

ALEKSANDRA WIECZOREK

Instytut Języka Polskiego Polskiej Akademii Nauk

ORCID: 0000-0002-8829-559X

Metody badań nieciągłego szyku przydawki przymiotnej w Średniopolskim Korpusie Zależnościowym

A b s t r a k t: Artykuł przedstawia rozważania na temat metod badań nad nieciągłym szykiem przydawki w zdaniach siedemnasto- i osiemnastowiecznych na materiale korpusu anotowanego składniowo w schemacie zależnościowym. Duży materiał, oznakowany pod względem morfosyntaktycznym oraz składniowym, daje nowe możliwości analiz, jednak wymaga też opracowania sposobu ekskserpcji i klasyfikacji badanych konstrukcji. Do porównania wykorzystano materiał współczesny z Polskiego Banku Drzew Zależnościowych.

S ł o w a k l u c z e: nieciągły związek składniowy; przydawka przymiotna; elektroniczny korpus zależnościowy; polszczyzna XVII i XVIII wieku

Wstęp

Związek składniowy jest nieciągły, gdy przerywa go składnik nienależący do niego strukturalnie (por. Szpakowicz i Świdziński 1992: 183; Świdziński 1996: 113; Derwojedowa 2000: 117). Teksty polskie z XVII i XVIII wieku w powszechnym mniemaniu cechuje częstsze niż w tekstach współczesnych rozbijanie związków składniowych przez inne elementy zdania. Celem artykułu jest weryfikacja tego twierdzenia na materiale korpusowym. Do analizy wybrałam jeden typ zależności składniowej: dwuelementową grupę nominalną składającą się z nadrzędnika i przydawki przymiotnej. Przydawka przymiotna może być wyrażona przymiotnikiem, zaimkiem

przymiotnym, liczebnikiem przymiotnikowym lub imiesłowem przymiotnikowym (por. Gębka-Wolak 1997: 12); jest uzgodniona z nadrzędnikiem pod względem liczby, przypadku i rodzaju (nieliczne wyjątki zostaną opisane dalej). W artykule będę używała określeń „przydawka” bądź „przydawka przymiotna” w tym zawężonym znaczeniu. Wyłączam zatem z badań przydawkę dopełniaczową, apozycyjną oraz w postaci wyrażenia przyimkowego. Nadrzędnikiem tak rozumianej przydawki, czyli reprezentantem grupy nominalnej, może być rzeczownik, zaimek osobowy, gerundium bądź liczebnik główny lub zbiorowy (por. Szpakowicz i Świdziński 1990: 10–11; Gębka-Wolak 1997: 6). W szczególnych przypadkach może to też być przymiotnik lub inna część mowy; zagadnienie to zostanie rozwinięte dalej.

Grupa nominalna może nie zawierać żadnej przydawki lub zawierać ich kilka. Szczególny typ grupy nominalnej wybrany do badań, czyli dwuskładnikową grupę nominalną z przydawką przymiotną, będę dla uproszczenia nazywała „parą nadrzędnika z przydawką”. W cytacie *przerwali wesole swe pienia* mamy jedną grupę nominalną (*wesole swe pienia*) oraz dwie pary nadrzędnika z przydawką (*wesole pienia* oraz *swe pienia*).

Artykuł jest w przeważającej mierze poświęcony metodzie kolekcjonowania i analizy materiału. Najważniejsze pytania, na które będę szukać odpowiedzi w tym materiale, to: czy rozbieżność grup przydawkowych jest rzeczywiście częstsza w tekstach z XVII i XVIII wieku niż we współczesnych; czy można wskazać jakieś czynniki gramatyczne współwystępujące z rozbijaniem badanych par przydawkowych; czy w obrębie badanej epoki można zaobserwować różnice, jeśli chodzi o nieciągłe związki nadrzędnika z przydawką, związane z czynnikami pozatekstowymi (np. czas powstania tekstu, rodzaj tekstu, tematyka). Nie jest natomiast przedmiotem moich badań kolejność przydawek przymiotnych w jednej grupie nominalnej ani też antepozycyjne lub postpozycyjne położenie przydawki. Problemem wykraczającym poza ramy tego opracowania jest również stopień oddalenia nadrzędnika od podrzędnika.

Badania nad nieciągłym szykiem przydawki w polszczyźnie historycznej – wybór

Badacze polszczyzny XVII i XVIII wieku zwracają w swoich pracach uwagę na rozbijanie związków składniowych, w tym na odsuwanie przydawki

od wyrazu określanego, tak w odniesieniu do poezji, jak i do innych rodzajów tekstów (Wierzbicka 1964: 20; Karplukówna 1967: 43; Perczyńska 1978: 62; Skubalanka 1984: 178; Smolińska 1983: 120; Wiśniewska 1994: 129–130). Są to jednak ogólne obserwacje, poparte zaledwie kilkoma przykładami. Przyczyny występowania w tych tekstach nieciągłości składniowych upatruje się we wpływie klasycznej składni łacińskiej, która miała oddziaływać na polskie piśmiennictwo już w dobie Odrodzenia (Klemensiewicz 2009: 424 i 428; Wierzbicka 1966; Smolińska 1983: 120).

Szczególną uwagę poświęcił szykowi przydawki W. Śliwiński. Wykorzystując propozycje A. Wierzbickiej (1963 i 1964), badacz zaproponował schemat opisu różnych grup przydawkowych ze względu na układ występujących w nich wyrazów (Śliwiński 1982), a następnie zastosował ów schemat do porównania pod tym względem wybranych utworów Kochanowskiego i Morsztyna (Śliwiński 2006). Badanie to pokazało, że u poety barokowego częściej występuje oddalenie przydawki od jej nadrzędnika, a także przerywanie tego związku przez inne składniki zdania (Śliwiński 2006: 115). Najistotniejszym wskazanym przez A. Wierzbicką (za: Fitałow 1962) i rozwiniętym przez W. Śliwińskiego podziałem grup nominalnych z przydawką jest podział na grupy konfiguracyjne, czyli takie, w obręb których nie wchodzi składnik spoza tej grupy, oraz niekonfiguracyjne, czyli rozbite przez składnik z zewnątrz (Wierzbicka 1963 i 1964; Śliwiński 1982).

Material

Materiałem badawczym jest zbiór 2017 zdań oznakowanych składniowo w formacie zależnościowym, pochodzących ze Średniopolskiego Korpusu Zależnościowego¹ (pierwszą jego wersję opisałam w artykule: Wieczorek 2025). Zdania te pochodzą z tekstów siedemnasto- i osiemnastowiecznych, a każde z nich ma reprezentację w postaci tzw. drzewa zależności składniowych. W takim drzewie każdy segment² jest przyporządkowany

¹ Zasób powstał w ramach projektów: NPRH nr 0413/NPRH7/H11/86/2018 oraz NCN Miniatura nr 2023/07/X/HS2/00111.

² W rozumieniu NKJP (Przepiórkowski *et al.* 2012: 61–62) oraz KorBy (Gruszczyński *et al.* 2020: 33). Zgodnie z tą konwencją niektóre formy gramatyczne składają się z dwóch lub większej liczby segmentów (np. formy czasu przeszłego czasownika składające się z dawnego imiesłowu i tzw. aglutynantu, czyli ruchomej końcówki,

dokładnie jednemu nadrzędnikowi i oznakowany odpowiednią etykietą zależności składniowej. Element główny całego zdania ma przypisany arbitralny nadrzędnik *root*. Zdania w ŚKZ pochodzą z ręcznie anotowanego podkorpusu Elektronicznego Korpusu Tekstów Polskich z XVII i XVIII Wieku, znanego jako KorBa (por. Gruszczyński *et al.* 2022; www.korba.edu.pl). W ŚKZ nie ma zdań z tekstów pisanych wierszem ani zdań zawierających elementy obce. System znakowania relacji składniowych został zapożyczony z Polskiego Banku Drzew Zależnościowych (Polish Dependency Bank, dalej: PDB, <http://zil.ipipan.waw.pl/PDB>; Wróblewska, Rybak 2019). PDB również jest korpusem zdań oznakowanych składniowo w konwencji zależnościowej i zawiera 22 152 zdań, pochodzących przeważnie z Narodowego Korpusu Języka Polskiego. W moich badaniach PDB pełni rolę współczesnego materiału porównawczego.

Przykładowe drzewo zależności składniowych z ŚKZ przedstawia rys. 1.

Pozyskiwanie i porządkowanie materiału

W analizie zastosowałam metodę automatycznej ekscerpcji nieciągłych par z przydawką przymiotną z ŚKZ i PDB. W korpusach tych każdy wyraz ma przypisane informacje gramatyczne oraz numer porządkowy w zdaniu, a także informację o tym, którego wyrazu jest podrzędnikiem. To oznakowanie pozwoliło na wydobywanie interesujących mnie par oraz określenie odległości między nimi, liczonej w segmentach.

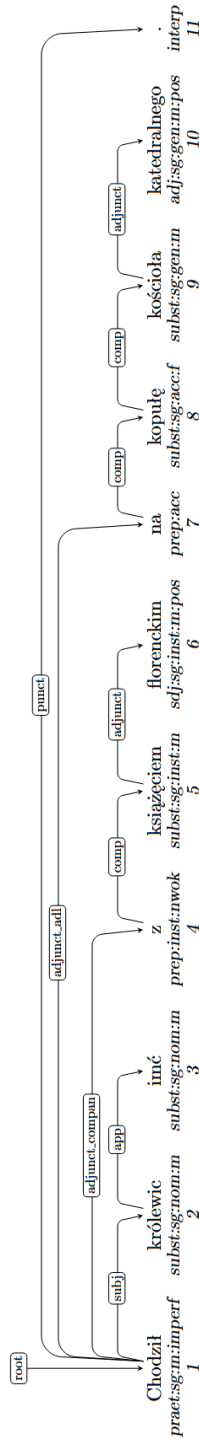
Wybór par: nadrzędnik-przydawka

Z korpusów zostały wydobyte pary wyrazów połączonych relacją nadrzędnno-podrzędną, spełniające następujące kryteria³:

np. *robił-em*). W zastosowanym schemacie składniowym elementy takich form gramatycznych są połączone strzałkami oznaczającymi zależności między nimi, jednak ich etykiety wskazują na to, że nie są to zależności składniowe, lecz morfologiczne, np. *aglt* dla aglutynantów.

³ Dziękuję Dorocie Komosińskiej za przeprowadzenie prac informatycznych.

Rysunek 1. Przykładowe drzewo zależności składniowych



Objaśnienie etykiet zależności składniowych: *root* – element główny zdania, *subj* – podmiot, *app* – przydawka apozycyjna, *adjunct_compan* – okolicznik towarzyszenia, *adjunct_adl* – okolicznik kierunku, *comp* – podrzędnik wymagany inny niż dopełnienie bliższe i dalsze rzeczownikowe, *adjunct* – przydawka. Wyjaśnienie wszystkich skrótów funkcji składniowych znajduje się na stronie <https://zil.ipipan.waw.pl/PDB/DepRelTypes>.

1. nadrzędnik należy do jednej z następujących klas gramatycznych⁴: rzeczownik (subst), przymiotnik⁵ (adj), liczebnik przymiotny (adjnum), skrót (brev), liczba (dig), gerundium (ger), liczebnik główny (num), liczebnik zbiorowy (numcol), zaimek pierwszo- lub drugoosobowy (ppron12), zaimek trzecioosobowy (ppron3), zaimek *siebie* (siebie);
2. podrzędnik należy do jednej z następujących klas gramatycznych: przymiotnik (adj), przymiotnik w formie odmiany niezłożonej (adjb), liczebnik przymiotny (adjnum), imiesłów przymiotnikowy bierny (ppas), imiesłów przymiotnikowy czynny (pact), imiesłów przeszły typu *zdumiały* ‘ten, który się zdumiał’ (ppraet), imiesłów przymiotnikowy czynny w formie odmiany niezłożonej (pactb), imiesłów przymiotnikowy bierny w formie odmiany niezłożonej (ppasb), imiesłów przeszły w formie odmiany niezłożonej (ppraetb);
3. nadrzędnik zgadza się z podrzędnikiem pod względem przypadku, liczby i rodzaju gramatycznego⁶ (z pewnymi zastrzeżeniami, które będą opisane dalej);
4. relacja łącząca te wyrazy ma etykietę *adjunct*.

Jak widać, wśród nadrzędników znalazły się zarówno rzeczowniki, jak i inne części mowy (przymiotniki, liczebniki przymiotnikowe, liczebniki główne i zbiorowe, zaimki osobowe, a także skróty i liczby). Przymiotniki i liczebniki przymiotne zostały tu uwzględnione, gdyż zastępują one rzeczownik w przypadku elipsy tego ostatniego, jak np. w zdaniu:

Kray zasłoneczny południowy, iest to tá część ziemi, ktorá się zamyka między cyrkulem solstycyonalnym południowym, y p o l a r n y m p o ł u d n i o w y m , dokoła szeroka ná gradusow 43.

⁴ Klasyfikacja na klasy gramatyczne jest bardziej szczegółowa niż klasyfikacja na części mowy (Bień, Saloni 1982). Segmentom w ŚKZ przypisane są klasy gramatyczne zgodnie z zasadami anotacji morfosyntaktycznej przyjętymi w KorBie, zob. <https://korba.edu.pl/manual>, stamtąd pochodzą również używane w ŚKZ znaczniki klas gramatycznych.

⁵ W rozumieniu KorBy, czyli również zaimek przymiotny; w kolejnych punktach tej listy będę używała terminu przymiotnik w tym właśnie znaczeniu.

⁶ Z tego powodu nie objęłam badaniem połączeń z zaimkami *co*, *coś*, *cóż*, *nic* typu *coś słodkiego*.

Natomiast przydawki ograniczyłam do przymiotników (w tym zaimków przymiotnych), liczebników przymiotnikowych oraz imiesłowów przymiotnikowych.

Dodatkowego wyjaśnienia wymaga punkt 3. Po pierwsze, w tekstach z XVII i XVIII w. zdarzają się jeszcze formy liczby podwójnej rzeczowników (np. *(dwie) bitwie*), natomiast nie notuje się już form liczby podwójnej przymiotników. Z tych powodów rozszerzyłam rozumienie zgody formalnogramatycznej na pary, gdzie jeden segment ma znacznik liczby mnogiej pl, a drugi – liczby podwójnej du, np. *bitwie*_{subst:du:acc:f} *okrutne*_{adj:pl:acc:f:pos} (por. Gruszczyński *et al.* 2022: 320). Drugi tego rodzaju wyjątek dotyczy znaczników rodzaju męskiego w niektórych przypadkach gramatycznych, które to znaczniki mają w KorBie uwzględniać kształtowanie się kategorii męskoosobowości. W bierniku liczby pojedynczej formy równe dopełniaczowi (np. *konia*) mają znacznik rodzaju manim2, natomiast formy równe mianownikowi (np. *wsiąść na k o Ń*) – znacznik m. W mianowniku, bierniku i wołaczu liczby mnogiej formy męskoosobowe (np. *królowie*) mają znacznik rodzaju manim1, natomiast niemęskoosobowe lub neutralne (np. *króle, wilki, mieszczanie*) – znacznik m (Gruszczyński *et al.* 2022: 320–321). Mogą się zatem zdarzyć pary rzeczownika z przydawką w tym samym rodzaju gramatycznym, gdzie jednak formalnie oznakowanie rodzajów nie jest identyczne, np. *pracowici*_{adj:pl:nom:manim1:pos} *mieszczanie*_{subst:pl:nom:m}. Z tych powodów przyjąłam, że jeśli przydawka przymiotna występuje w rodzaju manim1 lub manim2, a jej nadrzędnik w rodzaju m, to traktuję to jako zgodę formalnogramatyczną.

Są także wypadki, gdzie brak zgody formalnogramatycznej spowodował wykluczenie z badań par, które użytkownik języka polskiego mógłby uznać za mieszczące się w przyjętej przeze mnie definicji par nadrzędnik – przydawka. Nie było możliwe dołączenie ich do opisanej wyżej seryjnej grupy wyjątków, a na dodanie ręczne nie zdecydowałam się ze względu na duże ryzyko wystąpienia błędów. Należą tu następujące sytuacje (przykłady pochodzą z ŚKZ oraz PDB):

- przydawka w liczbie mnogiej określa nadrzędnik będący grupą szeregową, składającą się z elementów w liczbie pojedynczej, np. z *Augustyne* i *Hieronimem Świątymi*;
- przydawka w liczbie mnogiej określa nadrzędnik będący grupą szeregową, składającą się z elementów w liczbie mnogiej, ale

w różnych rodzajach, co musi powodować niezgodność rodzaju jednego z rzeczowników z przydawką, np.: *zalatnych*_{pl:gen:m3:pos} *słówek*_{pl:gen:n} i *żarcików*_{pl:gen:m2};

- przydawka w liczbie mnogiej określa nadrzędnik odnoszący się do bytu mnogiego, wyrażonego przez rzeczownik w liczbie pojedynczej z wyrażeniem przyimkowym, np. *Obórka ze stajenką w kupie, stare, także złe*;
- sytuacja, gdy liczebniki wchodzą w związki zgody z formami liczby pojedynczej rodzaju nijakiego przymiotników, np. *niecałe trzysta złotych*;
- zgoda semantyczna, np. *z Bracią Naszemi*;
- prawdopodobny błąd autora tekstu, drukarza bądź redaktora wydania uwspółcześnionego, np. *I para noży we srebro oprawne; odmienia się w wiele gwiazdy i promieni błyszczące*;
- błąd anotacji morfosyntaktycznej, np. *w czerwono-żółto-niebieskiej*_{sg:gen:f:pos} *kamizelce*_{sg:loc:f}.

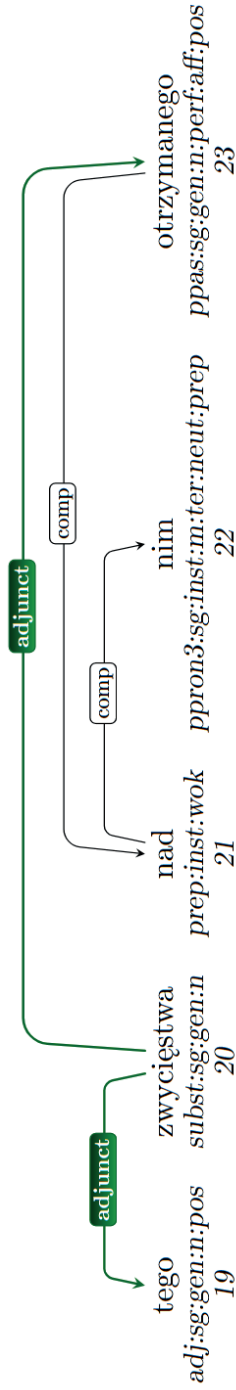
W ŚKZ wystąpiło ok. 20 tego rodzaju przypadków, a w PDB – ok. 90.

Wyszukiwanie relacji nieciągłych

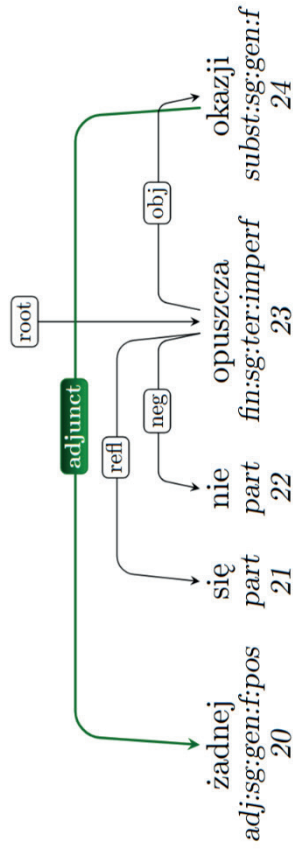
Numeracja segmentów w zdaniu pozwala określić liczbę segmentów stojących pomiędzy nadrzędnikiem a przydawką. W tym celu od numeru porządkowego przydawki (np. *tego*₁₉ oraz *otrzymanego*₂₃ na rys. 2) odejmujemy numer porządkowy nadrzędnika (*zwycięstwa*₂₀). W ten sposób otrzymujemy następujący wynik: odległość pomiędzy *tego* a *zwycięstwa* wynosi -1, a odległość między *zwycięstwa* a *otrzymanego* wynosi 3. Jak widać, liczby ujemne wskazują na antepozycyjny szyk przydawki, a liczby dodatnie – na szyk postpozycyjny. Przy poszukiwaniu nieciągłości będę brała pod uwagę odległość bezwzględna, czyli liczby ujemne zastępuję dodatnimi, co oznacza, że dla odległości -1 lub 1 odległość bezwzględna wynosi 1 itd.

Jest to sposób na wykrycie odległych relacji składniowych, jednak niekoniecznie nieciągłych. W następnym kroku należy zatem odróżnić sytuacje, gdzie pomiędzy nadrzędnikiem a przydawką znajdują się elementy tej samej grupy składniowej (jak na rys. 2) od tych, gdzie grupę rozdziela element należący do innej grupy składniowej, jak na rys. 3.

Rysunek 2. Fragment drzewa zależności składniowych zawierający parę *tego* zwyczajstwa oraz zwyczajstwa otrzymanego



Rysunek 3. Fragment drzewa zależności składniowych zawierający parę *żadnej okazji*.



Elementy, które mogą wystąpić pomiędzy nadrzędnikiem a przydawką, podzieliłam więc na następujące grupy:

1. para jest rozdzielona składnikiem będącym w relacji składniowej z jednym z elementów tej całości:
 - a) inna przydawka przymiotna, składnik, który nie mógłby wystąpić w innym miejscu (partykuła *-ż*, spójnik⁷) lub jest spoza poziomu języka (znak interpunkcyjny⁸, np. *potężnym srebrnym krzyżu, temu ż klasztorowi, sławnej jazdy/szlacheckiej*;
 - b) inne podrzędniki nadrzędnika lub przydawki, np. *w dalszym mowy przeciągu, ziele nam nieznajome, zwycięstwa nad nim otrzymanego*;
2. para jest rozdzielona składnikiem nienależącym do jej struktury, np. *żadnej się nie opuszcza okazji, dowody z uprzejmego płynące serca*.

Segmenty, które kwalifikują się do punktu 1., nie były brane pod uwagę podczas wyliczania skorygowanej odległości między nadrzędnikiem a przydawką. W praktyce od liczby, którą uzyskiwałam poprzez odjęcie numeru nadrzędnika od numeru przydawki, odejmowana była liczba odpowiadająca liczbie segmentów kwalifikujących się do punktu 1. (jeśli uzyskana początkowo odległość była ujemna, to należało dodać do niej liczbę elementów kwalifikujących się do punktu 1.). Zatem dla przykładu *zwycięstwa₂₀ nad₂₁ nim₂₂ otrzymanego₂₃*, gdzie początkowo wyliczona odległość między *zwycięstwa* a *otrzymanego* wynosiła 3, należy odjąć 2, gdyż segmenty *nad* i *nim* stanowią razem składnik podrzędny wobec przydawki. Ostatecznie odległość w tym przypadku wyniesie więc 1. Natomiast w przykładzie *żadnej₂₀ się₂₁ nie₂₂ opuszcza₂₃ okazji₂₄*, gdzie początkowo wyliczona odległość między *żadnej* a *okazji* wynosi -4, żaden z segmentów stojących pomiędzy nie jest (także pośrednio) podrzędny wobec nadrzędnika lub przydawki, a zatem ostateczna odległość pozostaje równa -4. Jeśli między nadrzędnikiem a przydawką odległość bezwzględna po skorygowaniu wynosi 2 lub więcej, to taką relację określam jako nieciągłą.

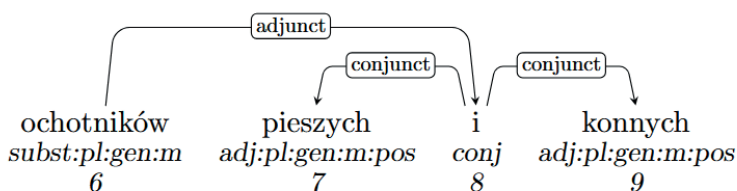
⁷ Obecność spójników w tym punkcie wyjaśnię w kolejnym paragrafie.

⁸ Znaki interpunkcyjne są traktowane jako samodzielne segmenty i w związku z tym mają swoje numery porządkowe w zdaniu.

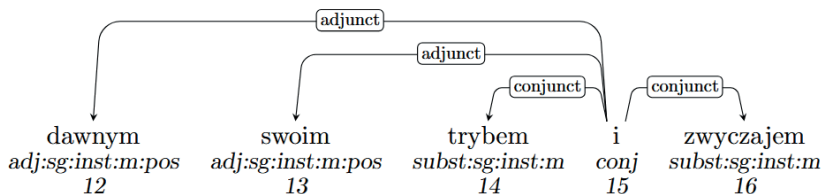
Przydawki lub nadrzędniki skoordynowane

Trudności metodologicznych przysparzają sytuacje, gdy nadrzędnik miał dwie lub więcej przydawek połączonych spójnikiem, np.: *ochotników pieszych i konnych*. W takich przypadkach zgodnie z założeniami PDB stosowane były następujące zasady anotacji zależnościowej: bezpośrednim podrzędnikiem określonego rzeczownika (tu: *ochotników*) jest spójnik bądź znak interpunkcyjny, oznakowany relacją *adjunct*, a przydawki są podrzędnikami tego spójnika bądź znaku interpunkcyjnego, z relacją *conjunct*; zob. rys. 4. Takich sytuacji jest w ŚKZ 411, a w PDB – 2 203. Analogiczne zasady anotacji są stosowane w przypadku koordynacji dwóch lub większej liczby nadrzędników, czyli spójnik jest traktowany jako reprezentant grupy, zob. rys. 5. Skoordynowanych nadrzędników przydawek przymiotnych jest w ŚKZ 67, natomiast w PDB – 219.

Rysunek 4. Fragment drzewa zależności składniowych zawierający pary *ochotników pieszych* i *ochotników konnych*

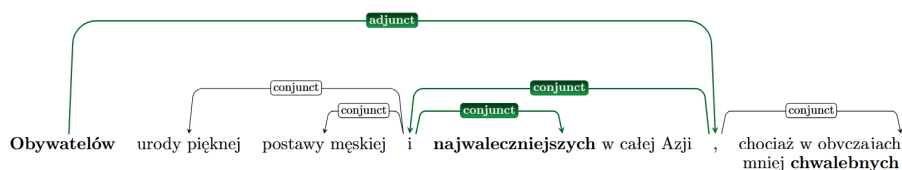


Rysunek 5. Fragment drzewa zależności składniowych zawierający pary *dawnym trybem* i *swoim trybem*



W takich przypadkach reguła „znajdź wyraz o wskazanym znaczniku kategorii gramatycznej, połączony relacją *adjunct* z podrzędnikiem o wskazanym znaczniku klasy gramatycznej i takich samych wartościach określonych kategorii gramatycznych” nie mogła zadziałać, gdyż zamiast jednego z segmentów pary mamy spójnik lub znak interpunkcyjny. Zatem w celu znalezienia przydawki, która jest elementem grupy skoordynowanej (np. *pieszych* w *pieszych i konnych*), należało przyjąć następujące warunki: „jeśli nadrzędnik jest połączony relacją *adjunct* z wyrazem należącym do klasy gramatycznej *spójnik współrzędny* lub *znak interpunkcyjny*, to szukaj wyrazów spełniających kryteria przydawki przymiotnej tego nadrzędnika wśród podrzędników tego spójnika bądź znaku interpunkcyjnego”. Niekiedy należało w ten sposób przeszukać więcej poziomów, jak w przykładzie pokazanym na rys. 6. Aby odnaleźć parę *obywatelów najwaleczniejszych*, trzeba było wziąć pod uwagę, że w sekwencji zależności składniowych pomiędzy segmentami *obywatelów* i *najwaleczniejszych* występuje nie tylko przecinek, lecz także spójnik *i*.

Rysunek 6. Uproszczony fragment drzewa zależności składniowych zawierający pary *obywatelów najwaleczniejszych* i *obywatelów chwalebnych*



Przyjęłam inne zasady analizy w przypadku koordynacji przydawek, a inne w przypadku koordynacji nadrzędników. Jeśli mamy skoordynowane przydawki, to każda z nich tworzy parę z nadrzędnikiem, a spójnik jest uznawany za jeden z tych segmentów, które należy pominąć przy obliczaniu odległości nadrzędnika od przydawki (dlatego w poprzednim paragrafie został wymieniony w punkcie 1a). Zatem dla przykładu na rys. 4 mamy dwie pary: *ochotników pieszych* i *ochotników konnych*, i dla każdej z nich odległość wynosi 1. Natomiast w przypadku skoordynowanych nadrzędników postanowiłam brać pod uwagę tylko jeden z nich, mianowicie ten, który znajduje

się bliżej przydawki w ciągu linearnym. Decyzja była podyktowana tym, że w przeciwnym razie niektóre przydawki byłyby liczone więcej niż raz, a tego chciałam uniknąć. Zatem na rys. 5 mamy dwie pary: *dawnym trybem* i *swoim trybem*, a dla każdej z nich odległość wynosi 1.

Baza danych

Dane zostały zgromadzone w tabeli, która liczy obecnie 43 kolumny. Zawiera ona następujące informacje: numer zdania; długość zdania; forma nadrzędnika; forma przydawki; klasa gramatyczna (pos) nadrzędnika; klasa gramatyczna przydawki; informacja, czy nadrzędnik lub przydawka mają postać grupy skoordynowanej; pozycja nadrzędnika w zdaniu (numer porządkowy segmentu); pozycja przydawki w zdaniu; odległość (mierzona jako różnica pozycji podrzędnika i nadrzędnika); szczegółowe informacje na temat tego, jakie segmenty znajdują się pomiędzy nadrzędnikiem a przydawką oraz czy ich podrzędniki lub nadrzędniki wychodzą poza obręb badanej pary; odległość skorygowana; wybrane metadane tekstu, m.in. półwiecze powstania, rodzaj i gatunek⁹. Poniżej prezentuję wybrane kolumny tej tabeli:

⁹ Dokładny opis metadanych znajduje się na stronie: <https://korba.edu.pl/teksty>.

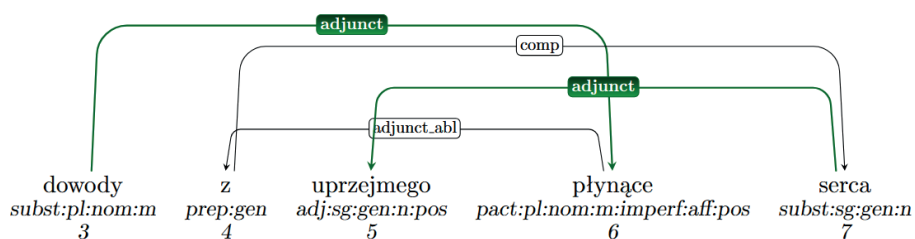
Tabela 1. Przykłady par nadrzędnika z przydawką z SKZ

nadrzędnik	przydawka	pos nadrzędnika	pos przydawki	pozycja nadrzędnika	pozycja przydawki	odległość	odległość skorygowana	półwiecze powstania	rodzaj
zwyczajstwa	tęgo	subst	adj	20	19	-1	-1	1601–1650	teksty perswazyjne
zwyczajstwa	otrzymanego	subst	ppas	20	23	3	1	1601–1650	teksty perswazyjne
okazji	żadnej	subst	adj	24	20	-4	-4	1601–1650	listy
krzyżu	potężnym	subst	adj	19	17	-2	-1	1701–1750	teksty naukowo-dydaktyczne i informacyjno-poradnikowa
krzyżu	srebrnym	subst	adj	19	18	-1	-1	1701–1750	teksty naukowo-dydaktyczne i informacyjno-poradnikowa
jazdy	szlacheckiej	subst	adj	3	5	2	1	1601–1650	teksty perswazyjne
ziele	nieznajome	subst	adj	7	9	2	1	1601–1650	teksty naukowo-dydaktyczne i informacyjno-poradnikowa
przeciągu	dalszym	subst	adj	28	26	-2	-1	1751–1800	teksty naukowo-dydaktyczne i informacyjno-poradnikowa
ochotników	pieszych	subst	adj	6	7	1	1	1751–1800	wiadomości prasowe i druki ulotne
ochotników	konnych	subst	adj	6	9	3	1	1751–1800	wiadomości prasowe i druki ulotne
serca	uprzejmego	subst	adj	7	5	-2	-2	1751–1800	dramat
dowody	płynące	subst	pact	3	6	3	1	1751–1800	dramat

Źródło: opracowanie własne.

Na uwagę zasługuje ostatni wiersz. We fragmencie *dowody z uprzejmego płynące serca* mamy do czynienia z przeplotem (według terminologii M. Świdzińskiego, Świdziński 1996: 116). Para *uprzejmego serca* jest rozbita przez segment *płynące*, który jest podrzędnikiem stojącego poza nią segmentu *dowody*, a para *dowody płynące* jest rozbita przez segment *uprzejmego*, który jest przydawką segmentu *serca*, stojącego poza tą parą (zob. rys. 7).

Rysunek 7. Pary *dowody płynące* i *uprzejmego serca*



A jednak algorytm do klasyfikacji materiału uznał, że segment *uprzejmego* jest podrzędnikiem jednego z elementów pary *dowody płynące* i tym samym nie powoduje nieciągłości. Stało się tak dlatego, że segment *uprzejmego* jest pośrednio podrzędnikiem segmentu *z*, a jeszcze bardziej pośrednio – podrzędnikiem segmentu *płynące*. Nie uwzględniono (jeszcze) takich sytuacji, gdy krawędzie drzewa wychodzą poza badaną parę, a potem wracają w jej obręb. Tego rodzaju błędy nie są częste i nie rzutują na wyniki analiz wstępnych, jednak dobrze byłoby je wyeliminować na dalszym etapie prac.

Te same kryteria zostały zastosowane przy wyborze danych z PDB, przy czym zbiór ten nie zawiera informacji metatekstowych. Poniżej przedstawiam fragment danych z PDB dla następujących fragmentów zdań: *w ubiegłym tygodniu; jesteś Wiesiu ssakiem i to mocno otyłym; kolejny taki przypadek; dom niewiele uszkodzony; z grupą partyzancką o zabarwieniu skrajnym, likwidującą komunistów i uciekinierów; tak młody człowiek jak pan – ciągnęła uczona – nie związany konserwatywnymi przesądami lingwistycznymi, a przeciwnie, ożywiony duchem.*

Tabela 2. Przykłady par nadrzędnika z przydawką z PDB

nadrzędnik	przydawka	pos nad- rzędnika	pos przy- dawki	pozycja nadrzęd- nika	pozycja przydaw- ki	odległość	odległość skorygo- wana
tygodniu	ubiegłym	subst	adj	5	4	-1	-1
ssakiem	otyłym	subst	adj	18	22	4	1
przypadek	kolejny	subst	adj	5	3	-2	-1
przypadek	taki	subst	adj	5	4	-1	-1
dom	uszkodzony	subst	ppas	8	10	2	1
grupą	likwidującą	subst	pact	5	11	6	1
człowiek	związany	subst	ppas	8	16	8	5
człowiek	ożywiany	subst	ppas	8	24	16	5

Źródło: opracowanie własne.

Wstępne wyniki badań

Pary z przydawką w Średniopolskim Korpusie Zależnościowym i Polskim Banku Drzew Zależnościowych – dane ogólne

Z materiału ŚKZ wyodrębniono 4 760 par składających się z nadrzędnika i jego przydawki. W PDB wystąpiło 33 813 takich par. Wśród tych par w ŚKZ było 358 par nieciągłych, co stanowi 7,5% wszystkich par. W PDB natomiast zanotowano jedynie 150 par nieciągłych, czyli 0,4% wszystkich par. Dane te potwierdzają tezę, że nieciągłe grupy z przydawką przymiotną występowały w tekstach z XVII i XVIII wieku znacznie częściej niż dziś.

Tabela 3. Pary z przydawką przymiotną w ŚKZ i PDB

	ŚKZ	Procent	PDB	Procent
Liczba zdań	2 017	–	22 152	–
Liczba segmentów	47 232	–	349 978	–
Średnia długość zdania w segmentach	23,4	–	15,8	–

Tabela 3. Pary z przydawką... (cd.)

	ŚKZ	Procent	PDB	Procent
Liczba wszystkich par nadrzędnik-przydawka	4 760	—	33 813	—
Liczba par nieciągłych	358	7,5%	150	0,4%

Źródło: opracowanie własne.

Rozkład nieciągłych grup nominalnych w tekstach z XVII i XVIII wieku

Większa lub mniejsza tendencja do rozbijania par przydawkowych może zależeć od wielu czynników językowych i pozajęzykowych. Można przypuszczać, że wśród czynników językowych będą to: część mowy nadrzędnika i przydawki; funkcja składniowa, jaką pełni w zdaniu cała badana grupa nominalna; względy stylistyczne. Z faktów pozajęzykowych wpływ może tu mieć rodzaj lub gatunek tekstu; data jego publikacji; tematyka; preferencje autora. Poniżej przyjrzyć się tylko dwóm z tych czynników w ŚKZ: dacie powstania i rodzajowi tekstu.

Rozkład wszystkich par składających się z nadrzędnika i przydawki, jak również rozkład par nieciągłych w czterech półwieczach XVII–XVIII wieku przedstawia tabela 4.

Tabela 4. Występowanie nieciągłych par z przydawką w czterech półwieczach XVII–XVIII wieku

Półwiecze	Liczba wszystkich par nadrzędnik-przydawka	Liczba par nieciągłych	Procent par nieciągłych wśród wszystkich par
1601–1650	1 129	50	4,4%
1650–1700	844	75	8,9%
1701–1750	979	78	8,0%
1751–1800	1 808	155	8,6%

Źródło: opracowanie własne.

Różna liczba par z przydawką w wyróżnionych podokresach w dużej mierze wynika z różnej reprezentacji tych podokresów w korpusie KorBa, co z kolei jest odzwierciedleniem piśmiennictwa epoki (por. <https://korba.edu.pl/overview>). Jeśli chodzi o występowanie nieciągłości w parach nadrzędnik – przydawka, wyróżnia się tu tylko pierwsze półwiecze badanego okresu, gdzie procent par nieciągłych jest zdecydowanie mniejszy od średniej.

W kolejnej tabeli przedstawiam występowanie par nieciągłych z podziałem na rodzaje tekstu. Tu znowu liczba par z przydawką w poszczególnych wierszach jest bardzo zróżnicowana, co ponownie w pewnej mierze odpowiada nierównym proporcjom między tekstami różnych rodzajów w KorBie. Jeśli chodzi o procent par nieciągłych, to zdecydowanie wyróżnia się epika, w których jest tych par znacznie więcej niż w każdym innym rodzaju. Nieco powyżej średniej są także dramat, literatura faktograficzna, wiadomości prasowe i druki ulotne oraz listy. Bardzo mało par nieciągłych jest natomiast w tekstach urzędowo-kancelaryjnych, utworach synkretycznych i Biblii.

Tabela 5. Występowanie nieciągłych par z przydawką w poszczególnych rodzajach tekstów wyszczególnionych w KorBie

Rodzaj tekstu	Liczba wszystkich par nadrzędnik-przydawka	Liczba par nieciągłych	Procent par nieciągłych wśród wszystkich par
Epika	229	41	17,9%
Dramat	141	13	9,2%
Literatura faktograficzna	632	57	9,0%
Wiadomości prasowe i druki ulotne	502	43	8,6%
Listy	108	9	8,3%
Teksty naukowo-dydaktyczne lub informacyjno-poradnikowe	1 402	99	7,1%
Teksty perswazyjne	1 105	72	6,5%
Teksty urzędowo-kancelaryjne	471	21	4,5%
Utwory synkretyczne	22	1	4,5%
Biblia	148	2	1,4%

Źródło: opracowanie własne.

Dane z dwóch powyższych tabel uzupełnią się, jeśli sprawdzimy liczbę tekstów z rodzaju „epika” w poszczególnych półwieczach w korpusie, co przedstawia tabela 6.

Tabela 6. Występowanie tekstów z rodzaju „epika”
w czterech półwieczach XVII–XVIII wieku w KorBie

Półwiecze	Liczba tekstów z rodzaju „epika”
1601–1650	1
1651–1700	5
1701–1750	14
1751–1800	42

Źródło: opracowanie własne.

Liczba tekstów epickich na początku i pod koniec badanej epoki jest nieporównywalna. Przy badaniach takiego materiału będzie zatem należało przyjmować odpowiednie przeliczniki. Na razie można sądzić, że mała liczba par nieciągłych w pierwszym półwieczu badanej epoki wynika z tego, że w KorBie jest tylko jeden tekst epicki z tego podokresu.

Wnioski i plany badawcze

Przedstawione powyżej wstępne analizy pozwoliły na wytyczenie dalszych kierunków badań. Pewnego dopracowania wymaga jeszcze klasyfikacja danych w tabeli, tak aby możliwie wszystkie nieciągłości zostały prawidłowo oznakowane. Już teraz w tabeli zostały zebrane różnorodne informacje na temat cech gramatycznych tak elementów badanej pary, jak i segmentów stojących pomiędzy nimi. Są to informacje o klasie gramatycznej tych segmentów, o ich dalszych powiązaniach składniowych, a także o wzajemnym oddaleniu nadrzędnika i przydawki oraz o ich pozycji w zdaniu. Informacje te zostaną wykorzystane do analizy czynników gramatycznych powiązanych z nieciągłością związku składniowego przydawki i jej nadrzędnika. Przy badaniu wpływu czynników pozatekstowych na nieciągłość analizowanych par należy zwrócić baczną uwagę na wzajemne korelacje tych czynników oraz na nierównomierne występowanie różnego typu tekstów w bazie źródłowej, jaką jest korpus KorBa. Dalszym etapem badań będzie analiza

kolejnych zjawisk językowych: antepozycji bądź postpozycji przydawki, stopnia oddalenia przydawki od jej nadrzędnika, a wreszcie próba ujęcia całych grup nominalnych.

Finansowanie i podziękowania

Praca powstała w wyniku realizacji działania naukowego nr 2023/07/X/HS2/00111 finansowanego ze środków Narodowego Centrum Nauki w ramach konkursu MINIATURA.

Dziękuję anonimowym recenzentom za wskazówki, które skłoniły mnie do doprecyzowania niektórych kwestii w tym artykule.

Bibliografia

- BIEŃ J.S., SALONI Z., 1982, Pojęcie wyrazu morfologicznego i jego zastosowanie do opisu fleksji polskiej (wersja wstępna), *Prace Filologiczne* XXXI, s. 31–45, [online:] <http://bc.klf.uw.edu.pl/63/>.
- DERWOJEDOWA M., 2000, *Porządek linearny składników zdania elementarnego w języku polskim*, Warszawa: Dom Wydawniczy Elipsa.
- FITIAŁOW S.J. – ФИТИАЛОВ, 1962, О моделировании синтаксиса в структурной лингвистике, w: *Проблемы структурной лингвистики*, Москва: Издательство Академии Наук СССР, s. 100–114.
- GĘBKA-WOLAK M., 1997, O szyku składników grupy nominalnej, *Acta Universitatis Nicolai Copernici. Filologia Polska* 48, s. 3–34, [online:] <http://repozytorium.umk.pl/handle/item/4934>.
- GĘBKA-WOLAK M., 2000, *Związki linearne między składnikami grupy nominalnej we współczesnym języku polskim*. Toruń: Wydawnictwo UMK.
- GRUSZCZYŃSKI W., ADAMIEC D., BRONIKOWSKA R., KIERAŚ W., MODRZEJEWSKI E., WIECZOREK A., WOLIŃSKI M., 2022, The Electronic Corpus of 17th- and 18th-century Polish Texts, *Language Resources and Evaluation* 56, s. 309–332, [online:] <https://doi.org/10.1007/s10579-021-09549-1>.
- GRUSZCZYŃSKI W., ADAMIEC D., BRONIKOWSKA R., WIECZOREK A., 2020, Elektroniczny Korpus Tekstów Polskich z XVII i XVIII w. – problemy teoretyczne i warsztatowe, *Poradnik Językowy* 8, s. 32–51. <https://doi.org/10.33896/PorJ.2020.8.3>.
- KARPLUKÓWNA M., 1967, *Język Jana Jurkowskiego (około 1580–1639)*. Wrocław–Warszawa–Kraków: Zakład Narodowy im. Ossolińskich.

- KLEMENSIEWICZ Z., 2009, *Historia języka polskiego*, Warszawa: Wydawnictwo Naukowe PWN.
- PERCZYŃSKA N., 1978, Uwagi o kolokwializmach składniowych w języku pamiętników Jana Chryzostoma Paska, w: J. Twardzikowa (red.), *Studia z polskiej składni historycznej II*, Wrocław–Warszawa–Kraków–Gdańsk: Zakład Narodowy im. Ossolińskich, s. 59–64.
- PRZEPIÓRKOWSKI A., BAŃKO M., GÓRSKI R.L., LEWANDOWSKA-TOMASZCZYK B. (red.), 2012, *Narodowy korpus języka polskiego*, Warszawa: Wydawnictwo Naukowe PWN, [online:] https://nkjp.pl/settings/papers/NKJP_książka.pdf.
- SKUBALANKA T., 1984, *Historyczna stylistyka języka polskiego. Przekroje*, Wrocław–Warszawa–Kraków–Gdańsk–Łódź: Zakład Narodowy im. Ossolińskich.
- SMOLIŃSKA B., 1983, *Polszczyzna północnokresowa z przełomu XVII i XVIII w. Na podstawie rękopisów Jana Władysława Poczobuta Odlanickiego i Antoniego Kazimierza Sapiehy*, Wrocław–Warszawa–Kraków–Gdańsk–Łódź: Zakład Narodowy im. Ossolińskich.
- SZPAKOWICZ S., ŚWIDZIŃSKI M., 1990, Formalna definicja równorzędnej grupy nominalnej we współczesnej polszczyźnie pisanej, *Studia Gramatyczne IX*, Wrocław, s. 9–54.
- ŚLIWIŃSKI W., 1982, W poszukiwaniu metody opisu szyku wyrazów w grupach nominalnych dzisiejszej polszczyzny pisanej, *Pamiętnik Literacki LXXIII*, z. 1–2, s. 145–166.
- ŚLIWIŃSKI W., 2006, Inwersje barokowe w Kanikule i Lutni Jana Andrzeja Morsztyna, *LingVaria* 1(1), s. 103–116.
- ŚWIDZIŃSKI M., 1996, *Własności składniowe wypowiedników polskich*, Warszawa: Dom Wydawniczy Elipsa.
- WIECZOREK A., 2005, Towards Middle Polish Dependency Treebank, w: A. Wilczek, M. Marczevska (red.), *Native Language in the 21st Century: System, Communication Practices and Education*, V & R Unipress GmbH, s. 37–49.
- WIERZBICKA A., 1963, Lingwistyczne narzędzia w stylistycznej analizie szyku wyrazów, *Pamiętnik Literacki* z. 2, s. 511–533.
- WIERZBICKA A., 1964, O metodach opisu szyku wyrazów, *Język Polski* z. 1, s. 14–26.
- WIERZBICKA A., 1966, *System składniowo-stylistyczny prozy polskiego renesansu*, Warszawa: Państwowy Instytut Wydawniczy.
- WIŚNIEWSKA H., 1994, *Kulturalna polszczyzna XVII wieku. Na przykładzie Zamościa*, Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej.
- WRÓBLEWSKA A., RYBAK P., 2019, Dependency Parsing of Polish, *Poznań Studies in Contemporary Linguistics* 55(2), s. 305–337, [online:] <https://doi.org/10.1515/psicl-2019-0012>.

**Methods of research on the discontinuous order of adjectival attributes
in the Middle Polish Dependency Corpus**

(s u m m a r y)

This article explores a methodology for researching discontinuous word order of adjectival modifier in seventeenth- and eighteenth-century sentences, using a dependency treebank. This extensive material, marked morphosyntactically and syntactically, offers new analytical possibilities; however, methods for excerpting and classifying the studied structures are necessary. Contemporary material from the Polish Dependency Bank is used for comparison.

K e y w o r d s: discontinuous syntactic relationship; adjectival modifier; dependency treebank; 17th- and 18th-century Polish