieronima Wietora z 1522 r., egzemplarz Biblioteki Seminarium Metropolitalnego Archidiecezji Warszawskiej, sygn. O. 24.1 (74118).
- 10 *Żywot świętej Anny*, dalej ŻSA – unikat drugiego wydania znajduje się w Bibliotece Jagiellońskiej, sygn. Cim. 147.

interesującego fragmentu, wyświetla się transkrypcja tekstu i ewentualne noty do niej. W polu pracy z tekstem możliwy jest również ogłód transliteracji tekstu i not do niej, a także zapoznanie się z innymi transkrypcjami danego tekstu, źródłami łacińskimi i miejscami paralelnymi w innych tekstach apokryficznych.

Wstawienie interpunkcji do tekstu staropolskiego jest zawsze jego interpretacją, a transkrypcja odzwierciedla myślenie badacza o dawnej składni (por. np. Krążyńska, Mika 2007; Stramczewska 2014; Mika, Ziółkowska 2017). Z tego względu uznaliśmy, że możliwość porównania edycji przygotowanych w różnym czasie przez różnych badaczy będzie wartościowa oraz przydatna w badaniach składniowych, i zaplanowaliśmy zakładkę *Inne transkrypcje*, w której znajdują się teksty interesujących nas apokryfów z wcześniejszych wydań różnych badaczy¹¹.

Nie sposób badać kultury literackiej i umysłowej średniowiecza bez refleksji nad źródłami tekstów dawnych. W zakładce *Źródła* można znaleźć zestawione z fragmentami polskich tekstów łacińskie i czeskie dzieła, które mogły być inspiracją staropolskich twórców (mogli z nich korzystać bezpośrednio lub pośrednio). Zostały one podzielone na pięć grup tekstów w zależności od ich statusu (kanoniczny, patrystyczny), gatunku i czasu powstania: 1) Nowy Testament, 2) Stary Testament, 3) Ojcowie i Doktorzy Kościoła, 4) Apokryfy i legendy, 5) Inne pisma chrześcijańskie¹².

Zakładka *Wątki* pozwala na eksplorację tekstów apokryficznych pod kątem rozgrywających się w nich wydarzeń. Wątki zostały podzielone na *Wydarzenia* i *Osoby*. Dzięki temu użytkownik może porównać np. sposób przedstawiania wjazdu do Jerozolimy czy modlitwy w Ogrójcu we wszystkich tekstach lub wyświetlić wszystkie wątki związane przykładowo z Klaudią Prokulą czy z Judaszem. Funkcjonalność umożliwia szybkie i precyzyjne zlokalizowanie określonego wydarzenia bądź osoby w nim działającej w całości badanych tekstów apokryficznych.

Narzędzie zaplanowano dla przedstawicieli różnych dyscyplin: badaczy średniowiecznej polszczyzny (fleksji, składni, semantyki, a także stylu), recepcji określonych dzieł kultury w polskim średniowieczu, osób zainteresowanych przekładem w epokach dawnych czy obecnością określonych wątków i motywów w tekstach staropolskich. W artykule skupiliśmy się na funkcjonalnościach zaplanowanych z myślą o językoznawcach: komentarzu gramatycznym, wyszukiwarce po słowach oraz wyszukiwarce gramatycznej. Celem naszego opracowania jest prezentacja tych trzech funkcjonalności oraz siatki klas i kategorii, a także wskazanie problemów związanych z anotacją tekstu dawnego.

11 Do tej pory zestawiono: Brückner 1904, 1907; Wydra, Rzepka, Kuraszkiewicz 1984/2004; Rzepka, Wydra, Adamczyk 1996; Mazurkiewicz 2000; Wydra, Wójcik, Krzak-Weiss 2014; Rojszczak-Robińska 2017; Mazurkiewicz s.a.; Twardzik IJP PAN; Twardzik s.a. Więcej zob.: *Opracowania, O innych wydaniach* na stronie apocrypha.amu.edu.pl.

12 Zob. *Źródła* na stronie apocrypha.amu.edu.pl; więcej zob.: Rojszczak-Robińska 2021.

Co i jak anotujemy?

Anotacji gramatycznej poddano teksty transkrybowane, a co za tym idzie – już zinterpretowane. Przyjęte przez zespół badawczy rozwiązania edytorskie zostały poddyktowane kilkoma czynnikami, które szczegółowo omówiliśmy na stronie *Staropolskie apokryfy Nowego Testamentu* (dalej także: *Apocrypha*):

Ich przyjęcie warunkowane jest wieloma czynnikami:

- zakorzeniem w tradycji edytorskiej tekstów dawnych wyznaczonej przez *Zasady wydawania tekstów staropolskich. Projekt* i wielokrotnym odchodzeniem od niektórych zaleceń w praktyce w rozlicznych edycjach, nawet tych, w których deklarowano ich stosowanie,
- rozwiązaniami przyjętymi i diskutowanymi wobec najstarannie wydanych w ostatnich latach zabytków filologicznych (*Kazania świętokrzyskie, Rozmyślanie przemyskie*, inne teksty Biblioteki Zabytków Polskiego Piśmiennictwa Średniowiecznego czy Korpusu Tekstów Staropolskich IJP PAN),
- wynikami najnowszych badań nad niektórymi aspektami tekstów staropolskich (składnia, interpunkcja i delimitacja tekstu, wielowarstwowość genetyczna tekstów, stosunek do źródeł, błędy pisarskie),
- funkcjonalnością narzędzia, które z jednej strony ma umożliwić badania nad apokryfami specjalistom z różnych dziedzin (nie tylko językoznawcom czy filologom), z drugiej obejmuje zarówno rękopisy, jak i druki (narzędzie musi przeszukiwać jedną i drugą grupę tekstów) (*Apocrypha*: zakładka *O projekcie, Przyjęte rozwiązania edytorskie*).

Pełny spis przyjętych rozwiązań czytelnik znajdzie na cytowanej powyżej stronie *Apocrypha*.

Ze względu na specyfikę materiału konieczna była anotacja ręczna. Polegała ona na jednostkowym przypisaniu każdego wyrazu tekstowego do określonej klasy gramatycznej oraz nadaniu mu wartości kategorii gramatycznych, przysługujących klasie i konkretnej jednostce¹³, a także wskazaniu lematu¹⁴, który jest traktowany jak forma słownikowa, pewna etykieta dla zbioru wszystkich form danego wyrazu.

Jednym z podstawowych założeń naszej anotacji gramatycznej jest adekwatność zestawu klas i kategorii do materiału staropolskiego. Poszczególne jednostki zostały wyodrębnione na podstawie dwóch kryteriów: formalnego (morfosyntaktycznego) i semantycznego. Decyzja ta związana jest z tym, że w języku staropolskim kategoryzacja semantyczna była równie istotna co kategoryzacja strukturalna (por. Krążyńska, Mika, Słoboda 2015: 97–98). Skupienie się na czysto formalnych kryteriach byłoby nakładaniem na staropolszczyznę kategorii wykształconych dopiero w języ-

¹³ Więcej na temat klas i kategorii gramatycznych zob. w dalszej części artykułu.

¹⁴ Więcej na temat lematu w narzędziu *Staropolskie apokryfy Nowego Testamentu* zob. na stronie *apocrypha.amu.edu.pl* (dostęp: 26 IV 2022).

ku epok późniejszych, a co za tym idzie – stanowiłoby nieuzasadnioną interpretację języka staropolskich apokryfów.

Szczególnie wyraźnym przykładem, który obrazuje konieczność uwzględnienia kryterium semantycznego, jest liczebnik (tradycyjnie wyodrębniana w ten sposób część mowy została przez nas podzielona na siedem odrębnych klas, por. część *System klas i kategorii*). Anotacja czysto formalna wymagałaby zaliczenia np. liczebnika porządkowego do klasy przymiotników (ma on takie same jak przymiotnik kategorii gramatyczne). Innym przykładem tak wyodrębnionych klas są zaimki. Zdecydowaliśmy, aby nie przypisywać zaimkom klas analogicznych części mowy (przymiotnik, rzeczownik itd.), ale stworzyć dla nich osobne klasy: zaimek przymiotny, rzeczowny itd. Do nich poszczególne wyrazy tekstowe są zaliczane na podstawie kryterium formalnego, zrezygnowaliśmy natomiast z kryterium funkcjonalnego (zaimek osobowy, pytajny, dzierżawczy itp.). Obie te części mowy – liczebnik i zaimek – na podstawie kryterium semantycznego zostały wydzielone jako odrębne klasy. Kryterium formalne pozwoliło natomiast przypisać do tych klas poszczególne wyrazy tekstowe.

Podjęcie się anotacji tekstów dawnych wymusiło zaplanowanie siatki klas i kategorii, a następnie nałożenie jej na analizowany materiał. Zdajemy sobie jednak sprawę, że przedmiotem naszej refleksji są teksty pochodzące z okresu intensywnych przemian gramatycznych, nasze decyzje zatem mają charakter interpretacyjny – wiążą się z przyporządkowywaniem form do klas także w sytuacji, gdy ich status morfologiczny może być zmienny:

Klasyfikacja jednostek języka polega na wyodrębnieniu spośród całego zasobu leksykalnego kilku bardzo ogólnych klas wyrazowych, które tworzą podstawowy poziom składników gramatycznych języka (por. Wajszczuk 2010: 15). W ujęciu wyłącznie dynamicznym stałych klas wyrazowych nie ma. Mamy tu bowiem do czynienia z ciągłym ruchem: zaimki zmieniają się w spójniki czy partykuły (por. Sobotka 2014a, 2014b), imiesłowy się adiektywizują bądź adverbializują, adwerbia stają się przyimkami lub partykułami itd. Nieodmienne zaimki i przysłówki chętnie nasycają się nowymi treściami bądź eksponują swoje funkcje sekundarne i zaczynają pracować na wyższych niż przedmiotowy poziomach języka (por. też Danielewiczowa 2012: 98; Żabowska 2015: 73–84) (Sobotka, Żabowska 2017: 8).

W staropolszczyźnie funkcjonowały wyrazy, których przynależność do danej klasy może być niejednoznaczna (por. np. Kleszczowa, Szczepanek 2014; Mika, Słoboda 2015; Pawelec 2015). Ich przykładem są partykuły i spójniki. Zrezygnowaliśmy z tworzenia klas przejściowych, potencjalnie przydatnych wtedy, gdy nie można jednoznacznie stwierdzić, czy mamy do czynienia ze spójnikiem czy z partykułą. Co za tym idzie, w wypadku tych dwóch części mowy decyzja o zaliczeniu danego wyrazu do jednej z klas podejmowana była jednostkowo, często w efekcie pogłębionej analizy danego fragmentu (nie tylko językowej, ale także stylistycznej i źródłowej).

W projektowaniu anotacji gramatycznej tylko częściowo punktem wyjścia mogły być dla nas istniejące już narzędzia do anotacji języka dawnego. Ograniczenia wynikały przede wszystkim z tego, że nawet w narzędziach służących analizie tekstów dawnych zastosowano system anotacji wypracowany w analizie tekstów współczesnych (np. w ramach Narodowego Korpusu Języka Polskiego (NKJP)). Najbardziej oczywistym punktem odniesienia mogła być dla nas KorBa – Elektroniczny Korpus Tekstów Polskich XVII i XVIII w. (do 1772 r.) – ponieważ zawiera materiał chronologicznie najbliższy tekstom zamieszczonym w narzędziu *Apocrypha*. Zastosowany w KorBa zestaw klas i kategorii został oparty jednak na Narodowym Korpusie Języka Polskiego:

Niniejsza instrukcja opisuje sposoby korzystania z zasobów zgromadzonych w *Elektronicznym Korpusie Tekstów Polskich z XVII i XVIII w. (do 1772 r.)* za pomocą wyszukiwarki MTAS. Dokument ten stanowi zmodyfikowaną wersję instrukcji opisującej sposób działania wyszukiwarki MTAS, opracowanej przez Witolda Kierasia na potrzeby *Korpusu tekstów polskich z lat 1830–1918*. Z kolei podstawą tej instrukcji była *Ściągowka do Narodowego Korpusu Języka Polskiego*, czyli instrukcja użytkownika wyszukiwarki Poliqarp, wykorzystującej, podobnie jak MTAS, język zapytań znany pod nazwą *Corpus Query Language* (CQL). *Ściągowka* została opracowana przez Adama Przepiórkowskiego, a następnie poprawiona i rozszerzona przez Jakuba Wilka i Aleksandra Buczyńskiego. Jej pełna wersja znajduje się w repozytorium wyszukiwarki Poliqarp (INSTR: 1).

W KorBa, a także NKJP, podstawową jednostką opisu gramatycznego jest fleksem, rozumiany jako „[zbiór] tylko tych form, które można scharakteryzować za pomocą tych samych kategorii gramatycznych” (ibid.: 2). Przyjęcie podobnego, przede wszystkim formalnego, sposobu wyodrębniania jednostek języka staropolskiego mogłoby spowodować „dopasowywanie” wyrazów tekstowych do jednostek wyodrębnionych na podstawie formalnych wyróżników. Stąd przyjęte w zespole grantowym rozwiązania są tylko częściowo kompatybilne z już istniejącymi narzędziami, uznano bowiem, że istotniejsze jest możliwie pełne uchwycenie specyfiki średniowiecznej polszczyzny niż zachowanie zgodności tagsetów.

Jak wskazaliśmy powyżej, zdecydowaliśmy się na anotację ręczną zamiast automatycznej, zastosowanej w innych narzędziach. Konieczne było też zaprojektowanie odrębnego systemu klas i kategorii. Wynika to przede wszystkim ze specyfiki materiału – w wypadku staropolszczyzny decyzja o przypisaniu danego wyrazu tekstowego do klasy i określeniu jego kategorii niejednokrotnie jest decyzją interpretacyjną, ponieważ granice między klasami nie były wówczas w takim stopniu ustabilizowane, jak w języku współczesnym. Ponadto wybrany korpus tekstów jest zbyt mały, by efektywnie wytrenować na nim automatyczne narzędzie do anotacji.

System klas i kategorii¹⁵

Wyodrębnienie klas i kategorii gramatycznych było ściśle związane ze specyfiką języka doby staropolskiej. W planowaniu siatki klas i kategorii przyświecała nam myśl, by z jednej strony w pełni oddać charakter materiału, który będzie anotowany, z drugiej – jasno wskazać ograniczenia wynikające z dostępnej postaci tekstu. Na kształt tekstów, które przyszło nam anotować, mają wpływ zarówno czynniki wewnętrzne, jak i zewnętrzne. Mamy do czynienia z rękopisami oraz starodrukami. W wielu miejscach interpretowaliśmy zapisy, co bezpośrednio przekłada się na anotację i identyfikację zarówno klasy, jak i wartości kategorii. Ze względu na etap rozwoju polskiej grafii niemożliwe jest jednoznaczne odczytanie zapisów z rękopisów (np. zapis: *anyely szly*, transkrypcja *anieli szli* – rodzaj męski, lub *aniely szly* – rodzaj niemęski), jak i przypisanych klas (zapis: *osądzycze*, transkrypcja *osądzicie* *impt* lub *osądzicie* *fut_simp*). Samo wstawienie interpunkcji powoduje, że z dwóch lub więcej możliwości interpretacji i określenia form gramatycznych pozostaje jedna (np. *O, szczęsne obiecanie!* [OBIECANIE subst sg nom n] lub *O szczęsne obiecanie!* [OBIECANIE subst sg voc n]).

Kłopotliwe okazały się także uszkodzenia całych kodeksów (jak RD, zob. Ziółkowska 2016), szczątkowy stan zachowania niektórych rękopisów (jak KR, zob. Rojszczak-Robińska 2017), uszkodzenia pojedynczych kart (jak w SCh) czy braki i niekonsekwencje wynikające z wielowarstwowości genetycznej tekstów (jak RP, zob. Mika, Rojszczak-Robińska 2016). Decyzje o niepoprawianiu i nieuzupełnianiu tekstu, które zapadły już na etapie tworzenia transkrypcji, miały wpływ także na anotację gramatyczną. Anotujemy ten materiał, który jest podstawą wszystkich wyszukiwarek w narzędziu, czyli tekst w transkrypcji. Wyrazy tekstowe w całości uzupełnione (np. <krolowie>) lub w całości usunięte (np. [krolowie]) są wyłączone z anotacji. Nie anotujemy wyrazów skreślonych przez pisarza. Nie anotujemy też możliwych innych odczytań, sygnalizowanych w notach. Anotujemy formy częściowo uzupełnione (np. wyraz<owi>) lub częściowo ucięte (np. wyraz[owi]) oraz zmienne (np. wyraz).

Teksty, na których pracujemy, pochodzą z przełomu XV i XVI w. Duża część z nich zachowała się jedynie w kopiach. W efekcie w obrębie jednego tekstu poświadczono są zarówno formy czy konstrukcje starsze, jak i nowsze. To także, jak wskazaliśmy wyżej, czas dynamicznych zmian w języku i wyodrębniania się funkcji składniowych niektórych elementów funkcyjnych (zob. Kleszczowa 2015; Mika, Słoboda 2015; Sobotka, Żabowska 2017).

Do wyodrębnienia klas gramatycznych zastosowaliśmy przede wszystkim kryteria morfosyntaktyczne oraz semantyczne. Wynikało to z potrzeby dostosowania się do specyfiki języka doby staropolskiej. Siatka klas obejmuje:

15 Serdecznie dziękujemy wszystkim członkom Zakładu Historii Języka Polskiego, z którymi na bieżąco mogliśmy dyskutować wszelkie problemy i wątpliwości. Szczególne podziękowania kierujemy w stronę profesor Agnieszki Słobody.

- rzeczownik (tag: subst) – do tej grupy należą wszystkie rzeczowniki, niezależnie od znaczenia i pochodzenia słowotwórczego, a więc także rzeczowniki odczasownikowe;
- przymiotnik (tag: adj) – w tej klasie znajdują się wszystkie przymiotniki zarówno w odmianie prostej, jak i złożonej. Decyzje o przynależności elementów do tej klasy zostały podjęte ze względu na kryterium formalne;
- przysłówek (tag: adv) – tu kwalifikujemy tradycyjnie rozumiane przysłówki oraz zleksykalizowane wyrażenia przyimkowe.

Tradycyjnie wyodrębniane jako osobne części mowy zaimek i liczebnik zostały przez nas podzielone ze względu na kryterium formalne. Poszczególne klasy wyodrębniono ze względu na kategorie (gramatyczne i słownikowe), które można im przypisać. W wyniku tego zaimki i liczebniki zostały zaanotowane jako:

- zaimek rzeczowny (tag: pron_subst),
- zaimek przymiotny (tag: pron_adj),
- zaimek liczebnny (tag: pron_num),
- zaimek przysłowny (tag: pron_adv),
- liczebnik główny (tag: num_card),
- liczebnik porządkowy (tag: num_ord),
- liczebnik zbiorowy (tag: num_col),
- liczebnik zespołowy (tag: num_incl),
- liczebnik krotny (tag: num_freq),
- liczebnik wieloraki (tag: num_mult),
- liczebnik ułamkowy (tag: num_fract).

Zamiast ogólnej klasy czasownika wyróżniliśmy kilka klas odpowiadających poszczególnym formom czasownikowym w zależności od tego, które wartości kategorii przyjmują i jakie konstrukcje budują. Odrębne klasy tworzą:

- forma teraźniejsza (tag: praes) – czasowniki w formie osobowej, mające znaczenie czasu teraźniejszego (dotyczy to też czasownika *być*);
- forma przyszła prosta (tag: fut_simp) – czasowniki w formie osobowej w znaczeniu czasu przyszłego (odnosi się to też do czasownika *być*);
- wykładnik czasu przyszłego BYĆ (tag: fut) – formy osobowe czasu przyszłego czasownika *być* pełniące funkcję wykładnika czasu przyszłego. Co istotne, forma taka wchodzi w skład konstrukcji złożonej z wykładnikami czasu przyszłego i bezokolicznika lub pseudoimiesłowu;
- wykładnik czasu zaprzeszłego BYĆ (tag: plusq) – formy osobowe czasu przeszłego pełniące funkcję wykładnika czasu zaprzeszłego. W tym przypadku formy te również wchodzi w skład konstrukcji złożonej;
- aoryst (tag: aor) – formy aorystyczne czasownika *być*;
- rozkaznik (tag: imp) – formy czasownika w trybie rozkazującym, zarówno z wykładnikiem zerowym, jak i z wykładnikiem *-i*;

- bezokolicznik (tag: inf) – nieosobowe formy czasownika wchodzące w skład różnych konstrukcji składniowych;
- winien/powinien (tag: obl) – czasowniki wyrażające powinność, które mają charakterystyczne cechy gramatyczne i składniowe. Są nimi czasowniki *winię* i *powinię*;
- predykatyw (tag: pred) – ośrodki predykcji takie jak *można*, *trzeba*;
- aglutynant BYĆ (tag: aglt) – ruchoma końcówka osobowa dołączana do pseudoimiesłowu lub czasownik *być* w formie osobowej;
- aglutynant aorystu BYĆ (tag: aglt_aor) – osobowa końcówka aorystyczna do-dawana do form trybu przypuszczającego oraz czasu przeszłego.

Osobną grupę tworzą imiesłowy, które podzieliliśmy na następujące klasy:

- pseudoimiesłów (tag: praet) – tak zwany imiesłów elowy, który z czasem stał się samodzielnym wykładnikiem czasu przeszłego;
- imiesłów czasu teraźniejszego czynny nieodmienny (tag: pcon) – wszystkie nieodmienne imiesłowy czasu teraźniejszego czynne zakończone na *-ę*, *-ęcy*, *-ąc*. Jednostki te, pomimo iż historycznie są kontynuantami form nom i acc m i f, uznajemy za w pełni zleksykalizowane;
- imiesłów czasu teraźniejszego czynny odmienny (tag: pact) – imiesłowy współ-cześnienazywaneprzymiotnikowymi, a więc formy zakończone na *-ący*, *-ąca*, *-ące*;
- imiesłów czasu przeszłego czynny (tag: pant) – formy zakończone na *-w*, *-wszy*, *-szy*;
- imiesłów bierny (tag: ppas) – formy proste i złożone imiesłowu biernego zakończone na *-n*, *-ny*, *-t*, *-ty*, a także formy czasownika zwane bezosobnikiem zakończone na *-no*, *-to*.

Pozostałe klasy obejmują elementy tradycyjnie uznawane za nieodmienne:

- przyimek (tag: prep) – do tej grupy należą wszystkie przyimki. Formy wokaliczne i niewokaliczne traktujemy jako osobne elementy, dlatego mają one osobne lematy;
- spójnik (tag: conj) – w tej klasie znajdują się elementy językowe, które tradycyjnie rozumiane są jako spójniki, a więc wiążą zewnętrznie lub wewnętrznie człony jednej wypowiedzi;
- partykuła (tag: qub) – do tej klasy zaliczamy elementy o charakterze modalnym, które nie mają samodzielnego znaczenia. Pełnią one w tekście różne funkcje: wzmacniającą, pytającą, życzącą bądź przeczącą;
- interiekcja (tag: interj) – to grupa, do której przynależą formy wykrzyknikowe;
- zgramatykalizowane SIĘ (tag: gram_SIĘ) – w tej klasie uwzględniamy wyrazy tekstowe *się/sie/si* stanowiące wykładnik jakiejś informacji gramatycznej. Może on być morfemem słotwórczym, wykładnikiem zwrotności, bierności lub bezosobowości;
- aglutynant trybu przypuszczającego BY (tag: aglt_cond) – do tej grupy zaliczany jest wykładnik trybu przypuszczającego *by*.

Osobną klasę stanowią wyrazy obce (tag: xen), czyli wszystkie wyrazy, które nie są zaadaptowane do polszczyzny. Co istotne, nie dokonujemy anotacji gramatycznej wyrazów z tej klasy.

Przy anotacji zostały uwzględnione przede wszystkim kategorie fleksyjne oraz te kategorie słownikowe, które mają wykładniki formalne, morfologiczne. Przykładem może być rodzaj rzeczownika. Nie wyodrębniamy natomiast aspektu przy czasownikach, ponieważ jest to kategoria czysto słownikowa, niemająca charakteru morfologicznego. Kategorie gramatyczne są zależne od danej klasy gramatycznej, tzn. każda klasa ma przypisany zbiór kategorii (patrz: załączona w aneksie tabela).

Poniżej zamieszczamy wykaz kategorii oraz wartości, które mogą one przyjąć.

Liczba:

- liczba pojedyncza (tag: sg),
- liczba podwójna (tag: du),
- liczba mnoga (tag: pl).

Wartości liczby mogą przyjmować rzeczowniki, przymiotniki, niektóre liczebniki i zaimki, większość form czasownikowych, pseudoimiesłowy, imiesłowy czasu teraźniejszego czynne odmienne i imiesłowy bierne.

Osoba:

- osoba pierwsza (tag: pri),
- osoba druga (tag: sec),
- osoba trzecia (tag: ter).

Wartości osoby określamy w przypadku czasowników w formie teraźniejszej i w formie przyszłej prostej, czasownika *być* w funkcji wykładnika czasu przyszłego i w formie aorystycznej, rozkazników oraz aglutynantów.

Przypadek:

- mianownik (tag: nom),
- dopełniacz (tag: gen),
- celownik (tag: dat),
- biernik (tag: acc),
- narzędnik (tag: inst),
- miejscownik (tag: loc),
- wołacz (tag: voc).

Kategoria przypadku jest przynależna rzeczownikom, przymiotnikom, zaimkom rzeczowym, zaimkom przymiotnym, liczebnikom głównym, porządkowym, zbiorowym, niektórym liczebnikom krotnym, wielorakim i ułamkowym, zaimkom liczebnym, imiesłowom czasu teraźniejszego czynnym odmiennym i imiesłowom biernym. Przyporządkowanie danego przypadku do formy w tekście jest zależne od kontekstu składniowego, co ma szczególną wagę w sytuacji przypadków synkretycznych.

Rodzaj:

- rodzaj męski (tag: m),

- rodzaj żeński (tag: f),
- rodzaj nijaki (tag: n),
- rodzaj niemęski (tag: nm).

Nie zdecydowaliśmy się na wprowadzenie kategorii męskoosobowości, ponieważ taka kategoria w staropolszczyźnie dopiero się kształtowała. Zamiast tego wprowadzamy wartość kategorii *rodzaj niemęski*. Ma to na celu oddanie zróżnicowania form rodzajowych bez jednoczesnego przypisywania staropolszczyźnie kategorii wykształconej dużo później. Rodzaj określamy w następujących klasach: rzeczownik, przymiotnik, zaimek rzeczowny, zaimek przymiotny, zaimek liczebny, liczebnik, wykładnik czasu zaprzeszłego BYĆ, formy *winien/powinien*, pseudoimiesłów, imiesłów czasu teraźniejszego czynny odmienny, imiesłów bierny. Wartość w tej kategorii wybieramy na podstawie formy tekstowej lub kontekstu składniowego.

Stopień – nie jest on wprowadzany jako kategorią gramatyczną, został jednak wyodrębniony ze względu na wygodę wyszukiwania. Ze względu na to, że w staropolszczyźnie czasem trudno odróżnić stopień wyższy od najwyższego, za Krystyną Kleszczową (2004) zdecydowano się na wprowadzenie tylko dwóch wartości:

- równy (tag: pos),
- porównawczy (tag: comp).

Kategorię stopnia przyjmują przymiotniki i przysłówki.

Aglutynacja:

- nieaglutynowany (tag: nagl),
- aglutynowany (tag: agl).

Aglutynację przyjmują jedynie pseudoimiesłowy. Jest ona rozumiana nieco inaczej niż w innych narzędziach korpusowych. Za aglutynowane uznajemy wszystkie pseudoimiesłowy, które wraz z innymi elementami tworzą osobne formy fleksyjne (np. czas zaprzeszły, czas przeszły złożony, tryb przypuszczający itd.). Jest to związane z przesłankami zarówno historycznojęzycznymi (analizowane przez nas teksty pochodzą z okresu, w którym m.in. stabilizuje się końcówka fleksyjna, a zanika wyraz posiłkowy), jak i narzędziowymi (anotacja konstrukcji nieciągłych, por. część *Druga anotacja...*). Zagadnieniu temu zostanie poświęcony odrębny artykuł.

Ze względu na specyfikę materiału nie zawsze było możliwe wskazanie wartości niektórych kategorii dla wybranych klas. Gdy dana kategoria nie dotyczy anotowanej formy, np. kategoria stopnia nie obejmuje niestopniowalnych przymiotników (typu *Piotrowy* czy *drewniany*), jej wartość jest oznaczana jako o. To samo dotyczy np. rodzaju zaimków rzeczownych dla 1. i 2. osoby. Wartość × natomiast oznacza, że nie można określić danej kategorii z powodu czynników wewnętrznych (np. kategoria liczby i rodzaju liczebników głównych w mianowniku, gdy nie wiadomo, czy mamy do czynienia ze starą czy nową odmianą) lub zewnętrznych (uszkodzenie tekstu). Rozwiązanie to wynika z przyczyn zarówno narzędziowych, jak i merytorycznych. Temat ten będzie przedmiotem osobnego opracowania.

Druga anotacja – czym jest i jak działa?

Jednym z narzędzi służących wyszukiwaniu odpowiednich form i konstrukcji w anotowanym korpusie jest składnia zapytań¹⁶. Z technicznego punktu widzenia przy przyjętym przez nas zestawie klas i kategorii jej stosowanie byłoby trudne dla użytkownika. Zdecydowano się zatem na inne rozwiązanie. By umożliwić użytkownikom odnalezienie w materiale zarówno prostych, jak i złożonych form czasów i trybów, materiał poddano dodatkowej anotacji. Na poziomie tzw. drugiej anotacji oznaczamy formy samodzielne, złożone tylko z jednego elementu, np. *zrobię* (ZROBIĆ futurum), *czytała* (CZYTAĆ praeteritum), *em* (BYĆ praesens), a także formy, które mogą składać się z więcej niż jednego elementu morfologicznego, np. pseudoimiesłowu i aglutynantu trybu przypuszczającego (*dałby* – DAĆ BYĆ conditionalis) lub wykładnika czasu przyszłego i bezokolicznika (*będę pisać* – BYĆ PISAĆ futurum). Formy dodatkowo zaanotowane użytkownik może wyszukać za pomocą wyszukiwarki *Czas i tryb*.

Wprowadzenie dwóch poziomów anotacji jest rozwiązaniem, które umożliwia wyszukiwanie zarówno np. samych pseudoimiesłów, jak i wszystkich konstrukcji, które mogą one tworzyć, np. czasu przeszłego złożonego, czasu zaprzeszłego itd. (wyszukiwanie czasów i trybów – zob. więcej w części *Funkcjonalności...*).

Na poziomie drugiej anotacji zostały zaanotowane następujące formy:

- czas teraźniejszy (tag: praesens),
- czas przeszły (tag: praeteritum),
- czas zaprzeszły (tag: plusquamperfectum),
- czas przyszły (tag: futurum),
- tryb rozkazujący (tag: imperativus),
- tryb przypuszczający (tag: conditionalis).

W wypadku form złożonych (w językoznawstwie korpusowym tradycyjnie nazywanych formami nieciągłymi) poszczególne elementy mogą być od siebie znacznie oddalone i występować w dowolnej kolejności. Są one wyświetlane w takim szyku, w jakim wystąpiły w tekście, np. dla konstrukcji *iżbych podniósł* forma złożona czasu przeszłego pojawi się jako *ch podniósł*.

Na poziomie drugiej anotacji nie powtarzamy informacji o wartościach kategorii osoby, liczby i rodzaju. Jest ona przejmowana za formami opisanymi na poziomie anotacji gramatycznej, np. dla czasu przeszłego złożonego z pseudoimiesłowu i wykładnika czasu zaprzeszłego informacja o osobie jest przejęta z wartości przypisanych na poziomie anotacji gramatycznej plusq, a o liczbie i rodzaju – z praet.

16 Por. np. KorBa, NKJP.

Funkcjonalności przydatne w badaniach językoznawczych

Tak opracowana anotacja gramatyczna jest podstawą działania trzech funkcjonalności narzędzia dostępnego na stronie *Apocrypha*: wyszukiwarki gramatycznej i leksykalnej oraz komentarza gramatycznego.

W podstawowym polu pracy z tekstem oprócz możliwości wyświetlenia transkrypcji, transliteracji, a także zestawienia innych transkrypcji, źródeł i wątków, możliwe jest również wyświetlenie komentarza gramatycznego dla każdego wyrazu tekstowego. Użytkownik otrzymuje informacje o przypisanym lemacie, klasie i wartościach kategorii. Kolejność, w jakiej wyświetlane są informacje gramatyczne, jest stała i odpowiada podanemu wyżej porządkowi klas i kategorii. Jeśli dany wyraz wchodzi w skład formy anotowanej także na poziomie drugiej anotacji, zostaje to uwzględnione w komentarzu gramatycznym.

Możliwe jest wyświetlenie komentarza gramatycznego dla wszystkich jednostek tekstu, które zostały objęte anotacją gramatyczną, a więc także dla wyrazów nieodmiennych i tych uznanych za obce, dla których wyświetlana jest tylko informacja o klasie i lemacie.

Kolejną z ważnych dla językoznawców funkcjonalności narzędzia dostępnego na stronie *Apocrypha* jest wyszukiwanie dowolnej jednostki leksykalnej (zakładka *Słowa*)¹⁷. W tej części narzędzia są prezentowane wszystkie „słowa”, przez które rozumiemy lematy ułożone w kolejności alfabetycznej. Po wpisaniu konkretnego lematu otrzymane wyniki można filtrować, zawężając pole eksploracji do określonych tekstów apokryficznych (filtr *Tekst*) bądź do określonej klasy (filtr *Klasa*).

Możliwe jest także wykorzystanie wyszukiwarki w badaniach słotwórczych: po wpisaniu odpowiedniej kombinacji znaków (por. instrukcja wyszukiwania zamieszczona na stronie *Apocrypha*) oraz cząstek. Wynikiem takiego wyszukiwania jest lista lematów, np. zapytanie **ciel* wyświetli wszystkie lematy zawierające formant *-ciel*.

Po wybraniu konkretnego lematu wyświetlą się jego wszystkie użycia tekstowe wraz z lokalizacją, komentarzem gramatycznym i kontekstem. Wybranie jednostkowego wyniku przenosi do pola pracy z tekstem. Wyświetli się także lista wariantów fonetycznych wybranego lematu.

W narzędziu badawczym możliwe jest także znalezienie wszystkich form gramatycznych danego słowa bądź form gramatycznych danej klasy wyrazów, które wystąpiły w tekstach apokryfów. Służy temu wyszukiwarka gramatyczna (zakładka *Formy gramatyczne*). Daje ona możliwość szukania form na kilka sposobów:

- 1) Po lemacie (po wpisaniu lematu w pasku wyszukiwania *Szukaj tylko w obrębie lematu* otrzymujemy wszystkie formy wyrazowe przynależące do danego lematu).

17 Zob. *O projekcie, Instrukcje* na stronie apocrypha.amu.edu.pl (dostęp: 26 IV 2022).

- 2) Za pomocą wybieranych filtrów profilujących zakres wyszukiwania do określonych tekstów i/lub zjawisk gramatycznych. Interesująca użytkownika forma gramatyczna może zostać znaleziona po klasie, liczbie, osobie, przypadku, rodzaju, stopniu, aglutynacji. Możliwe jest też dalsze precyzowanie i zawężanie wyniku wyszukiwania poprzez wybór kolejnych filtrów (odpowiadających nazwom poszczególnych klas – w wypadku wyszukiwania za pomocą klasy – bądź wartościom kategorii gramatycznych – w wypadku wyszukiwania za pomocą kategorii).
- 3) Poprzez połączenie obu tych sposobów.

Najistotniejszy jest fakt, że wyszukiwarka ta pozwala niezwykle precyzyjnie odnaleźć interesującą nas formę (przy jednoczesnym wyborze filtrów odpowiadających różnym kategoriom) we wszystkich lub w wybranych tekstach. Użytkownik zainteresowany na przykład repartycją końcówek celownika rzeczowników rodzaju męskiego lub form liczby pojedynczej rozkaznika czasowników może zawęzić wyszukiwanie wyłącznie do tej formy gramatycznej¹⁸. W ten sam sposób działa wyszukiwarka form czasów i trybów, anotowanych na poziomie drugiej anotacji. Tu dodatkowo użytkownik może zawęzić wyniki wyszukiwania do konkretnego sposobu budowania czasu i trybu, np. wybrać tylko formy trybu rozkazującego złożone z czasu teraźniejszego i partykuły.

Obie te wyszukiwarki umożliwiają także prowadzenie badań frekwencyjnych dzięki licznikom znajdującym się przy poszczególnych filtrach.

Podsumowanie

Anotacja gramatyczna na każdym etapie – od planowania siatki klas i kategorii oraz projektowania wyszukiwarek po ręczną anotację – wymaga uwzględnienia dwóch kwestii: specyfiki materiału staropolskiego oraz wymogów i ograniczeń technicznych. Staraliśmy się oddać właściwości języka, w którym granice między klasami gramatycznymi były płynne, niektóre kategorie się kształtowały, inne zanikały, a paradygmaty poszczególnych leksemów dopiero się stabilizowały. Pomimo że w narzędziu nie anotujemy poziomu składniowego, związki składniowe i zmiany w ich obrębie były kolejnymi czynnikami, które musieliśmy wziąć pod uwagę, analizując poszczególne formy. Ponadto staropolskie apokryfy Nowego Testamentu nie tylko są tekstami pochodzącymi z różnych okresów (od najstarszego *Listu Lentulusa* po XVI-wieczne *Rozmyślenia dominikańskie*), lecz także same przechowują warstwy językowe pochodzące z różnych etapów rozwoju polszczyzny¹⁹. Wszystkie te kwestie musieliśmy uwzględnić, jednocześnie jednak anotacja gramatyczna wymagała pewnych jednoznacznych rozstrzygnięć, np. w zakresie granic klas i kategorii. Zależało

¹⁸ Dotychczas tego typu badania prowadzone były ręcznie, por. np. Osiewicz 2020, 2021.

¹⁹ Por. np. Rzepka, Twardzik 1996a, 1996b, 1997a, 1997b, 1997c, 1999.

nam także na tym, by użytkownicy mieli możliwość wszechstronnego przeszukiwania materiału. Ostateczny kształt anotacji gramatycznej – zarówno z perspektywy anotatora, jak i użytkownika – jest wypadkową tych wszystkich czynników. Sposób anotowania wynika z naszych badań nad materiałem apokryficznym i jest podstawą stworzonego narzędzia, które jednocześnie otwiera kolejne perspektywy badawcze.

Tabela 1. System klas i kategorii

kategoria klasa	liczba	osoba	przy- padek	rodzaj	stopień	aglutynacja
rzeczownik	+		+	+		
przymiotnik	+		+	+	+	
przysłówek					+	
zaimek rzeczowny	+/-		+	+/-		
zaimek przymiotny	+/-		+	+/-		
zaimek liczebny			+			
zaimek przysłowny						
liczebnik główny	+/-		+	+/-		
liczebnik porządkowy	+		+	+		
liczebnik zbiorowy	+/-		+	+/-		
liczebnik zespołowy	+/-		+/-	+/-		
liczebnik krotny						
liczebnik wieloraki	+		+	+		
liczebnik ułamkowy						
forma teraźniejsza	+	+				
forma przyszła prosta	+	+				
wykładnik czasu przyszłego BYĆ	+	+				
wykładnik czasu zaprzeszłego BYĆ	+	+		+		
aoryst	+	+				
rozkaznik	+	+				
bezokolicznik						
<i>winien/powinien</i>	+	+		+		
predykatyw						
aglutynant BYĆ	+	+				
aglutynant aorystu BYĆ	+	+				
pseudoimiesłów	+			+		+

kategoria klasa	liczba	osoba	przy- padek	rodzaj	stopień	aglutyna- cja
imiesłów czasu teraźniejszego czynny nieodmienny						
imiesłów czasu teraźniejszego czynny odmienny	+		+	+		
imiesłów czasu przeszłego czynny						
imiesłów bierny	+		+	+		
przyimek						
spójnik						
partykuła						
interiekcja						
zgramatyzalizowane SIE						
aglutynant trybu przypuszczającego BY						
obce						

Źródło: opracowanie własne.

Literatura

- BRÜCKNER A., 1904, *Literatura religijna w Polsce średniowiecznej*, t. 3: *Legends i modlitewniki. Szkice literackie i obyczajowe*, „Biblioteka Dzieł Chrześcijańskich”, z. 40, Warszawa.
- BRÜCKNER A. (wyd.), 1907, *Rozmyślanie o żywocie Pana Jezusa. Z rękopisu grecko-katolickiej kapituły przemyskiej*, „Biblioteka Pisarzy Polskich”, nr 54, Kraków.
- DANIELEWICZOWA M., 2012, *W głąb specjalizacji znaczeń. Przysłówkowe metapredykaty atestacyjne*, Warszawa.
- INSTR: W. Gruszczyński, R. Bronikowska, *Instrukcja korzystania z wyszukiwarki do Elektronicznego Korpusu Tekstów Polskich z XVII i XVIII wieku (do 1772 r.)*, [on-line:] <https://korba.edu.pl/manual>.
- KLESZCZOWA K., 2004, *Kategoria stopnia staropolskich przymiotników*, [w:] H. Kurek, J. Labocha (red.), *Studia linguistica Danutae Wesołowska oblata*, Kraków, s. 85–92.
- KLESZCZOWA K., 2015, *U źródeł polskich partykuł. Derywacja funkcjonalna, przemiany, zaniżki*, „Prace Naukowe Uniwersytetu Śląskiego w Katowicach”, nr 3368, Katowice.
- KLESZCZOWA K., SZCZEPANEK A. (red.), 2014, *Wyrażenia funkcyjne w perspektywie diachronicznej, synchronicznej i porównawczej*, „Prace Naukowe Uniwersytetu Śląskiego w Katowicach”, nr 3141, Katowice.
- KORBA: Elektroniczny Korpus Tekstów Polskich XVII i XVIII w. (do 1772 r.), [on-line:] korba.edu.pl.
- KRĄŻYŃSKA Z., MIKA T., 2007, *Transkrypcja a granice interpretacji*, [w:] R. Laskowski, R. Mazurkiewicz (red.), *Amoenitates vel lepores philologiae*, Kraków, s. 160–170.

- KRĄŻYŃSKA Z., MIKA T., SŁOBODA A., 2015, *Składnia średniowiecznej polszczyzny*, cz. 1: *Konteksty – metody – tendencje*, Poznań.
- MAZURKIEWICZ R. (wstęp, wybór, oprac.), 2000, *Teksty o Matce Bożej. Polskie średniowiecze*, „Beatam me Dicent...”, t. 12, Niepokalanów.
- MAZURKIEWICZ R. (oprac.), s.a., *Rozmyślania dominikańskie (fragmenty)*, [on-line:] http://www.staropolska.pl/średniowiecze/biblia_i_apokryfy/dominikanskie.html.
- MIKA T., ROJSZCZAK-ROBIŃSKA D., 2016, *Kanon wiedzy o Rozmyślaniu przemyskim*, [w:] J. Bartmiński, A. Timofiejew (red.), *Rozmyślanie przemyskie. Świadectwo średniowiecznej kultury religijnej*, Przemysł, s. 13–24.
- MIKA T., SŁOBODA A., 2015, *Wyrażenia funkcyjne w średniowiecznej polszczyźnie z perspektywy składniowej. Wybrane problemy badawcze*, Poznań.
- MIKA T., ZIÓŁKOWSKA O., 2017, *Jak wprowadzać interpunkcję do tekstu staropolskiego?*, [w:] K. Borowiec, D. Maslej, T. Mika, D. Rojszczak-Robińska (red.), *Jak wydawać teksty dawne*, „Staropolskie Spotkania Językoznawcze”, t. 2, Poznań, s. 105–125.
- Narodowy Korpus Języka Polskiego, praca zbiorowa, red. A. Przepiórkowski, M. Bańko, R.L. Górski, B. Lewandowska-Tomaszczyk, Warszawa 2012.
- NKJP: Narodowy Korpus Języka Polskiego, [on-line:] nkjp.pl.
- OSIEWICZ M., 2020, *Formy celownika, narzędnika i miejscownika liczby mnogiej rzeczowników w Sprawie chędogiej o męce Pana Chrystusowej i w Historii Trzech Kroli (kodeks Wawrzyńca z Łaska, 1544)*, „Biuletyn Polskiego Towarzystwa Językoznawczego” LXXVI, s. 243–255.
- OSIEWICZ M., 2021, *Formy dopełniacza i celownika liczby pojedynczej rzeczowników rodzaju męskiego w Sprawie chędogiej o męce Pana Chrystusowej i Ewangelii Nikodema (kodeks Wawrzyńca z Łaska, 1544). Przyczynek do chronologizacji*, „Język Polski” CI, s. 89–99.
- PAWELEC R. (red.), 2015, *Słownik zapomnianych wyrażen funkcyjnych*, Warszawa.
- ROJSZCZAK-ROBIŃSKA D., 2017, Karta Rogawskiego. *Zapomniany zabytek języka staropolskiego*, „Slavia Occidentalis” 74/1, s. 119–134.
- ROJSZCZAK-ROBIŃSKA D., 2021, *Staropolscy apokryfiści o swoich źródłach. Strategie cytowania*, [w:] D. Rojszczak-Robińska, A. Deskur, W. Stelmach (red.), *Źródła staropolskich apokryfów. Pytania, problemy, perspektywy*, Poznań, s. 131–158.
- RZEPKA W.R., TWARDZIK W.B., 1996a, *Archaizmy fleksyjne w Rozmyślaniu przemyskim*, cz. 1, „Język Polski” LXXVI, s. 97–102.
- RZEPKA W.R., TWARDZIK W.B., 1996b, *Archaizmy fleksyjne w Rozmyślaniu przemyskim*, cz. 2, „Język Polski” LXXVI, s. 321–327.
- RZEPKA W.R., TWARDZIK W.B., 1997a, *Archaizmy fleksyjne w Rozmyślaniu przemyskim*, cz. 3, „Język Polski” LXXVII, s. 41–46.
- RZEPKA W.R., TWARDZIK W.B., 1997b, *Archaizmy fleksyjne w Rozmyślaniu przemyskim*, cz. 4, „Język Polski” LXXVII, s. 102–111.
- RZEPKA W.R., TWARDZIK W.B., 1997c, *Archaizmy fleksyjne w Rozmyślaniu przemyskim*, cz. 5, „Język Polski” LXXVII, s. 298–302.
- RZEPKA W.R., TWARDZIK W.B., 1999, *Archaizmy fleksyjne w Rozmyślaniu przemyskim*, cz. 6, „Język Polski” LXXIX, s. 35–46.
- RZEPKA W.R., WYDRA W. (wyd.), ADAMCZYK M. (wstęp), 1996, *Cały świat nie pomieściłby ksiąg. Staropolskie opowieści i przekazy apokryficzne*, Warszawa – Poznań.
- SOBOTKA P., 2014a, *Od zaimka do wyrażenia funkcyjnego. Dzieje prastłowiańskiego zaimka *vъs-akъ i jego kontynuantów*, [w:] A. Moroz, P. Sobotka, M. Żabowska (red.), *Maiuscula*

- linguistica. Studia in honorem professori Matthiae Grochowski sextuagesimo quinto dedicata*, Warszawa, s. 141–161.
- SOBOTKA P., 2014b, *Partykuły greckie – pochodzenie i wybrane problemy opisu*, [w:] K. Kleszczowa, A. Szczepanek (red.), *Wyrażenia funkcyjne w perspektywie diachronicznej, synchronicznej i porównawczej*, „Prace Naukowe Uniwersytetu Śląskiego w Katowicach”, nr 3141, Katowice, s. 133–167.
- SOBOTKA P., ŻABOWSKA M., 2017, *Wyodrębnianie, dekodowanie i klasyfikacja historycznych jednostek języka*, „LingVaria” nr 2 (24), s. 113–133, <https://doi.org/10.12797/LV.12.2017.24.09>.
- STRAMCZEWSKA O., 2014, *Różnice w interpretacji składni zabytku a interpretacja transkrypcji na przykładzie Rozmyślenia przemyskiego*, „Roczniki Humanistyczne” LXII, z. 6, s. 161–171, [on-line:] <https://ojs.tnku.pl/index.php/rh/article/view/6158/5953>.
- TWARDZIK IJP PAN: W. Twardzik (oprac.), s.a, *Rozmyślanie przemyskie*, [on-line:] <https://ijp.pan.pl/wp-content/uploads/2018/10/Rozm.pdf>.
- TWARDZIK W. (oprac.), s.a. „Wypisy z Opeca”, [on-line:] http://www.staropolska.pl/srednio-wieczne/biblia_i_apokryfy/Twardzik_Opec.html.
- WAJSZCZUK J., 2010, *Functional Class (so Called ‘Part of Speech’) Assignment as a Kind of Meaning-Bound Word Syntactic Information*, „Studia Kognitywne – Études Cognitives” nr 10, s. 15–33.
- WYDRA W., RZEPKA W.R. (wybór i oprac.), KURASZKIEWICZ W. (red.), 1984/2004, *Chrestomatia staropolska. Teksty do roku 1543*, wyd. 3, Wrocław.
- WYDRA W., WÓJCIK R. (wyd.), KRZAK-WEISS K. (wstęp ikonograficzny), 2014, *Żywot Pana Jezusa Krysta Baltazara Opeca*, Poznań.
- ZIÓŁKOWSKA O., 2016, *Specyfika języka Rozmyślań dominikańskich w świetle składni i leksyki*, Poznań.
- ŻABOWSKA M., 2015, *O godzeniu przeciwieństw – funkcja wyrażen jednocześnie i zarazem w języku naturalnym oraz w testowaniu hipotez semantycznych*, „Linguistica Copernicana” 12, s. 71–105, <https://doi.org/10.12775/LinCop.2015.003>.

Grammatical Annotation of an Old Polish Text: Problems and Challenges

Abstract

The paper is focused on grammatical annotation of Old Polish texts. The authors discuss the problems faced by the annotator in creating an annotation tool suitable for the analysis of texts written during the period of intensive changes in the Polish language. The article presents the main assumptions of grammatical annotation developed under the grant: *The origins of Polish language and religious culture in the light of the medieval apocrypha of the New Testament. A universal tool for the study of Polish apocryphal texts*, as well as the system of classes and categories developed in the course of the research on Old Polish apocrypha. The functionalities of the tool for the linguistic research are also presented: a grammatical and lexical search engine and a grammatical commentary.