

WITOLD KIERAŚ, MARCIN WOLIŃSKI*

INSTYTUT PODSTAW INFORMATYKI POLSKIEJ AKADEMII NAUK, WARSZAWA

*Słownik gramatyczny języka polskiego – wersja internetowa*¹

Słowa kluczowe: fleksja, paradygmatyka, słownik.

W artykule przedstawimy nowe wydanie *Słownika gramatycznego języka polskiego* (SGJP). Inaczej niż jego dwa poprzednie wydania (Saloni i in. 2007b, 2012), które były udostępniane w postaci programu instalowanego w komputerze (zob. Saloni i in. 2007a), teraz słownik jest w całości dostępny w Internecie, a jego funkcje użytkowe zostały rozszerzone. Skupimy się przede wszystkim na widocznych dla użytkownika zmianach w stosunku do poprzednich wersji.

Koncepcja stworzenia słownika gramatycznego języka polskiego powstała co najmniej 35 lat temu, a bezpośrednią inspiracją dla takiego pomysłu był niewątpliwie *Słownik gramatyczny języka rosyjskiego* A. Zalizniaka (1977/2003). Prace przygotowawcze mające doprowadzić do jego powstania obejmowały stopniowo formalny opis deklinacji polskich rzeczowników pospolitych (Gruszczyński 1989), koniugacji (Saloni 2001), aż wreszcie zaowocowały pierwszym wydaniem SGJP w 2007 roku. W przeciwieństwie do wcześniejszych prac SGJP od początku funkcjonował jako słownik elektroniczny, a nie papierowy. Przeniesienie go do Internetu rozszerza grono jego potencjalnych użytkowników i stwarza możliwość szybkiej i łatwej aktualizacji jego danych.

Autorami najnowszego wydania słownika są Zygmunt Saloni, Włodzimierz Gruszczyński, Robert Wołosz, Marcin Woliński i Danuta Skowrońska. Głównym wykonawcą aplikacji webowej prezentującej SGJP jest Jan Szejkó.

1. Materiał słownika

SGJP dysponuje rozległą podstawą leksykalną. Zasadnicza jego część składa się z listy haseł SJPDor, którego zasób słownictwa sięgał nawet ostatnich dziesięcioleci XVIII wieku, w związku z tym słownik zawiera sporo słownictwa przestarzałego, dziś nieużywanego nawet w języku literackim. Z drugiej strony jednak materiał SGJP był uzupełniany przez jego autorów na różnych etapach prac, w tym również o słownictwo najnowsze. W efekcie zatem materiał SGJP jest obecnie znacznie obszerniejszy niż SJPDor i obejmuje ponad 330 000 haseł zawierających w swoich paradygmatach łącznie blisko 4 300 000 form wyrazowych generowanych przez

* wkieras@ipipan.waw.pl, wolinski@ipipan.waw.pl

¹ Współautorstwo obejmuje wszystkie etapy pracy nad artykułem. Autorzy dziękują prof. Zygmuntowi Saloniemu za cenne uwagi redakcyjne i pomoc w doborze przykładów.

1114 ręcznie tworzonych wzorów odmiany. Szczegółowe różnice liczbowe między pierwszym wydaniem SGJP z 2007 roku a obecnym stanem w wersji online słownika z podziałem na klasy gramatyczne prezentuje tabela 1.

Tabela 1. Szczegółowe porównanie ilościowe zmian w SGJP między pierwszym wydaniem a obecną wersją internetową

	1. WYDANIE		3. WYDANIE (ONLINE)	
	HASEŁ	WZORÓW	HASEŁ	WZORÓW
w sumie	244 669	1 095	334 845	1 114
prefiksy	81	2	112	2
leksemy	244 588	1 095	334 733	1 116
o odmianie rzeczownikowej	135 529	762	172 178	767
zaimki osobowe	6	6	6	6
odśłowniki	29 590	2	29 526	2
odprzym. nazwy cech na <i>-ość</i>	28 980	1	62 445	1
pozostałe rzeczowniki	76 953		80 201	
nazwy własne	8 782		10 710	
pospolite	68 171		69 491	
o odmianie przymiotnikowej	65 671	71	103 761	77
imiesłowy	34 301		36 304	
czynne	13 931	1	13 877	1
bierne	20 370		22 427	4
przymiotniki	31 370	71	67 457	77
stopnia wyższego	950	1	1 065	1
stopnia równego	30 420	70	66 392	77
przysłówki odprzymiotnikowe	11 146	1	26 577	1
stopnia wyższego	1 106	1	1 243	1
stopnia równego	10 040	1	25 512	1
liczebniki	98	45	117	47
czasowniki	29 532	215	29 955	222
predykatywy	35	1	34	1
odmienne	29 497	214	29 925	221
pozostałe	2 612	2	1 967	2
pozostałe przysłówki	491	1	502	1
partykuły	193	1	201	1
przyimki	113	2	118	2
spójniki	121	1	121	1
wykrzykniki i zbliżone	458	1	490	1
skrót	1 117	1	392	17
inne	119	1	143	1

Inaczej niż większość innych słowników SGJP nie objaśnia znaczeń wyrazów. Co prawda w części haseł pojawiają się różnego rodzaju informacje sygnalizujące znaczenie lub przybliżające je, ale spełniają one jedynie zadanie pomocnicze. Konstytutywną cechą hasła słownikowego w SGJP jest postać paradygmatu odmiany i wypełnienie jego klatek formami gramatycznymi – w konsekwencji dwa różne leksemy o identycznej odmianie będą opisane w jednym hasle. Przykładowo, dwa leksemy MATKA o znaczeniach ‘kobieta mająca dziecko’ oraz ‘mała mata’, słowotwórczo zupełnie niepowiązane, są opisane w jednym hasle, ponieważ ich paradygmaty niczym się od siebie nie różnią. Z drugiej strony istnieją w SGJP trzy leksemy TYTAN: 1. ‘przen. osoba bardzo silna’, 2. ‘rodzaj chrząszcza’, 3. ‘pierwiastek chemiczny’, różniące się rodzajem gramatycznym i w konsekwencji również postacią paradygmatu – układem synkretyzmów form biernikowych, a także obecnością form męskoosobowych w mianowniku i wołaczu liczby mnogiej (w wypadku pierwszego z nich) i zakończeniem dopełniacza liczby pojedynczej (-u w trzecim znaczeniu, -a zaś w pozostałych).

W przeciwieństwie do słowników ogólnych lub specjalistycznych zawierających informacje o odmianie, które podają zwykle tylko wybrane formy lub zakończenia niektórych charakterystycznych dla leksemu form, SGJP prezentuje wszystkie formy fleksyjne w postaci osobnych tabel dla każdego leksemu. Jest to główne zadanie słownika – wszelkie inne informacje, np. na temat nacechowania stylistycznego, zakresu tematycznego, znaczenia czy pochodzenia, są wyłącznie pomocnicze i nie mają charakteru systematycznego.

2. Kuźnia

W trakcie prac nad pierwszym i drugim wydaniem słownika autorzy posługiwali się sześcioma bazami danych programu Microsoft Access, z których każda miała inną budowę, dostosowaną do specyfiki wzorów poszczególnych typów. W efekcie każda baza miała swojego opiekuna i tylko on dokonywał w niej zmian. Wtedy, gdy te bazy powstawały, nie wszyscy autorzy dysponowali stałym szerokopasmowym dostępem do Internetu i taki tryb pracy wydawał się najpraktyczniejszy. Z czasem jednak widać było również zasadnicze ograniczenia tego sposobu pracy, które hamowały możliwość szybkiej aktualizacji haseł lub poprawiania błędów dostrzeżonych przez użytkowników. Przeniesienie SGJP do Internetu, przy jednoczesnym założeniu, że słownik będzie rozwijany, wymagało stworzenia zupełnie nowych narzędzi nie tylko do prezentacji danych, ale i do ich redagowania, jak również scalenia osobnych baz dla różnych części mowy w jedną dużą bazę danych. W efekcie powstała aplikacja webowa o nazwie Kuźnia (Szejko 2014), dzięki której redaktorzy dysponujący odpowiednimi uprawnieniami mogą dokonywać zmian w całym materiale słownika, a efekt ich pracy jest widoczny od razu. Kuźnia stanowi zatem interfejs edycyjny dla nowej bazy danych, jak również interfejs czytelnika. W tym drugim wypadku jest to oczywiście poziom dostępu ograniczony jedynie do oglądania zawartości słownika (czyli haseł i wzorów), nie umożliwia zaś jego edycji.

Nowy interfejs edycyjny zasadniczo zmienił sposób pracy jego redaktorów. W poprzedniej wersji SGJP redaktor tworzący nowe hasło musiał utworzyć rekord w bazie danych, nie mógł jednak od razu zobaczyć bezpośredniego efektu swojej pracy w postaci tabeli odmiany danego leksemu, a co za tym idzie – nie mógł też sprawdzić poprawności paradygmatu. W wersji online autor nie tylko może zobaczyć, w jaki sposób zmienia się paradygmat odmiany leksemu pod

wpływem zmiany wzoru, ale może też liczyć na odpowiedź, jaki wzór wybrać. Narzędzie podpowiadające wzory odmiany dla wpisanej formy hasłowej proponuje je na podstawie zakończenia tej formy oraz informacji o tym, jak odmieniają się hasła obecne już w bazie o takich samych zakończeniach oraz – w wypadku rzeczowników – o takim samym rodzaju gramatycznym. Jest to szczególnie przydatne przy wprowadzaniu do słownika słownictwa najnowszego, bazującego na produktywnych formantach. Przykładowo, słowu **TOFUCZNICA** ‘danie z tofu naśladujące jajecznicę’, któremu leksykograf przypisał część mowy i rodzaj, mechanizm podpowiadania wzorów w pierwszej kolejności przypisał taki sam wzór, jaki w słowniku ma słowo **KLUCZNICA** (a także m.in. **JAJECZNICA**). Jest to zatem odpowiedź poprawna, choć – jak widać na rysunku 1 – dostępne są również inne propozycje, w tym m.in. dwie o odmianie przymiotnikowej. Nie jest to oczywiście mechanizm niezawodny, a jedynie pomoc dla redaktora ograniczająca liczbę błędów związanych z przypisaniem wzoru odmiany.

Kuźnia jest aplikacją, która może być używana nie tylko do rozwijania SGJP, ale też do wszelkich innych słowników fleksyjnych polszczyzny współczesnej i historycznej zgodnych z modelem SGJP. Wymaga oczywiście dostarczenia odpowiednich danych leksykalnych oraz bazy wzorów odmiany.

Podpowiadacz wzorów

klucznica ·a f	nazwa pospolita	D2+(-fneut)	sg:nom	tofucznica
bejc·a f	nazwa pospolita	D2	sg:gen	tofucznicy
chodząc·a f	nazwa pospolita	A4Ż	sg:dat	tofucznicy
ow·ca f	nazwa pospolita	D2cei	sg:acc	tofucznice
abatys·a f	nazwa pospolita	D4	sg:inst	tofucznica
abiologi·a f	nazwa pospolita	D1	sg:voc	tofucznico
abrakadabr·a f	nazwa pospolita	D4r	pl:nom	tofucznice
Abramowsk·a f	nazwisko	A3Ż	pl:gen:fneut	
absyd·a f	nazwa pospolita	D4d	pl:gen:fchar	tofucznica
acerol·a f	nazwa pospolita	D1+(ów)	pl:dat	tofucznicom
			pl:inst	tofuczniciami
			pl:loc	tofucznicach

☒ Uwzględniaj rodzaj
 ☐ Uwzględniaj pospolitość
 ☒ Pomijaj wzory nietypowe

Anuluj

Wybierz

Rysunek 1. Podpowiadanie wzorów

3. Filtry

Dla użytkowników SGJP prawdopodobnie najistotniejszą nowością jest znacznie bardziej rozbudowany system filtrów. We wcześniejszych wersjach słownika użytkownik mógł wyszukiwać formy hasłowe lub gramatyczne w całym słowniku, ewentualnie ograniczając wyszukania do klas gramatycznych (czasowników, rzeczowników itd.). W wersji internetowej możliwości wyszukiwania i filtrowania wyników są znacznie większe. Użytkownik może szukać leksem ze względu na kształt jego formy hasłowej lub dowolnej formy w jego paradygmacie, w szczególności zaś ze względu na zawieranie odpowiednich podciągów znaków czy odpowiadanie wzorcowi zadanemu przez wyrażenie regularne². Może również wyszukiwać leksemy ze względu na to, według jakich wzorów lub typów wzorów się odmieniają. Wreszcie, może wyszukiwać leksemy ze względu na ich cechy gramatyczne – przynależność do odpowiedniej klasy gramatycznej, wartość kategorii rodzaju (dla rzeczowników) lub aspektu (dla czasowników i ich derywatów), przechodniość.

Możliwość tworzenia kryteriów filtrowania dotyczy również kwalifikatorów i informacji o tym, czy dany rzeczownik jest nazwą własną, a jeśli tak, to jakiego typu. Należy przy tym pamiętać, że kwalifikatory stylistyczne i zakresowe oraz typologia nazw własnych w słowniku stanowią jedynie informację pomocniczą, mającą na celu ułatwienie rozszyfrowania znaczeń, zwłaszcza w wypadku haseł homonimicznych, nie stanowią jednak systematycznej i w pełni uporządkowanej informacji słownikowej.

Różne kryteria filtrowania można ze sobą łączyć, definiując w jednym zapytaniu kilka warunków połączonych za pomocą koniunkcji lub alternatywy, czyli w taki sposób, że każdy wyszukany leksem musi spełniać każdy warunek albo każdy leksem musi spełniać przynajmniej jeden z warunków.

Żałujemy, że chcemy odnaleźć w słowniku wszystkie rzeczowniki żeńskie zakończone na *-yni* typu JĘZYKOZNAWCZYNI, ZDOBYWCZYNI, będące derywatami od odpowiednich rzeczowników męskich. W tym celu można nałożyć filtr na klasę gramatyczną (rzeczownik) i rodzaj (f). W tym konkretnym wyszukiwaniu określenie klasy gramatycznej jest zresztą nadmiarowe, ponieważ jedyną – oprócz właściwych rzeczowników – wyodrębnioną w SGJP klasą leksemów, które mają przypisany rodzaj gramatyczny, są odsłowniki, nie mogą one jednak mieć rodzaju żeńskiego i nie charakteryzują się wymienionym wyżej zakończeniem. Następnie można dodać warunek na postać formy hasłowej: *Hasło kończy się na yni*. Użytkownik powinien uzyskać w ten sposób listę 130 rzeczowników żeńskich o poszukiwanym zakończeniu – ich lista pojawi się w oknie haseł po lewej stronie, a dokładna ich liczba zostanie wyświetlona na górnym pasku (w tytule karty) przeglądarki internetowej. Jeśli chcemy rozszerzyć zapytanie tak, by lista objęła również rzeczowniki zakończone na *-ini* typu CZŁONKINI, BOGINI, nie uzyskamy pożądanego efektu przez dodanie kolejnego warunku opartego na postaci zakończenia formy hasłowej, ponieważ te dwa warunki wykluczają się wzajemnie (żaden wyraz nie może mieć jednocześnie zakończeń *-yni* i *-ini*). Nie można też zmienić sposobu łączenia warunków

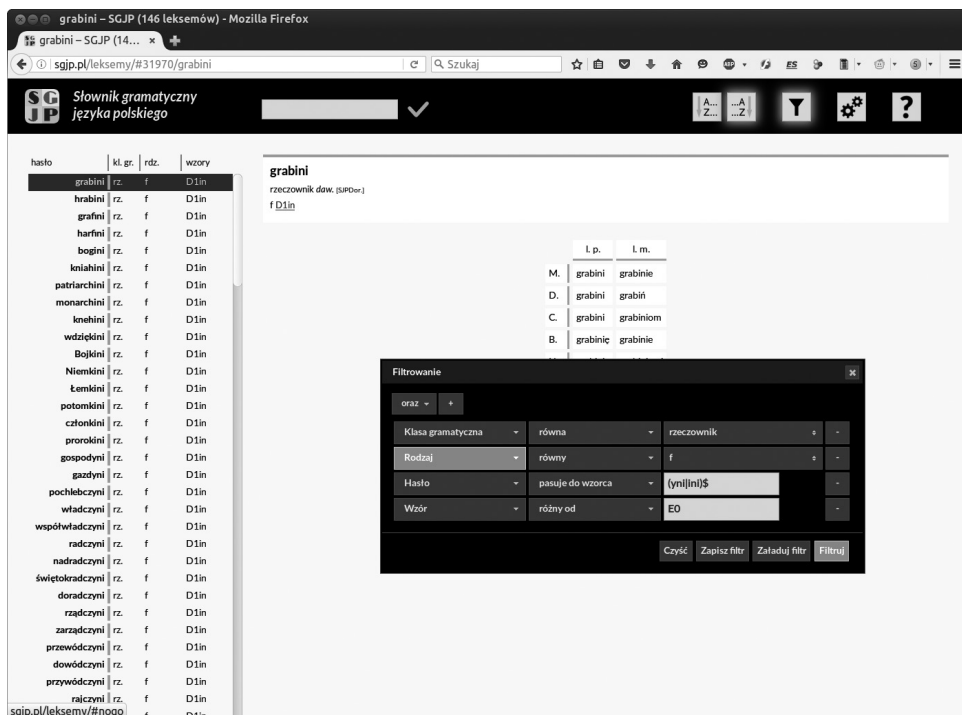
2 Wyrażenia regularne to wzorce, zapisane w zwięzły, formalny sposób, opisujące ciągi znaków pewnego typu.

z koniunkcji na alternatywę (nawet po pominięciu nadmiarowego warunku opartego na kategorii gramatycznej), ponieważ do zbioru wynikowego wpadłyby zarówno wszystkie rzeczowniki żeńskie, jak i wszystkie leksemy o formie hasłowej zakończonej na *-yni* (np. BURSZTYNI, DZIEWCZYNI) lub *-ini* (np. MARTINI, BIKINI, PAGANINI). Możemy oczywiście wywołać dwa oddzielne zapytania (każde dla jednego z dwu zakończeń) i obejrzeć wyniki obu wyszukiwań, co jednak może okazać się mało wygodne. Dlatego ciekawszą możliwością będzie zmodyfikowanie zapytania w taki sposób, by połączyć w jednym warunku dwa typy zakończeń. W tym celu należy skorzystać z opcji *Hasło pasuje do wzorca...* i jako wzorzec podać: $(yni | ini) \$$. Wzorzec jest wyrażeniem regularnym, zapisanym według powszechnie stosowanej konwencji opisanej w wielu miejscach, w tym również w polskiej Wikipedii. Ujęte w nawiasy okrągłe wyrażenie składa się z dwu zakończeń oddzielonych znakiem pionowej kreski oznaczającej alternatywę oraz znaku $\$$ mówiącego, że szukane leksemy mają mieć formy hasłowe o zakończeniu określonym bezpośrednio przed tym znakiem. Dzięki tak zmodyfikowanemu filtrowi lista wyszukanych leksemów powiększy się do 147, ale pojawi się na niej jeden niepożądany element: rzeczownik MINI o podwójnej kwalifikacji rodzajowej (żeńsko-nijakiej) – spełnia on wszystkie warunki formalne filtru, lecz nie jest to derywat żeński i w związku z tym nie pasuje do grupy rzeczowników, których szukamy. Przy konstruowaniu nieco bardziej złożonych filtrów dość często zdarza się, że na liście wyników pojawiają się tego typu elementy niepożądane, które można jednak eliminować na różne sposoby. W tym wypadku można np. dodać warunek wykluczający leksemy, którym przypisano jakikolwiek inny rodzaj niż żeński (filtr: *Rodzaj różny od...*), lub wykluczający rzeczowniki formalnie nieodmienne, które zawsze mają przypisany wzór E0 (*Wzór różny od E0*). Dzięki temu na liście wyników uzyskamy 146 leksemów spełniających warunki określone na początku (por. rysunek 2).

Użytkownik łatwo zauważy, że wszystkie leksemy na liście odmieniają się według tego samego wzoru: D1in, czyli wzoru odmiany typu żeńskiego o temacie zakończonym na spółgłoskę miękką, charakteryzującym się w zapisie wymianą końcowego *ni* w formie podstawowej na *ń* (w dopełniaczu liczby mnogiej). Można zatem spróbować sprawdzić, czy wśród derywatów żeńskich tego typu znajdują się takie, które mają zakończenie *-ni*, ale nie *-yni* lub *-ini*. W tym celu można odfiltrować wszystkie leksemy żeńskie, które odmieniają się według tego wzoru – powinno ich być 153, czyli o siedem więcej niż w poprzednim wyszukiwaniu. Istnieje zatem siedem leksemów o takiej cesze, można je znaleźć, rozszerzając filtr o warunek niepasowania do wzorca $(yni-ini) \$$. Na liście powinny znaleźć się rzeczowniki: ACANI, ASANI, KSIENI, ŁANI, MAHARANI, WACANI, WASANI. Pięć z nich to faktycznie derywaty żeńskie odpowiednio od rzeczowników: ACAN, ASAN, MAHARADŹA, WACAN i WASAN. Pozostałe dwa zaś – KSIENI i ŁANI – nie mają swoich męskich odpowiedników, choć reprezentują ten sam typ odmiany.

W podobny sposób można wyszukać np. wszystkie rzeczowniki męskie, które w wołaczu mają (jako oboczną lub jako jedyną) nietypową formę na *-cze*, np. GŁUPIEC (*głupcze*), SZEWC (*szewcze*). W tym wypadku wystarczy zdefiniować filtr ograniczający wyniki do rzeczowników o rodzaju innym niż żeński, nijaki lub przymnogi (czyli w praktyce w jednym z trzech rodzajów męskich), których forma hasłowa kończy się na *-c* i które mają w swoim paradygmacie formę kończącą się na *-cze*. Na liście powinno znaleźć się ponad sześćset takich leksemów. W tym

wypadku użytkownik może zauważyć, że wszystkie leksemy na liście będą miały przypisany wzór zawierający znak wykrzyknika (!). Znak ten informuje o tym, że wzór uwzględnia nietypową końcówkę wołacza liczby pojedynczej (nie tylko końcówkę *-cze*, ale też np. końcówkę *zerową*; zob. Saloni 2016). Oczywiście nie wszystkie zjawiska powierzchniowofleksyjne są opisane za pomocą odpowiednich symboli we wzorze, ale w wielu wypadkach mogą się one okazać przydatne przy wyszukiwaniu obszernych list leksemów o konkretnych cechach.



Rysunek 2. Przykład filtru składającego się z czterech warunków, z których trzeci ilustruje możliwość wykorzystania wyrażeń regularnych do przeszukiwania zawartości słownika

Filtry w postaci dostępnej w nowej wersji SGJP dają użytkownikowi znacznie większe możliwości niż w wersjach wcześniejszych. Oczywiście wraz z siłą wyrazu zwiększa się też stopień skomplikowania struktury filtrów. Niektóre decyzje związane z filtrami były również podyktowane względami praktycznymi, przede wszystkim szybkością przeszukiwania obszernej bazy danych.

4. Wzory

Ważną zmianę w stosunku do poprzednich wydań SGJP stanowi prezentacja wzorów odmiany jako integralnej i pełnoprawnej części słownika. O ile wcześniej wzory odmiany z punktu widzenia użytkownika były jedynie enigmatycznymi etykietami przypisanymi do haseł,

których szczegółowe znaczenie dostrzegali jedynie najbardziej wnikliwi czytelnicy, o tyle teraz stanowią one komponent, który użytkownik może przeglądać i wykorzystywać przy wyszukiwaniach (pokazano to wyżej). Same oznaczenia wzorów również uległy zmianie, a system nazewniczy wzorów został opisany i udostępniony na stronie słownika, jak również w postaci osobnego artykułu (por. Saloni 2016). Nowy sposób etykietowania wzorów nawiązuje do oznaczeń grup deklinacyjnych i koniugacyjnych Jana Tokarskiego w postaci znanej z SJPDor i powielanej w wielu późniejszych słownikach, powinien więc być stosunkowo łatwy do zrozumienia dla czytelników obytych z tym tradycyjnym opisem odmiany w polszczyźnie. Wzory SGJP są oczywiście znacznie bardziej szczegółowe i jest ich dużo więcej niż grup Tokarskiego, ponieważ muszą uwzględniać wszystkie możliwe alternacje, często bardzo rzadkie i specyficzne, które w wypadku grup Tokarskiego były opisywane w hasle słownikowym.

B4+e
wzór rzeczownikowy
Typ: m
Sumaryczna liczba leksemów: 48

Rodzaj	m2	m3	p2	p3
Liczba leksemów	3	2	10	44

Przykład: ☒ kontredans ☐ pasjans ☐ Tyszwice ☐ sztućce

Etykieta	Forma bazowa
sg:nom	kontredans-
sg:gen	kontredans-a
sg:dat	kontredans-owi
sg:inst	kontredans-em
sg:loc	kontredans-ie
sg:voc	kontredans-ie
pl:nom	kontredans-e
pl:gen	kontredans-ów
pl:dat	kontredans-om
pl:inst	kontredans-ami
pl:loc	kontredans-ach

Rysunek 3. Widok wzorów w słowniku. Na ilustracji widać wzór B4+e, czyli wzór odmiany męskiej o zakończeniu tematu na spółgłoskę twardą, charakteryzujący się dodatkowo specyficzną końcówką -e w mianowniku liczby mnogiej. Wzór jest używany przez 48 różnych leksemów o czterech różnych rodzajach: dwóch męskich i dwóch przymnognich

Opisany wcześniej system filtrów jest dostępny również w widoku wzorów, choć oczywiście kryteria ich budowania są znacznie skromniejsze – ograniczają się głównie do nazwy wzoru i do typu odmiany. Dzięki nim można np. sprawdzić, ile jest w całym słowniku wzorów czasownikowych, czyli takich, których nazwa zaczyna się od litery V, lub ile wariantów w SGJP ma konkretna grupa koniugacyjna Tokarskiego (np. grupa V została rozbita aż na

27 wzorów, bardzo regularna grupa IV zaś nie wymagała tworzenia żadnych dodatkowych wariantów). Przejrzysty sposób tworzenia etykiet wzorów pozwala też na ustalenie słownikowej frekwencji poszczególnych grup deklinacyjnych i koniugacyjnych w widoku leksemów. Przykładowo, czasowników ze wspomnianej grupy V w SGJP jest ponad cztery razy mniej niż czasowników z grupy IV.

5. Przyszłość słownika

Specyfika słownika elektronicznego funkcjonującego w medium internetowym powoduje, że powinien się on znajdować niemal w ciągłym rozwoju. Dotyczy to zresztą wszelkiego rodzaju innych zasobów i narzędzi lingwistycznych: np. korpusów, gramatyk formalnych, parserów. Gdy przestają być rozwijane i poprawiane, szybko się dezaktualizują lub powiązane z nimi oprogramowanie przestaje działać. Słownik elektroniczny wymaga zatem stałej opieki, która z kolei wymaga nakładów czasu i pracy ze strony jego redaktorów. Boddźcem do pracy jest nie tylko powiększone grono użytkowników publicznie dostępnego słownika internetowego, ale także zainteresowanie ze strony innych projektów, w których jego dane mogą okazać się przydatne. Dane SGJP są wykorzystywane przynajmniej w dwóch takich projektach: stanowią źródło informacji gramatycznej dla *Wielkiego słownika języka polskiego PAN* (Żmigrodzki 2015) oraz są podstawą słownikową analizatora morfologicznego Morfeusz w jednej z jego wersji (Woliński 2014), używanego przez wiele innych programów. W obu tych istotnych zastosowaniach SGJP powinien być wykorzystywany również w perspektywie co najmniej najbliższych kilku lat. Ponadto słownik jest też coraz częściej przytaczany jako źródło informacji gramatycznej w innych zasobach internetowych, zwłaszcza od czasu udostępnienia wersji online. Przykładowo, w Wikisłowniku (<https://pl.wiktionary.org>) SGJP jest przywoływany co najmniej kilkaset razy.

Pomimo swojej rozległej podstawy leksykalnej SGJP ma też pewne obszerne luki, wśród nich są przede wszystkim nazwy własne, które nigdy nie były w nim notowane w sposób systematyczny. Niewielka próbka nazw własnych obecnych w SGJP obejmuje około 11 tysięcy haseł, z czego ponad 4300 to nazwiska i ich człony, blisko 3100 to imiona, a 2700 to różnego rodzaju nazwy geograficzne i ich człony (głównie nazwy miejscowości). Nazwy własne stanowią zatem niespełna 3,3 proc. wszystkich jednostek zanotowanych w słowniku, co nie odzwierciedla faktycznego udziału tego typu słownictwa w tekstach polskich. Uzupełnienie tej luki jest możliwe, wymaga jednak zaplanowania i zrealizowania poważniejszych prac badawczych i leksykograficznych.

Bibliografia

- Gruszczyński W. 1989: *Fleksja rzeczowników pospolitych we współczesnej polszczyźnie pisanej*, Zakład Narodowy im. Ossolińskich, Warszawa–Wrocław.
- Saloni Z. 2001: *Czasownik polski*, Wiedza Powszechna, Warszawa.
- Saloni Z. 2016: *Systematyzacja wzorów fleksyjnych dla „Słownika gramatycznego języka polskiego”*, „Polonica” XXXVI, s. 5–18.
- Saloni Z., Gruszczyński W., Woliński M., Wołosz R. 2007a: *Grammatical dictionary of Polish. Presentation by the authors*, „Studies in Polish Linguistics” 4, s. 5–25.

- Saloni Z., Gruszczyński W., Woliński M., Wołosz R. 2007b: *Słownik gramatyczny języka polskiego*, Wiedza Powszechna, Warszawa.
- Saloni Z., Woliński M., Wołosz R., Gruszczyński W., Skowrońska D. 2012: *Słownik gramatyczny języka polskiego*, wyd. 2, Warszawa.
- SJPDo: *Słownik języka polskiego*, red. W. Doroszewski, t. 1–4, Wiedza Powszechna, Warszawa 1958–1962, t. 5–11, Państwowe Wydawnictwo Naukowe, Warszawa 1963–1969.
- Szejko J. 2014: *Aplikacja webowa do opracowywania słowników fleksyjnych*, praca magisterska, Uniwersytet Warszawski, Warszawa.
- Woliński M. 2014: *Morfeusz reloaded*, [w:] N. Calzolari, K. Choukri, T. Declerck, H. Loftsson, B. Maegaard, J. Mariani, A. Moreno, J. Odijk, S. Piperidis (red.), *Proceedings of the Ninth International Conference on Language Resources and Evaluation, LREC 2014*, ELRA, Reykjavík, Iceland, s. 1106–1111 (online: <http://www.lrec-conf.org/proceedings/lrec2014/index.html>).
- Zalizniak A.A. 2003: *Grammaticzeskij słowar' russkogo jazyka*, wyd. 4, Russkij jazyk, Moskwa.
- Żmigrodzki P. 2015: *Wielki słownik języka polskiego PAN – historia, stan obecny i perspektywy rozwoju po 2018 roku*, „Biuletyn Polskiego Towarzystwa Językoznawczego” LXXI, s. 177–187.

Summary

Grammatical Dictionary of Polish – an online version

Keywords: inflection, surface morphology, dictionary.

In the paper we present a new online version of the *Grammatical Dictionary of Polish*. We focus on new features which have been introduced since the previous editions of the dictionary which were available as a desktop application. The most interesting changes include a full view of inflectional patterns together with their new systematization and significantly extended searching tools allowing to filter entries of the dictionary based on various criteria. Some practical examples of using these new features are presented.