

# Numeryczne kody gramatyczne we wzbogaconym korpusie *Słownika frekwencyjnego polszczyzny współczesnej*

Redaktorzy:  
Janusz S. Bien<sup>\*</sup>, Marcin Woliński<sup>\*\*</sup>

17 grudnia 2001 roku

Niniejszy tekst stanowi jeden ze składników dokumentacji wzbogaconego korpusu *Słownika frekwencyjnego polszczyzny współczesnej* — w skrócie WKSF.

Stosowane w korpusie kody liczbowe są konsekwencją założeń, przedstawionych w każdej z niżej wymienionych publikacji. Szczegółowe ustalenia zostały dokonane w instrukcjach redakcyjnych, które również zostały opublikowane.

Prawie cały tekst wstępu oraz wspomniane instrukcje zostały wprowadzone do komputera przez Martę Nazarczuk [8], która jednocześnie wprowadziła pewne zmiany typograficzne. Pominięte fragmenty wstępu to punkty *Cel i charakter niniejszej publikacji*, które — jak wskazuje tytuł — zawierały informacje specyficzne dla konkretnego tomu. Każdy z tomów zawierał również specyficzną dla niego listę źródeł, stąd różnice w objętości wstępu.

Warto wspomnieć, że publikacja [7] zawiera także omówienie podstawowych założeń, również w języku angielskim.

Poniżej załączamy informacje powtórzone za cytowanymi niżej pracami w wersji przygotowanej przez Nazarczuk. Tekst ten znajduje się

- w tomie [1] na s. 5-60,
- w tomie [2] na s. 5-36,
- w tomie [6] na s. 3-35,
- w tomie [3] na s. 5-57,
- w tomie [4] na s. 7-53.

Na s. 32 pisała ona mianowicie

W każdym tomie zdającym sprawę z prac nad listami frekwencyjnymi [*Kurcz i in.*], we wstępie zamieszczono instrukcje wpisywania tekstów na fiszki oraz „Opis gramatyczny”. Jest to cenne źródło informacji o kodzie, użytym w korpusie do odróżnienia form homonimicznych. Tekst drukowany w latach siedemdziesiątych wydaje się jednak mało czytelny<sup>1</sup>. Poniżej przytaczam fragmenty wstępu trzeciego tomu, z wprowadzonymi nawiasami w miejsce pojedynczych ukośników i pojedynczych ukośników w miejsce ukośników podwójnych. Podkreślenia wyrazów linią przerywaną, odróżniające w tekście przykłady analizowanych form, zastąpiłam przez pochylony krój czcionki. Używam polskiego cudzysłowu. Cyfry sta-

<sup>\*</sup> Zakład Zastosowań Informatycznych, Instytut Orientalistyczny Uniwersytetu Warszawskiego, patrz <http://www.orient.uw.edu.pl/~zzi/>

<sup>\*\*</sup> Zespół Inżynierii Lingwistycznej, Instytut Podstaw Informatyki PAN

<sup>1</sup> Autorzy książki nie mieli możliwości pisania pewnych znaków, np. nie mogli używać nawiasów — w ich miejsce wpisywali ukośnik; posługiwali się także podwójnym ukośnikiem.

nowiące oznaczenia gramatyczne, jeżeli zawierają tzw. cyfrę na pierwszej pozycji (oznaczenie części mowy), podaję w nawiasach kwadratowych (tak, jak w komputerowej wersji próbek).

Poniżej podajemy wspomniane instrukcje w takiej formie, w jakiej znajdują się one w pracy Nazarczuk.

# Opis kodu korpusu

## Wstęp

### 0.1. Jednostka badania

Ze względu na leksykalno-morfologiczny typ podjętych badań frekwencyjnych podstawową jednostką analizy w tekście jest słowoforma, czyli forma fleksyjna wyrazu. Wyraz definiujemy — z pewnymi wyjątkami — według definicji graficznej, określając go jako ciąg liter zawartych między dwoma kolejnymi pauzami. Takie ujęcie było konieczne przy maszynowym opracowaniu danych. Z tego też względu formy czasów złożonych typu *będę jechał, pojechali byli* traktuje się jako składające się z dwóch odrębnych słowoform; funkcje tych słowoform jako składowych form czasów złożonych wskazuje jedynie specjalny kod liczbowy (por. Kod fleksyjny str. 40). Podobnie — zgodnie z zasadą analityczności — wyróżnia się dwie odrębne słowoformy w zestawieniach nominalnych typu *Stany Zjednoczone*; jako ciągi osobnych słowoform traktujemy także zestawienia liczebników porządkowych i głównych typu *sześćdziesiąty piąty, tysiąc dwieście czternaście* itp.

Od reguły tej stosowano jednakże wyjątki, ujmując niekiedy jako jedną słowoformę segment tekstu składający się z dwóch formalnie różnych członów. Dotyczyło to następujących wypadków:

1. Dwa formalnie odrębne człony stanowią jednostkę zleksykalizowaną, przy czym przynajmniej jeden z członów utracił samodzielność, stanowiąc archaizm leksykalny bądź fleksyjny — są to wyrażenia typu *po prostu, za młodu, po omacku, na bakier* itp.
2. Dwa formalnie odrębne człony nie dopuszczają wstawienia między nie innego członu, np. *w ogóle, przede wszystkim, na razie* itp. Tę zasadę stosowano tylko do niektórych leksemów wchodzących w skład utartych połączeń.
3. Częstki *de, von* stojące przed obcymi nazwiskami stanowią wraz z nazwiskiem jedną słowoformę (*de Gaulle, von Hassel*).

Jako jedną słowoformę traktuje się także:

1. imiesłowy przymiotnikowe i rzeczowniki odczasownikowe z zaimkiem *się* (*skradający się, skradanie się*);
2. cząstki *co* i *jak* stojące przed przymiotnikami i przysłówkami w stopniu najwyższym wraz z danymi częściami mowy (*co naj-, jak naj-*);
3. cząstkę *do* + przysówek (typ *do dziś*);
4. cząstkę *na* + przysówek (typ *na gorąco, na bieżąco, na nowo, na zawsze*);
5. cząstkę *od* + przysówek (typ *od wczoraj, od dziś*);
6. cząstkę *z* + przysówek (typ *z zewnątrz*);
7. cząstkę *za* + przysówek/przymiotnik (typ *za dużo, za mało, za duży, za mały*).

Jeszcze inaczej potraktowano czasowniki zwrotne, zaliczając do nich z punktu widzenia formalnego wszystkie czasowniki z zaimkiem *się*. Ze względu na definicję graficzną wyrazu oraz biorąc pod uwagę fakt, że informacja o częstotliwości zaimka *się* jako odrębnej słowoformy jest ważna dla językoznawców, interpretowano czasownik zwrotny jako składający się z dwóch słowoform (typ *biją się*). Aby jednakże nie utracić informacji językowej o zwrotności danego czasownika, umieszczano przy słowoformie czasownika symbol 501 jako znak kodu fleksyjnego wskazujący na zwrotność (*bili 501 się*). Stąd też wszystkie słowoformy czasowników zwrotnych (z wyjątkiem imiesłów przymiotnikowych traktowanych

jako osobne hasła) opatrzone są symbolem 501. Program hasłowania pozwalał w toku opracowania na przekształcenie symbolu 501 na zaimbek *się* przy hasle, stąd „pełna” postać hasel czasowników zwrotnych (por. hasło *bić się*).

W wielu wypadkach wątpliwych dotyczących postaci hasła przyjmowano rozstrzygnięcia zawarte w „Słowniku Języka Polskiego” pod red. W. Doroszewskiego.

## 1. Instrukcja I

### 1.1. Rozpisywanie tekstów

1. Skróty rozwiązywano, traktując je jako normalne słowoformy: *prof.* = *profesor*, *dla dra Kowalskiego* = *dla doktora Kowalskiego*, podobnie *mgr*, *doc.*, *tow.*, *ob.* i wszelkie inne — takie jak: *np.* = *na przykład*, *itd.* = *i tak dalej*, *br.* = *bieżącego roku*, *itp.* = *i tym podobne*.
2. Jednostki miar pisane skrótami rozwiązywano: *kg* = *kilogram*, *cm* = *centymetr*, *godz.* = *godzina*.
3. Liczby pisane cyframi arabskimi i rzymskimi przepisywano słownie:

*po XX zjeździe* = *po dwudziestym zjeździe*, *o godz. 9* = *o godzinie dziewiątej*, *w 1966 r.* = *w tysiąc dziewięćset sześćdziesiątym szóstym roku*.

4. Złożenia z członem cyfrowym przepisywano słownie: *50-lecie* = *pięćdziesięciolecie*, *3-dniowy* = *trzydniowy*.
5. Skrótownice (literowce, głoskowce, sylabowce) nie były rozwiązywane i traktowano je jako osobne słowoformy: *bombardowanie DRW* = *bombardowanie DRW*, *rocznica PKWN* = *rocznica PKWN*, *prezydium PAN* = *prezydium PAN*, *w PAN-ie* = *w PAN-ie*, *w PGR-ach* = *w PGR-ach*, *Pafawagiem* = *Pafawagiem*.
6. Pominięto w tekście wszelkie wzory oraz symbole typu  $H_2O$ , odcinek *AB* itp.

## 2. Instrukcja II

### 2.1. Imiona własne

Imię własne znaczymy znakiem +.

1. Za imię własne uważa się:
  - a) imiona i nazwiska ludzi i zwierząt (*Jan + Kowalski +*, *Azor +*); także, jeśli są homonimiczne z apelatywami (np. *Puzon +*, *Gołąb +*);
  - b) nazwy umowne instytucji, pojazdów lądowych i wodnych (*kawiarnia Sarenka +*, *statek Orzeł +*, *ekspres Błękitna + Fala +*);
  - c) nazwy geograficzne, np. *Wielka + Brytania +*, *Związek + Radziecki +*, *Stany + Zjednoczone +*, *Morze + Czarne +*, *Bagno +* (dzielnica Warszawy), *Bliski + Wschód +*;
  - d) tytuły książek, czasopism, filmów, referatów (*Ogniem + i + mieczem +*, *Życie + Gospodarcze +*, *Tam + gdzie + rosną + poziomki +*, *O + aktualnej + sytuacji + politycznej + w + świecie +*).
2. Nie są imionami własnymi:
  - a) pełne, nieumowne nazwy instytucji typu *Spółdzielnia pracy szewców i cholewkarzy*, *Centrala Handlowa Przemysłu Mleczarskiego*;
  - b) przymiotniki utworzone od imion własnych, np. *bliskowschodni*, *amerykański*, *warszawski*;

- c) nazwy mieszkańców danego regionu typu *poznaniak*, *Amerykanin*, *Chińczyk*;
  - d) nazwy firmowe samochodów, papierosów, słodyczy, np. *moskwicz*, *wołga*, *warszawa*, *sporty*, *giewonty*, *katarzynki* itp.
3. Znakiem + znaczone także umowne symbole klas sportowych (*drużyna zakwalifikowała się do klasy a +*) oraz witamin (*witamina b +*).

### 3. Opis gramatyczny

#### 3.1. Kod fleksyjny

Leksykalno-morfologiczne dane frekwencyjne z naszego opracowania dotyczą częstości zarówno haseł, jak i form fleksyjnych. Przy ustalaniu haseł na podstawie słowoform występujących w izolacji napotykamy na różne typy form homonimicznych. Jak wiadomo, istnieją 3 typy homonimii:

1. homonimia leksykalna typu *pokój*, *zamek*;
2. homonimia morfologiczna dotycząca form fleksyjnych; jest nią w obrębie nomen homonimia przypadków i liczb (*sprawności* jako gen., dat., loc., sg. i gen. pl., *dobrym* jako instr., loc. sg. oraz dat. pl.);
3. homonimia syntaktyczna dotycząca poszczególnych części mowy, polegająca na tym, że dany wyraz może pełnić funkcje syntaktyczne rzeczownika bądź czasownika (*brak*, *potrzeba*), rzeczownika bądź przysłówka (*czasem*), rzeczownika bądź przymiotnika (*chory*, *pośpieszny*), czasownika bądź przysłówka (*wolno*), przysłówka bądź przyimka (*blisko*, *obok*, *wokół*) itd.

W opracowaniu naszym pominięto homonimię leksykalną, przede wszystkim ze względu na trudności teoretyczne w odróżnianiu homonimii leksykalnej od polisemii. Homonimię morfologiczną i syntaktyczną usuwano, różnicując formy homonimiczne przez dopisywanie umownych symboli cyfrowych. Ze względu na ograniczoną pojemność pamięci maszyny liczba symboli gramatycznych została ograniczona do 63; ułożony kod jest kodem pozycyjnym.

Zwracamy uwagę, że w opracowywanym materiale leksykalnym symbolizacja odpowiednich cech gramatycznych dotyczy tylko słowoform i haseł homonimicznych, nie obejmuje więc ona całości badanego słownictwa. Słowoformy i hasła nie kodowane oznaczają więc formy nie będące homonimiami, których funkcje morfologiczno-syntaktyczne łatwo odczytać z samej postaci wyrazu (por. *domami*).

Tylko w dwóch wypadkach — przy rzeczownikach i przy liczebnikach brak kodu ma jeszcze inne, odrębne znaczenie. Przy rzeczownikach oznacza on także nieodmienność członu, pełniącego funkcję przydawki rzeczownikowej (przykłady typu *w hucie Pokój +*, *do spółdzielni Jedność +*), a także występującego przy odmianie tytułów i nazwisk żeńskich (*z doktor Kowalską +*, *nagrodę przyznano Janinie + Bator +*). Przykłady dotyczące liczebników wymagają szerszego omówienia i podane są niżej (por. str. 50).

Homonimię syntaktyczną usuwamy, umieszczając przy homonimicznym hasle symbol części mowy. Są to cyfry od 1 do 9 występujące *w pierwszej pozycji* po słowoformie. Interpretuje się je następująco:

1. „rzeczownik” — symbol przypisywany rzeczownikom i substantywizowanym przymiotnikom;
2. „przymiotnik” — symbol przypisywany przymiotnikom, liczebnikom i zaimkom przymiotnikowym oraz imiesłowom przymiotnikowym;

3. „liczebnik” — symbol przypisywany liczebnikom głównym i zbiorowym oraz zaimkom liczebnym (*ile, tyle, wiele*);
4. „zaimek” — symbol oznaczający zaimki rzeczownikowe;
5. „czasownik” — symbol oznaczający formy czasownika i *prædicativa* (*potrzeba, brak*);
6. „przyimek” — symbol oznaczający przyimki pierwotne i wtórne (*blisko, obok*);
7. „wykrzyknik” — symbol odróżniający wykrzykniki od spójników (*a, o*);
8. „przysłówek” lub „partykuła” — symbol podawany przy słowoformach nieodmiennych nie będących przyimkami ani spójnikami;
9. „spójnik”.

Pozycja II występuje jedynie wówczas, gdy w pozycji I użyto symboli 1–6. W obrębie rzeczowników [1] i przymiotników [2] odróżniamy homonimie przyimków (symbole 1–6 na II pozycji) i liczb (symbole 1–2 na III pozycji). Nie uwzględniono natomiast homonimii w obrębie kategorii rodzaju, gdyż pomnożyłoby to znacznie liczbę symboli.

Przy czasowniku [5] II pozycja, na której występują cyfry od 1 do 4, pozwala odróżnić użycie bezokolicznika w funkcji składnika form czasu przyszłego od innych jego użyc, oraz trzy różne funkcje formy na -ł (składnik form czasu przyszłego, czas przeszły z ruchomą końcówką i tryb warunkowy z ruchomą partykułą). Symbol 5 na II pozycji wyróżnia użycie form czasu teraźniejszego w funkcji trybu rozkazującego. Z kolei symbole 6–7 na II pozycji zarezerwowano dla wyróżnienia niektórych funkcji gramatycznych czasowników *być, bywać, zostać* (człony składowe w czasach złożonych i w formach strony biernej).

Na III pozycji występuje cyfra 1 przy czasownikach zwrotnych lub 0 (czyli brak symbolu cyfrowego) przy czasownikach niezwrotnych.

### 3.2. Homonimia syntaktyczna

Ze względu na brak wypracowanych ścisłych kryteriów pozwalających na odróżnienie poszczególnych części mowy, przyjęto przy kodowaniu pewne sprawdziany o charakterze roboczym. Starano się ustalać kryteria wyłącznie syntaktyczne, abstrahowano całkowicie od kryteriów semantycznych. Szczegółowy opis przyjętych przez nas zasad rozróżniania części mowy został opublikowany osobno (Biuletyn PTJ nr XXIX, 1971), tu ograniczymy się do skrótego omówienia tychże zasad.

Przyjęte kryteria będziemy podawać, omawiając kolejno parami części mowy, których funkcje mogą pełnić opisywane homonimy syntaktyczne. Dla jasności oznaczamy określony rodzaj homonimii syntaktycznej symbolami (SV = subst./verbum = 1/5, SAdv, VA itp.).

#### 3.2.1. Rzeczownik czy czasownik (typ SV = 1/5)

Homonimia ta dotyczyła wyrazów *potrzeba, brak, szkoda, strach, żal, wstyd*.

Kryterium: kwalifikator czasownika [5], jeżeli w danym zdaniu występuje względnie można wstawić człon (*było*) *nam* = zaimek osobowy w celowniku przy nieobecności innego członu predykatywnego.

S: *brak*[1] *ten odczuwano bardzo boleśnie*;

V: *w dyskusji brak*[5] *nam było ładu*.

#### 3.2.2. Przymiotnik czy czasownik (typ AV = 2/5)

Homonimiami syntaktycznymi tego typu są hasła *winien* albo *winny*. Hasło *winien/winny* kwalifikowano:

1. jako czasownik [5], jeżeli w zdaniu towarzyszy mu bezokolicznik: *weryfikacja winna [5] objąć wszystkich członków*;
2. jako przymiotnik [2], jeżeli pełni funkcję przydawki lub orzecznika: *człowiek winny [2] zbrodni; winni [2] są rodzice, nie dzieci; także jest mi winien [2] sto złotych*.

### 3.2.3. Przysłówek czy czasownik (typ AdvV=8/5)

Homonimia ta dotyczy wyrazów *wolno* oraz *przeszło*.

*Wolno* kwalifikujemy:

1. jako przysłówek [8], jeżeli jest możliwa substytucja wyrazów *pomału, powoli: jechał wolno* [8];
2. jako czasownik [5], jeżeli w danym zdaniu występuje względnie można wstawić człon (*było nam* = zaimek osobowy w celowniku: *nie wolno (nam palić)*). *Przeszło* kwalifikujemy jako przysłówek [8], jeżeli jest możliwa substytucja wyrazów *więcej niż*, w pozostałych wypadkach kwalifikujemy jako czasownik [5]: *wydobyto przeszło [8] tysiąc ton węgla*.

### 3.2.4. Rzeczownik czy przymiotnik (typ SA = 1/2)

Homonimia ta dotyczy substancywizowanych przymiotników. Ze względu na różne stopnie tej substancywizacji przyjęto sformułowane przez S. Jodłowskiego formalne kryterium, które wskazuje przy przymiotnikach substancywizowanych na ograniczenie form rodzaju do dwóch lub nawet do jednego rodzaju. Np. *chory* ma tylko 2 rodzaje w funkcji rzeczownika [1]. Używając tego kryterium kwalifikowano jako rzeczowniki [1] takie wyrazy jak *przewodniczący, biały, młody* itp. A także słownictwo typu *wytyczna, średnia, dane*, również ułamki (*dziesiąta, setna*). Przykłady kontekstów: *otrzymano nowe wytyczne*[1], *średnia*[1] *krajowa w zakresie spożycia, dane*[1] *liczbowe* itp.

### 3.2.5. Rzeczownik czy przyimek (typ SPrp = 1/6)

Homonimia ta dotyczy niektórych rzeczowników, które w pewnych wypadkach mogą pełnić funkcje przyimków. W badanych tekstach były to rzeczowniki *celem, drogą, koło, skutkiem, względem*.

Jako kryteria pozwalające uznać dany człon za przyimek [6] przyjęto:

1. możliwość zastosowania określonych transformacji lub substytucji,
2. relację danego członu z innymi członami ciągu syntaktycznego (kontekst).

*Celem* + gen. nom. actionis jest przyimkiem [6], jeżeli jest możliwa transformacja tego wyrażenia w podrzędne zdanie okolicznikowe celu, tzn. gdy zachodzi relacja:

*celem* + gen. nom. act. → *aby* + infinitivus/verbum finitum.

Przykład: *Rada Bezpieczeństwa zebrała się celem [6] przedyskutowania sytuacji politycznej* → ...*aby przedyskutować sytuację polityczną*. Jeżeli podobna transformacja nie jest możliwa, *celem* uzyskuje kwalifikator rzeczownika [1].

*Drogą* + gen. nom. actionis jest przyimkiem [6], jeżeli jest możliwa substytucja wyrażenia przyimkowego *przez* + acc. nom. act.: *drogą [6] pertraktacji uzyskano ustępstwa* → *przez pertraktacje*... Natomiast *drogą* jest rzeczownikiem [1], jeżeli:

1. jest członem określanym,
2. nie jest możliwa powyższa substytucja.

*Skutkiem* + gen. oraz *względem* + gen. kwalifikujemy jako przyimki [6], jeżeli są możliwe odpowiednie substytucje: dla *skutkiem* substytucja *wskutek*, dla *względem* substytucja *wobec*:

*skutkiem [6] nieostrożności spowodowano pożar* → *wskutek nieostrożności... , mieć względem [6] kogoś zobowiązania* → *mieć wobec kogoś...* Jeżeli zaś substytucje te nie są możliwe, kwalifikujemy oba te człony jako rzeczowniki [1].

*Koło* + gen. kwalifikujemy jako przyimek [6], jeżeli jest możliwa substytucja wyrazu *obok*: *koło [6] lasu* → *obok lasu*. W wypadku dwuznaczności (*koło wozu*) o możliwości lub niemożliwości tej substytucji rozstrzyga kontekst, a więc relacja z innymi członami ciągu syntaktycznego: *koło [6] wozu pasy się konie*, ale *pękło nam tylne koło [1] wozu*.

### 3.2.6. Zaimki

Zaimki klasyfikujemy wyłącznie na podstawie kryterium syntaktycznego, wyodrębniając tylko dwie klasy zaimków: zaimki o odmianie przymiotnikowej, mające homonimiczne formy przypadka i liczby (*których, jakie*) oraz zaimki o odmianie rzeczownikowej, mające homonimiczne formy przypadka (*kogo, czym*). Zaimki przymiotnikowe otrzymują symbol [2], zaimki rzeczownikowe — symbol [4]. Tak więc nie uwzględniamy — z bardzo małymi wyjątkami (por. „Hasła kodowane umownie”, str. 54) klasyfikacji semantycznej zaimków (zaimki względne, pytajne, nieokreślone itp.).

Rozróżnienie między zaimkami przymiotnikowymi i rzeczownikami nie jest kłopotliwe, ulega ono jedynie zmianie w odniesieniu do zaimków *to, wszystko, samo, jedno*. Mianowicie:

1. Zaimki *ten, ta, to; wszystek, wszystka, wszystko; sam, sama, samo; jeden, jedna, jedno* są traktowane jako zaimki przymiotne [2], jeżeli występują :
  2. w liczbie mnogiej,
  3. w liczbie pojedynczej w funkcji przydawki. Przykłady: *ci ludzie; zaprosz tych, których znasz; zastanawia mnie to zagadnienie; wszyscy ludzie; wszystkim się zdawało, że Wojski wciąż gra jeszcze; kot wypił wszystko mleko; w zebraniu uczestniczyli sami cudzoziemcy; dokonały tego zupełnie same, bez niczyjej pomocy; błąd tkwił już w samym założeniu; jedne pary szły na prawo, drugie na lewo; jedni chcą tak, inni inaczej; jedna jaskółka wiosny nie czyni; jedna pani powiedziała, że widziała twoją córkę z chłopcem.*
4. Zaimki *to, wszystko, samo, jedno* są traktowane jako zaimki rzeczownikowe [4], jeżeli występują w liczbie pojedynczej w funkcji podmiotu, dopełnienia lub przydawki dopełniaczowej. Przykłady:

*to* [4]: *dążył do tego, aby się zemścić; chodziło nam o to, aby jak najszybciej ukończyć studia; bardzo się temu sprzeciwiano; rzecz w tym, że już nie dysponujemy na to funduszami; także jest to niebagatelny problem;*

*wszystko* [4]: *troska o wszystko; niedobrze we wszystkim dopatrywać się zła; wszystkiego się można po nim spodziewać; wszystkiemu można by zaradzić;*

*jedno* [4]: *jedno jest dla mnie oczywiste — że brak im wytrwałości; przykład ten świadczy o jednym — o ich wzajemnym przywiązaniu;*

*samo* [4]: *to samo mam na myśli; tym samym projekt zatwierdzono; chodziło nam wszystkim o to samo.*

Uwaga: w połączeniach *to samo, to wszystko, to jedno* przyjęto umownie *to* jako człon określający (przydawkę = [2]), *samo, wszystko, jedno* jako człon określany (podmiot, dopełnienie = [4]): *to samo chciałem powiedzieć, mieliśmy szczęście w tym wszystkim; to jedno warto zapamiętać.*

Zaimek *to* jest także homonimiczny ze spójnikiem [9]: *jeżeli... to, chociaż... to, skoro... to, to... to*, jednakże odróżnienie to nie sprawiało przy opisie żadnego kłopotu.

Człon *to* kwalifikowano także jako partykułę [8], jeżeli w danym ciągu skła-



dniowym człon ten można było zerować: *coraz to* [8] *chłodniej* ~ *coraz chłodniej*; *gdyby to* [8] *chodziło o was* ~ *gdyby chodziło o was*; *kiedy to* [8] *wreszcie skończą się kłopoty* ~ *kiedy wreszcie skończą się kłopoty*; także: *brat długi zwrócił i to* [8] *jeszcze przed terminem* ~ *brat długi zwrócił i jeszcze przed terminem*.

Zaimek *tym* (= forma fleksyjna zaimka *ten*) jest homonimiczny ze spójnikiem [9] w połączeniach *i. ... tym: im* [9] *dalej w las, tym* [9] *więcej drzew* oraz ze spójnikiem w połączeniach stopnia wyższego przysłówków i przymiotników: *tym* [9] *bardziej, tym* [9] *większa radość* itp., odróżnienie to nie wymagało jednak formułowania osobnych kryteriów.

Bardziej szczegółowy opis uzyskał także zaimek rzeczownikowy *co*. Oprócz normalnych wyróżnień homonimicznych form przypadków (*co* nom. i acc., *czym* inst. i loc. itp.) kwalifikatorem mianownika [41] opatrywano *co* występujące przy stopniu wyższym przysłówka (*co gorzej, co więcej*) oraz wyróżniano umownie symbolem [6] *co* występujące przy miarach czasu i odległości (typu *co* [64] *godzinę, co* [62] *roku*). Także umownie wyróżniano *co* w funkcji zaimka względnego, opatrując je symbolem [9]: *ten, co* [9] *nie miał szczęścia*, oraz *co* w funkcji partykuły [8]: *dopiero co*.

Na osobną wzmiankę zasługują wreszcie zaimki liczebnikowe *tyle, wiele*. Oba te człony traktowano:

1. jako liczebniki [3] mające homonimiczne formy przypadków, jeżeli były one członami określającymi przy nomen (*w tylu* [36] *sprawach, z tyłoma* [35] *ludźmi, w wielu sprawach, z wieloma ludźmi*;
2. jako przysłówki [8], jeżeli były one nieodmiennymi członami określającymi przy verbum (*wiele* [8] *można zrobić; tyle* [8] *widziałem, że nie wiem, od czego zacząć*), a także — umownie — jeżeli następuje po nich wyrażenie przyimkowe: *to rozwiązanie pozostawia wiele* [8] *do życzenia; miał jeszcze wiele* [8] *do zrobienia*.

W połączeniach *o tyle ... o ile, o tyle ... że, nie tyle ... ile, tyle ... że* człon *tyle* wchodzi w skład większego wyrażenia, stanowiącego spójnik; podobnie jak w wypadku członu *to*, także i ten przykład homonimii syntaktycznej nie stwarzał trudności przy opisie i nie wymagał stosowania żadnych specjalnych kryteriów.

### 3.2.7. Liczebniki

Skomplikowany i nieregularny charakter tej kategorii w języku polskim nastroczał mnóstwo kłopotów w opisie. Jak już wspomniano, wyodrębniono w zasadzie 2 klasy liczebników: liczebniki o odmianie przymiotnikowej [2] mające homonimiczne formy w zakresie przypadku i liczby (por. *pierwszym, jednej* oraz pozostałe mające homonimię tylko w zakresie przypadku (por. *pięciu, dwóch*).

Przy liczebnikach głównych odrębne znaczenie ma także brak kwalifikatora. Dotyczy to:

1. liczebników głównych występujących jako człony nieodmienne w wieloczłonowych zestawieniach mających funkcję liczebników porządkowych: *w sto dwudziestym pułku; od tysiąc dziewięćset piętego roku*;
2. liczebników głównych w funkcji nieodmiennych przydawek ilościowych (przykłady częste w prasie codziennej i w publicystyce): *zwyciężył Kowalski w czasie pięć godzin dziewięć minut piętnaście sekund; ze stopą wzrostu osiem procent, w latach tysiąc dziewięćset sześćdziesiąt – sześćdziesiąt trzy; zwyciężyli Belgowie w stosunku pięć – cztery*.

Dodatkowy opis dotyczy dwóch funkcji liczebnika *jeden*. Jak już wspomniano przy omawianiu zaimków, przy członach *jeden, jedna, jedno* nie rozróżniamy funkcji liczebnika i zaimka nieokreślonego. Natomiast wyraz *jeden* otrzymuje inne kwalifikatory w następujących dwóch wypadkach:

1. umownie kwalifikator [8], jeżeli jest on składnikiem odmiennego liczebnika głównego: *w dwudziestu jeden [8] krajach, delegaci trzydziestu jeden [8] państw*;
2. bez kwalifikatora, jeżeli występuje jako liczebnik główny w funkcji ilościowej przydawki nieodmiennej: *miara o długości jeden metr trzydzieści centymetrów; w czasie dwie godziny dwadzieścia jeden minut*.

### 3.2.8. Przysłówek czy przyimek (typ Adv/Prp = 8/6)

Jest to rozróżnienie bardzo trudne, a dotyczyło takich członów jak *blisko, obok, dokoła, ponad, wokół, poniżej, powyżej*. W gramatykach polskich brak szczegółowego opisu odrębnych (dwojakich) funkcji tych członów, także w opisie tych haseł w „Słowniku języka polskiego” pod red. W. Doroszewskiego nie rozgranicza się tych funkcji (podano je wyłącznie przy hasłach *około* i *obok*).

Kryterium przyjęte przez nas brzmi następująco:

1. badany człon zaliczamy do klasy przyimków [6], jeżeli przypadek następującego po nim nomen jest zdeterminowany przez człon badany;
2. badany człon zaliczamy do klasy przysłówek [8], jeżeli przypadek następującego po nim nomen jest zdeterminowany przez inny człon, a człon badany można zastąpić innym przysłówkiem. Kryterium to wyjaśnimy na przykładzie kilku spośród przytaczanych członów, cytując pod:
  - a) klasyfikację danego członu jako przyimka [6],
  - b) klasyfikację jako przysłówka [8].

Około:

1. *utwory te miały około [6] dwustu nagrań; ankieta dała około [6] tysiąca odpowiedzi; w strajku brało udział około [6] miliona pracowników; około [6] południa zaczęło zbierać się na deszcz;*
2. człon *około* można zastąpić wyrazami *mniej więcej, prawie*: *drużyna wróciła do kraju z około [8] dwustu punktami; cała liczba tomów równa się sumarycznie około [8] ośmiu tysiącom stron; artykuły przemysłowe stanowią około [8] czterdzieści procent obrotów*.

Blisko:

1. *pociąg zatrzymał się blisko [6] stacji; byli już blisko [6] celu swej wędrowki;*
2. człon *blisko* można zastąpić wyrazami *mniej więcej, serdecznie, wnikliwie*: *wydobyto blisko [8] trzy tysiące ton węgla; blisko [8] połowę załogi stanowili kolorowi; byli z sobą blisko [8] już od wielu lat; trzeba go znać bardzo blisko [8], aby móc go ocenić*.

Ponad:

1. *istnieją obecnie próby wznoszenia komunikacji ponad [6] ziemię; ponad [6] miastem unosiły się ciężkie chmury;*
2. człon *ponad* można zastąpić wyrazem *przeszło*: *drużyna wróciła do kraju z ponad [8] dwustu punktami; na eksport przeznaczają się ponad [8] dziesięć procent produkcji; przeciętnie gospodarstwa te posiadają ponad [8] sto hektarów ziemi*.

Wokół:

1. *dyskusja rozwinęła się wokół [6] poruszonego w referacie zagadnienia; wokół [6] stadionu zbierały się tłumy;*
2. człon *wokół* jest członem określającym przy czasowniku: *nie orientował się, co się wokół [8] dzieje*.

Analogiczne kryteria dotyczą członów *obok, dokoła, poniżej, powyżej*.

Na zakończenie uwag dotyczących opisu liczebników oraz omówionych powyżej typów przysłówek/przyimków trzeba przypomnieć, iż decyzja rozpisywania w formie słownej występujących w tekstach liczebników w znacznej mierze

uzależnienia formy rozpisywane od indywidualnej wymowy maszynistek, co — ze względu na dużą chwiejność w zakresie uzusu — sprawia, iż frekwencje pewnych form liczebników uznać trzeba za mało dokładne.

### 3.2.9. Przyimek czy spójnik (typ Prp/Coni. = 6/9)

Homonimia ta dotyczyła członów *podczas* oraz *mimo* i *pomimo*. Kwalifikowano je jako przyimki [6], gdy dany człon determinował przypadek następującego po nim nomen: *podczas* + gen., *mimo* + gen., acc., *pomimo* + gen., acc. (*podczas zimy, mimo gniewu, mimo to, pomimo licznych próśb, pomimo to*). Jako spójniki [9] uznano natomiast w całości hasła: *podczas gdy, podczas kiedy, mimo że, mimo iż, pomimo że, pomimo iż*.

### 3.2.10. Spójnik czy partykuła (typ Coni./Part. =9/8)

Rozróżnienie to zastosowano tylko do morfemów *by*, *aby*, zaniechano go przy morfemie *a* ze względu na trudność znalezienia dostatecznie wyrazistego kryterium. *By* oraz *aby* są kwalifikowane jako spójniki [9], gdy zaczynają zdania podrzędne okolicznikowe celu lub zdania dopełnieniowe: *pojechał na stację, by [9] spotkać się na dworcu ze znajomym; prosił by [9] go możliwie szybko odwiedzić; proszono uczestników wycieczki, aby [9] nie oddalali się od miejsca zbiórki; wyjechał, aby [9] uregulować swoje sprawy majątkowe*.

Morfem *by* kwalifikowano jako partykułę [8], gdy pełnił funkcję ruchomego morfemu w formach trybu warunkowego: *on by [8] to zrobił, trzeba by [8] się na to zdecydować*. Równocześnie leksem danego czasownika otrzymuje kwalifikator trybu warunkowego z ruchomą partykułą (*zrobił* [54]). Natomiast morfem *aby* uznawano za partykułę [8] w wypadku, gdy w danym ciągu syntaktycznym morfem ten można było wyzerować: *czy aby [8] nie za późno na wizytę ~ czy nie za późno na wizytę*.

### 3.3. Hasła kodowane umownie

Zbyt mała liczba (63) symboli kodu fleksyjnego nie pozwalała na bardziej szczegółowy opis gramatyczny. Jednakże w kilkunastu wypadkach, chcąc objąć kodowaniem pewne formy nie uwzględnione w spisie rejestrowanych cech gramatycznych, wykorzystano przyjęte symbole w sposób całkowicie umowny, niezgodny z przyjętymi znaczeniami kodu. O niektórych tego typu wypadkach już wspomniano powyżej.

Chcąc ułatwić Czytelnikom korzystanie z opublikowanych materiałów, podajemy kodowanie umowne hasła w formie listy alfabetycznej, co pozwoli odszukać szybko odmienne znaczenie danego symbolu przy określonym hasle.

aby	8	w kontekstach typu <i>czy aby</i> [8] <i>nie za późno</i> ;
co	61-64	przy miarach czasu i odległości, gdy jest możliwa substytucja członu <i>każdy</i> : <i>co</i> [61] <i>godzina, co</i> [64] <i>godzinę, co</i> [62] <i>roku</i> ;
co	8	w kontekstach typu: <i>wielu co</i> [8] <i>ambitniejszych, nie ma się co</i> [8] <i>dziwić</i> ;
co	9	w funkcji zaimka względnego: <i>ten, co</i> [9] <i>nie miał szczęścia; ci, co</i> [9] <i>dobiegli pierwsi</i> ;
ile	8	w kontekstach typu: <i>było tyle</i> [8] <i>wody, ile</i> [8] <i>się jej w wiadrze zmieści</i> ;

	jak	8	w funkcji zaimka pytajnego: <i>jak</i> [8] <i>to zrobić?</i> z <i>jak</i> [8] <i>dużą dokładnością?</i> ; także jako partykuła w kontekstach typu: <i>udało mi się i to jak</i> [8];
	jak	9	porównawcze: <i>gryka jak</i> [9] <i>śnieg biała</i> ; zaczynające zdanie niepytające: <i>jak</i> [9] <i>zdobędę skierowanie, przyjadę</i> ; <i>jak</i> [9] <i>wzeszło słońce, wojska ruszyły</i> ; + przysłówki w stopniu równym: <i>jak</i> [9] <i>dalece</i> , <i>jak</i> [9] <i>dotychczas</i> ;
	jako	61-66	w kontekstach typu: <i>on jako</i> [61] <i>znawca</i> , <i>jego jako</i> [64] <i>znawcę</i> ; z <i>nim jako</i> [65] <i>ze znawcą</i> ;
	jeden		bez kodu: w nieodmiennych przydawkach liczebnikowych odpowiadających na pytania który? jaki? — przykłady: <i>zwyciężył w czasie dziewięć godzin dwadzieścia jeden minut</i> ; w <i>latach sześćdziesiąt jeden sześćdziesiąt dwa</i> ;
	jeden	8	jako składnik odmiennego liczebnika głównego: w <i>dwudziestu jeden</i> [8] <i>krajach</i> ; <i>delegacje trzydziestu jeden</i> [8] <i>państw</i> ;
liczebniki główne			bez kodu: jako nieodmienne człony w zestawieniach liczebników porządkowych (w <i>sto dwudziestym pułku</i> ; w <i>tysiąc dziewięćset piątym roku</i> ); w nieodmiennych przydawkach liczebnikowych odpowiadających na pytanie który? jaki? (zwyciężył w <i>czasie dziesięć godzin dwadzieścia minut piętnaście sekund</i> ; <i>Stal pokonała Wartę trzy dwa</i> );
	niczym	9	w kontekstach typu: <i>niczym</i> [9] <i>słońce wszędzie rosa oczy wyje</i> ;
	raz	8	w znaczeniu <i>jednokrotnie</i> : <i>raz</i> [8] <i>na cztery lata</i> ; <i>jeszcze raz</i> [8] <i>spróbujemy</i> ;
	razem	8	jeżeli jest możliwa substytucja wyrazu <i>wspólnie</i> : <i>odkrycia tego dokonali Rosjanie razem</i> [8] z <i>Amerikanami</i> ;
	się	41	formy nieosobowe przy czasownikach niezwrotnych: <i>mówi się</i> [41] <i>po plaży chodzi się</i> [41] <i>boso</i> ;
	temu	8	przy miarach czasu: <i>godzinę temu</i> [8]; <i>kilka lat temu</i> [8];
	tym	9	w połączeniach spójnikowych <i>im. . . tym</i> , + stopień wyższy przysłówka lub przymiotnika: <i>tym</i> [9] <i>bardziej</i> , <i>tym</i> [9] <i>większy</i> ;
	ułamki		liczebniki porządkowe jako mianowniki ułamków mają kwalifikatory rzeczownika: <i>jedna setna</i> [1];
	z	8	w kontekstach typu: <i>było tego z</i> [8] <i>dziesięć sztuk</i> , <i>było tego ze</i> [8] <i>sto par</i> ;
	za	8	w kontekstach typu: <i>co za</i> [8] <i>pech</i> ;

#### Wykaz symboli używanych w opisie morfologicznym słowoform

- 111 rzeczownik w mianowniku l.poj.
- 112 rzeczownik w mianowniku l.mn.
- 121 rzeczownik w dopełniaczu l.poj.
- 122 rzeczownik w dopełniaczu l.mn.
- 131 rzeczownik w celowniku l.poj.

- 132 rzeczownik w celowniku l.mn.
- 141 rzeczownik w bierniku l.poj.
- 142 rzeczownik w bierniku l.mn.
- 151 rzeczownik w narzędniku l.poj.
- 152 rzeczownik w narzędniku l.mn.
- 161 rzeczownik w miejscowniku l.poj.
- 162 rzeczownik w miejscowniku l.mn.
- 171 rzeczownik w wołaczu l.poj.
- 172 rzeczownik w wołaczu l.mn.
- 211 przymiotnik w mianowniku l.poj.
- 212 przymiotnik w mianowniku l.mn.
- 221 przymiotnik w dopełniaczu l.poj.
- 222 przymiotnik w dopełniaczu l.mn.
- 231 przymiotnik w celowniku l.poj.
- 232 przymiotnik w celowniku l.mn.
- 241 przymiotnik w bierniku l.poj.
- 242 przymiotnik w bierniku l.mn.
- 251 przymiotnik w narzędniku l.poj.
- 252 przymiotnik w narzędniku l.mn.
- 261 przymiotnik w miejscowniku l.poj.
- 262 przymiotnik w miejscowniku l.mn.

- |    |                          |    |                       |
|----|--------------------------|----|-----------------------|
| 31 | liczebnik w mianowniku   | 41 | zaimek w mianowniku   |
| 32 | liczebnik w dopełniaczu  | 42 | zaimek w dopełniaczu  |
| 33 | liczebnik w celowniku    | 43 | zaimek w celowniku    |
| 34 | liczebnik w bierniku     | 44 | zaimek w bierniku     |
| 35 | liczebnik w narzędniku   | 45 | zaimek w narzędniku   |
| 36 | liczebnik w miejscowniku | 46 | zaimek w miejscowniku |

- 5 czasownik w formie nie wyróżnionej niżej, o ile dana forma jest homonimiczna (np. *napaść*[5]/*napaść*[1])
- 501 czasownik zwrotny w formach nie wyróżnionych niżej
  - 51 bezokolicznik jako forma składowa form czasu przyszłego, czasownik niezwrotny (*będzie pisać*[51])
  - 511 jak 51, ale czasownik zwrotny (*będzie się bawić*[511])
  - 52 forma na -ł jako składowa form czasu przyszłego, czasownik niezwrotny (*będzie pisał*[52])
  - 521 jak 52, ale czasownik zwrotny (*będzie się bawił*[521])
  - 53 forma na -ł jako czas przeszły z ruchomą końcówką, czasownik niezwrotny (*kiedyście przyjechali*[53])
  - 531 jak 53, ale czasownik zwrotny (*dlaczegoś się spóźnił*[531])
  - 54 forma trybu warunkowego z ruchomą partykułą, czasownik niezwrotny (*wczoraj bym napisał*[54])

- 541 jak 54, ale czasownik zwrotny (*ale bym się przestraszył*[541])
- 55 opisowa forma trybu rozkazującego 3 osoby, czasownik niezwrotny (*niech pisze*[55])
- 551 jak 51, ale czasownik zwrotny (*niech się bawi*[551])
- 56 czasownik *być* jako składowa form czasów złożonych (*będzie*[56] *писаł*[52], *napisał był*[56])
- 57 czasownik *być*, *bywać*, *zostać* jako składowe form strony biernej (*jest*[57] *napisany*, *bywał*[57] *nabywany*, *zostanie*[57] *napisany*)
- 61 przyimek z mianownikiem
- 62 przyimek z dopełniaczem (z[62] *domu*)
- 63 przyimek z celownikiem (ku[63] *domowi*)
- 64 przyimek z biernikiem (*na*[64] *stół*)
- 65 przyimek z narzędnikiem (z[65] *tobą*)
- 66 przyimek z miejscownikiem (*na*[66] *stole*)
- 7 wykrzyknik
- 8 partykuła lub przysłówek
- 9 spójnik

## Literatura

- [1] Kurcz, Ida; Lewicki, Andrzej; Sambor, Jadwiga; Woronczak, Jerzy. *Słownictwo współczesnego języka polskiego. Listy frekwencyjne. Tom I. Teksty popularnonaukowe*. Warszawa, 1974. Uniwersytet Warszawski.
- [2] Kurcz, Ida; Lewicki, Andrzej; Sambor, Jadwiga; Woronczak, Jerzy. *Słownictwo współczesnego języka polskiego. Listy frekwencyjne. Tom II. Drobne wiadomości prasowe*. Warszawa, 1974. Uniwersytet Warszawski.
- [3] Kurcz, Ida; Lewicki, Andrzej; Sambor, Jadwiga; Woronczak, Jerzy. *Słownictwo współczesnego języka polskiego. Listy frekwencyjne. Tom IV. Proza artystyczna*. Warszawa, 1976. Uniwersytet Warszawski.
- [4] Kurcz, Ida; Lewicki, Andrzej; Sambor, Jadwiga; Woronczak, Jerzy. *Słownictwo współczesnego języka polskiego. Listy frekwencyjne. Tom V. Dramat artystyczny*. Warszawa, 1977. Uniwersytet Warszawski.
- [5] Kurcz, Ida; Lewicki, Andrzej; Sambor, Jadwiga; Szafran, Krzysztof; Woronczak, Jerzy. *Słownik frekwencyjny polszczyzny współczesnej*, Kraków, 1990. Instytut Języka Polskiego PAN.
- [6] Lewicki, Andrzej; Masłowski, Władysław; Sambor, Jadwiga; Woronczak, Jerzy. *Słownictwo współczesnego języka polskiego. Listy frekwencyjne. Tom III. Publicystyka*. Warszawa, 1975. Uniwersytet Warszawski.
- [7] Kurcz, Ida; Lewicki, Andrzej; Sambor, Jadwiga; Szafran, Krzysztof; Woronczak, Jerzy. *Słownik frekwencyjny polszczyzny współczesnej*, Kraków, 1990. Instytut Języka Polskiego PAN.
- [8] Nazarczuk, Marta. *Wstępne przygotowanie korpusu „Słownika frekwencyjnego polszczyzny współczesnej” do dystrybucji na CD-ROM*. Praca magisterska napisana pod kierunkiem dra hab. Janusza S. Bienia. Warszawa, 1997. Instytut Języka Polskiego Uniwersytetu Warszawskiego. 59 s., płyta CD.