

Klasyfikacja semantyczna przymiotników polskich na podstawie ich własności syntaktycznych w języku polskim i angielskim

Krzysztof Jassem
Adam Mickiewicz University
Faculty of Mathematics and Computer Science
ul. Matejki 48-49
60-769 Poznań

jassem@math.amu.edu.pl

W poniższej pracy zaproponowano klasyfikację przymiotników polskich w oparciu o ich własności składniowo-semantyczne w języku polskim i angielskim. Poszczególne klasy scharakteryzowano semantycznie. Celem tworzenia takiej klasyfikacji – w przyszłych badaniach również dla innych części mowy – jest zaproponowanie nowego podejścia w budowaniu leksykonów gramatyczno-syntaktycznych języka polskiego: zamiast opisywać własności syntaktyczno-semantyczne każdego hasła-lexemu na podstawie żmudnej analizy korpusu tekstów proponuje się podzielenie lexemu na tzw. lexem-sensy i przydzielenie każdego lexem-sensu do odpowiedniej klasy – na przykład poprzez analogię semantyczną do haseł już sklasyfikowanych.

The paper proposes a classification of Polish adjectives on the basis of their syntactical-semantic features in Polish and English. The classes are characterised semantically. The aim of the classification – in future research also for other parts of speech – is to suggest a new approach in building grammatical-syntactical lexicons for Polish: instead of describing syntactical-semantic features of each lexeme-entry on the basis of arduous analysis of a text corpus, it is proposed to divide lexemes into so-called lexeme-senses and assign each lexeme-sense to an appropriate class – for example by semantic analogy to the entries already classified.

1. Motywacja klasyfikacji

Komputerowa analiza syntaktyczna i semantyczna języka naturalnego zakłada istnienie leksykonu gramatyczno-syntaktycznego, w którym dla każdego wyrazu podane są jego możliwe kolokacje syntaktyczno-semantyczne. Przez kolokacje syntaktyczne rozumie się współistnienie wyrazu w tekście z określonymi kategoriami gramatycznymi, natomiast kolokacje syntaktyczno-semantyczne to współistnienie wyrazu z wyrażeniami o określonej kategorii gramatycznej i określonym znaczeniu. Na przykład lexem *dać* (coś komuś) tworzy kolokacje syntaktyczne z wyrażeniem rzeczownikowym w bierniku i wyrażeniem rzeczownikowym w celowniku, oraz kolokacje syntaktyczno-semantyczne z wyrażeniem

rzeczownikowym – nieosobowym w bierniku oraz wyrażeniem rzeczownikowym – osobowym w celowniku.

Dla większości typów analizy (choć nie wszystkich – kontrprzykładem może być tzw. Link Grammar [1]) wystarczające jest podanie dla każdego hasła kolokacji z kategoriami mu podrzędnymi. Zbiór takich kolokacji będziemy nazywać własnościami syntaktyczno-semantycznymi.

Celem podjętych działań jest usystematyzowanie wyrazów języka polskiego pod względem ich własności syntaktyczno-semantycznych. W dotychczasowych próbach opisu polskich części mowy (patrz [2], [3], [4]) stosowano następującą metodę: dla każdego leksemu z osobna wymieniano zbiór jego możliwych kolokacji otrzymany na podstawie analizy pewnego korpusu tekstów. W tej pracy proponuje się inne podejście do tego zagadnienia. Polega ono na utworzeniu dla każdej części mowy klasyfikacji opartej na własnościach syntaktyczno-semantycznych. Do jednej klasy należeć będą te jednostki, które tworzą kolokacje z identycznymi kategoriami gramatycznymi.

Można spodziewać się, że tak utworzone klasy miałyby niewielką licznosc. Na przykład wyrazy bliskoznaczne „uważać” i „sądzić” nie mogłyby znaleźć się w tej samej klasie; co prawda tworzą one podobne konstrukcje typu „uważać (sądzić coś), uważać (sądzić) coś o kimś”, „uważać (sądzić, że”, ale tworzą też kolokacje charakterystyczne tylko dla siebie (sądzić kogoś po czymś, sądzić kogoś za coś, uważać na coś, uważać kogoś za kogoś). Aby uniknąć zbytniego rozdrobnienia takiej klasyfikacji proponuje się, aby jednostką opisu był nie leksem, lecz „leksem-sens” (odpowiednik zapożyczenia z języka angielskiego zbitki word-sense), czyli uporządkowana para: leksem i jego sens. Jako sens danego leksemu proponuje się jego wybrany odpowiednik w języku angielskim.

Przy takim założeniu istniałyby (co najmniej) cztery jednostki opisu dla leksemu *uważać* odpowiadające sensom: *think, mind, be careful, consider* oraz (co najmniej) trzy jednostki opisu dla leksemu *sądzić*: *think, judge try*. Jednostki (*uważać*: *think*), (*sądzić*: *think*) należałyby do tej samej klasy (oczywiście nie dlatego, że mają identyczny odpowiednik w j. angielskim, ale dlatego że tworzą takie same kolokacje).

Do jednej klasy proponowanej klasyfikacji należeć będą tylko te leksem-sensy, dla których kolokacje są identyczne zarówno w języku polskim i angielskim (wybór języka angielskiego jest tutaj arbitralny – spowodowany posiadaniem odpowiedniej bazy do analizy). Na przykład leksem-sensy *postawić*: *put*, i *zamienić*: *change* tworzą w języku polskim (trochę przypadkowo) takie same kolokacje (*coś na coś*), lecz w języku angielskim różnią się kolokacjami (odpowiednio: *something on something, something into something*) i dlatego w proponowanym podziale będą należeć do różnych klas.

Cele stworzenia takiej klasyfikacji są następujące:

- Zapewnienie jednorodności i pełności opisu kolokacji dla jednostek tego samego typu. W dotychczasowych opisach jednostki o tych samych kolokacjach były częstokroć opisywane na różne sposoby z powodu korzystania z ograniczonego korpusu oraz indywidualnych preferencji leksykografów.

- Ułatwienie pracy leksykograficznej – w celu opisania własności syntaktyczno-semantycznych hasła nie trzeba będzie badać korpusu lecz tylko umieścić hasło w odpowiedniej klasie (np. na podstawie analogii do haseł już opisanych).

- Umożliwienie tworzenia algorytmów przetwarzania języka w oparciu o klasy haseł – a nie pojedyncze hasła.

W obecnej stawia się następującą tezę: poszczególne klasy tak utworzonej klasyfikacji można scharakteryzować semantycznie. W tej pracy teza ta zostanie zweryfikowana na przykładzie przymiotników zaczerpniętych ze słownika systemu POLENG¹.

2. Kolokacje przymiotników

Można wyróżnić następujące kategorie gramatyczne, które mogą być podrzędne przymiotnikom, i jako takie, stać się przedmiotem obecnego opisu:

- frazy rzeczownikowe w określonym przypadku,
- frazy przyimkowe – z określonym przyimkiem i przypadkiem części rzeczownikowej,
- niektóre typy zdań podrzędnych.

3. Klasyfikacja przymiotników

Wydzielone i scharakteryzowane semantycznie zostały 23 klasy przymiotników o określonych kolokacjach w języku polskim i angielskim (niektóre z nich zawierają podklasy). Ponadto, wyszczególniono klasę przymiotników o kolokacjach specyficznych tylko dla siebie oraz klasę imiesłowów. Jako osobną (najbardziej liczną) klasę potraktowano pozostałe przymiotniki – czyli takie, które nie pełnią funkcji nadrzędnej w żadnych związkach zdaniowych. Dla każdej klasy (podklasy) podano kolokacje syntaktyczno-semantyczne należących do niej przymiotników w języku polskim i angielskim. Dla niektórych klas możliwych jest kilka różnych kolokacji. W celu algorytmicznego przetwarzania języka należy założyć, że w analizowanym zdaniu możliwe jest występowanie wszystkich przymiotników danej klasy we wszystkich kolokacjach.

¹ Słownik ten zawiera 15 000 leksemów występujących najczęściej w tekstach związanych z informatyką

Klasa 3. Przymiotniki wyrażające porównanie

Klasa 3.1. Stopień zgodności w relacji podobieństwa

kolokacje polskie	kolokacje angielskie
z czymś (kimś)	with something (someone)

dobry: well_matched, kompatybilny: compatible, porównywalny: comparable, niekompatybilny: incompatible, niezgodny: inconsistent, równoczesny: concurrent, sprzeczny: inconsistent, synonimiczny: synonymous, zbieżny: convergent, zgodny: compatible, zgodny: consistent,

Klasa 3.2. Stopień zgodności w relacji tożsamości

kolokacje polskie	kolokacje angielskie
z czymś [kimś]	to something [someone]

identyczny: identical, jednoznaczny: equivalent, niezgodny: contrary, równoprawny: equal, równy: equal, równoznaczny: tantamount, sprzeczny: contradictory, tożsamy: identical

Klasa 3.3. Porównania innego typu

kolokacje polskie	kolokacje angielskie
do [wobec, względem, w stosunku do] czegoś [kogoś]	to something [someone]

sporadycznie: z czymś (kimś)
analogiczny: analogous, odwrotny: contrary, podobny: similar, proporcjonalny: proportional, przeciwny: contrary, równoważny: equivalent, zbliżony: similar

Klasa 4. Przymiotniki wyrażające stopień niezbędności do wykonania jakiegoś działania

kolokacje polskie	kolokacje angielskie
1) na coś [do czegoś]	1) for something
2) na zrobienie czegoś [do zrobienia czegoś]	2) to do something
3) zrobić coś [by zrobić coś]	3) to do something
4) dla kogoś [komuś], na zrobienie czegoś, [do zrobienia czegoś]	4) for someone to do something
5) dla kogoś [komuś], by zrobić coś	5) for someone to do something

adekwatny: adequate, dogodny: convenient, fundamentalny: fundamental, idealny: perfect, konieczny: necessary, możliwy: possible, niedostateczny: insufficient, niemożliwy: impossible, nieodzowny: essential, nieprzydatny: useless, niewłaściwy: inappropriate,

niewystarczający: insufficient, niezastąpiony: irreplaceable, niezbędny: necessary, odpowiedni: fit, potrzebny: necessary, pożyteczny: useful, przydatny: useful, stosowny: appropriate, użyteczny: useful, właściwy: fit, zbędny: unnecessary, zbędny: unnecessary,

Klasa 5. Przymiotniki wyrażające gotowość lub kwalifikacje osoby do wykonania jakiegoś działania

kolokacje polskie

1) na coś [do czegoś]

2) na zrobienie czegoś

[do zrobienia czegoś]

3) zrobić coś [by zrobić coś]

kolokacje angielskie

1) for something

2) to do something

3) to do something

chętny: ready, gotowy: ready, kompetentny: qualified, nieodpowiedni: inadequate, niepowołany: not competent, niepowołany: unfit, nieupoważniony: unauthorized, nieuprawniony: unauthorized, niewłaściwy: unsuitable, skory: willing, skłonny: prone, uprawniony: entitled, zdeterminowany: determined, zobowiązany: obliged

Klasa 6. Przymiotniki wyrażające jakość w działaniu

kolokacje polskie

w czymś [z czegoś]

biegły: proficient, dobry: good, genialny: brilliant, kiepski: poor, nierównany: matchless, pomysłowy: ingenious, wybitny: outstanding

kolokacje angielskie

at something

Klasa 7. Przymiotniki wyrażające postawę lub skuteczność w działaniu

kolokacje polskie

w czymś

aktywny: active, konsekwentny: consistent, odczuwalny: noticeable, przodujący: leading, silny: robust, skuteczny: effective, staranny: careful, wpływowy: influential, wprawny: practised, wytrwały: persistent

kolokacje angielskie

in something

Klasa 8. Przymiotniki wyrażające możliwość ogarnięcia obiektu, zjawiska przy pomocy zmysłów lub rozumowania

kolokacje polskie

dla kogoś

akceptowalny: acceptable, bezbolesny: painless, czytelny: readable, jasny: clear, nieczytelny: illegible, niedostępny: inaccessible, nieprzewidywalny: unpredictable, niewidoczny: invisible, niezrozumiały: incomprehensible, strawny: acceptable, wykonalny: feasible

kolokacje angielskie

for someone

Klasa 9. Przymiotniki wyrażające wartość pewnego obiektu dla osoby

kolokacje polskie

kolokacje angielskie

dla kogoś

for someone

atrakcyjny: attractive, bezcenny: priceless, cenny: valuable, doniosły: critical, niebagatelny: substantial, niekorzystny: unprofitable, opłacalny: profitable, owocny: profitable, ważny: important

Klasy: 8. i 9. to jedyne dwie klasy o identycznych kolokacjach w obu językach – rozróżnienie między tymi klasami jest więc ściśle semantyczne.

Klasa 10. Przymiotniki wyrażające odróżnienie, separację

kolokacje polskie

kolokacje angielskie

od czegoś (kogoś) [niż coś (ktoś)]

from something (someone)

inny: different, oddzielny: separate, odseparowany: separate, odmienny: different

Klasa 11. Przymiotniki wyrażające oddalenie przestrzenne

kolokacje polskie

kolokacje angielskie

od czegoś o ...

from something (someone) by ...

daleki: far, oddalony: distant

Klasa 12. Przymiotniki opisujące zależność geometryczną

kolokacje polskie

kolokacje angielskie

do czegoś

to something (someone)

prostopadły: perpendicular, przeciwny: opposite, przyległy: adjacent, stychny: adjacent, zbieżny: convergent

Klasa 13. Przymiotniki wyrażające stopień odporności, wrażliwości

kolokacje polskie

kolokacje angielskie

na coś (rzadziej: kogoś)

to something (someone)

czuły: sensitive, głuchy: deaf, obojętny: insensible, odporny: resistant, odporny: immune, oporny: resistant, podatny: susceptible, uodporniony: immune, wrażliwy: sensitive, wytrzymały: resilient

Klasa 14. Przymiotniki wyrażające stopień przekonania do czegoś lub świadomości o czymś

kolokacje polskie

kolokacje angielskie

1) czegoś (kogoś)

1) of something (someone)

2) zdanie podrzędne „że”

2) zdanie podrzędne „that”

niepewny: uncertain, nieświadomy: unaware, pewny: sure, świadomy: aware

Klasa 15. Przymiotniki wyrażające stan będący konsekwencją jakiegoś czynu

kolokacje polskie

kolokacje angielskie

1) za coś

for something

2) sporadycznie: czegoś

of something

niewinny: innocent, odpowiedzialny: responsible, skazany: convicted,
wdzięczny: grateful, winny: guilty, zobowiązany: obliged

Klasa 16. Przymiotniki określające typowość obiektu dla pewnej klasy, którą obiekt reprezentuje

kolokacje polskie dla czegoś (kogoś)	kolokacje angielskie of something
--	---

charakterystyczny: characteristic, charakterystyczny: typical,
nietypowy: non-standard, nietypowy: untypical, reprezentatywny:
representative, specyficzny: specific, typowy: typical, zmienny:
characteristic

Klasa 17. Przymiotniki określające sławę

kolokacje polskie z czegoś (kogoś)	kolokacje angielskie for something
--	--

sławny: famous, słynny: famous, znany: well-known

Klasa 18. Przymiotniki określające stopień zasobności

kolokacje polskie w coś	kolokacje angielskie in something
-----------------------------------	---

bogaty: rich, obfity: rich, ubogi: poor

Klasa 19. Przymiotniki wyrażające celebrowanie czynności

kolokacje polskie przy czymś [w czymś]	kolokacje angielskie about something
--	--

ostrożny: careful, wybredny: choosy

Klasa 20. Przymiotniki wyrażające uzdolnienia

kolokacje polskie 1) do czegoś	kolokacje angielskie at something
--	---

uzdolniony: talented: utalentowany: gifted

Klasa 21. Przymiotniki opisujące nastawienie podczas oceniania

kolokacje polskie do [wobec, względem, w stosunku do] czegoś (kogoś)	kolokacje angielskie about something
--	--

obiektywny: objective, subiektywny: subjective, sceptyczny: sceptical

Klasa 22. Przymiotniki opisujące stopień zadowolenia

kolokacje polskie 1) czymś 2) zdanie podrzędne „że”	kolokacje angielskie with something zdanie podrzędne „that”
---	---

zadowolony: pleased, nieusatysfakcjonowany: unsatisfied,
niezadowolony: unsatisfied, usatysfakcjonowany: satisfied

Klasa 23. Przymiotniki wyrażające zasługiwanie na coś

kolokacje polskie	kolokacje angielskie
czegoś (kogoś)	of something (someone)
godny: worthy, niegodny: undeserving	

Klasa A. Inne mało liczne podklasy przymiotników

Klasa A zawiera przymiotniki, których własności syntaktyczne nie pasują do żadnej z wymienionych wyżej klas. Nie można wykluczyć, że analiza szerszego słownictwa pozwoliłaby na poszerzenie tych mało licznych grup a co za tym idzie, na ich semantyczną interpretację.

kolokacje polskie	kolokacje angielskie
-------------------	----------------------

A1.

od czegoś	on something
zależny: dependent	

A2.

od czegoś	from something
niezależny: independent	

A3.

na czymś	at something
obecny: present	

A4.

na czymś	from something
nieobecny: absent	

A5.

1) do czegoś	of something
2) zrobić coś [by zrobić coś]	of doing something
zdolny: capable	

A6.

czegoś	about something
ciekawyy: curious	

A7.

z czymś	of something
rozłączny: exclusive	

A8.

1) z czegoś	1) of something
2) zdanie podrzędne „że”	2) zdanie podrzędne „that”

PAGE

3) ze zrobienia czegoś

dumny: proud

3) to do something

A9.

w kimś

zakochany: enamoured

with someone

A10.

wśród [pośród] grupy

popularny: popular

with a group of people

Klasa B. Przymiotniki odczasownikowe (imiesłowy)

W leksykografii imiesłowy traktowane są niekonsekwentnie – niekiedy umieszczane są w słownikach jako oddzielne hasła, a niekiedy funkcjonują wyłącznie jako formy fleksyjne czasowników. Prawdopodobnie jednym z kryteriów decydujących o zaklasyfikowaniu imiesłowu przymiotnikowego jako oddzielnego hasła jest rola, jaką zwykle pełni w zdaniu (jeśli imiesłów ma tendencje do występowania w funkcji przydawki, to umieszczany jest jako osobny leksem, a jeśli w funkcji orzecznika, to traktowany jest jako forma fleksyjna czasownika)

Fakt występowania imiesłówów w słownikach jako osobnych leksemów – czyli traktowania ich tak samo jak przymiotników – wskazuje konieczność przeanalizowania ich własności syntaktyczno-semantycznych razem z innymi przymiotnikami. Z drugiej strony kolokacje imiesłówów są najczęściej takie same jak kolokacje ich czasowników bazowych (np. imiesłów *wtajemniczony*: *initiated* tworzy kolokacje *w coś*: *into something*, bo takie same kolokacje tworzy czasownik *wtajemniczyć*) – i podlegają zależnościom charakterystycznym dla czasowników, które nie są przedmiotem badań w tym miejscu.

Poniżej zajmiemy się tylko tymi kolokacjami imiesłówów, które nie dotyczą czasowników bazowych.

11. Imiesłowy wyrażające silne emocje

czymś

at something

przerażony: terrified, zdumiony: amazed, przejęty: excited

12. Imiesłowy wyrażające znużenie

czymś

of something

znudzony: bored, zmęczony: tired, znużony: weary

13. Imiesłowy wyrażające zatroskanie

czymś

about something

przejęty: worried, zatroskany: worried

I4. Inne imiesłowy**czymś [przez coś (kogoś)]****1) by something [someone]****Klasa N. Przymiotniki nie tworzące kolokacji z kategoriami podrzędnymi**

Tę najbardziej liczną klasę przymiotników – ze względu na jej zróżnicowanie – trudno jest zinterpretować semantycznie. Interpretacja syntaktyczna jest następująca: przymiotniki tej grupy nie przyłączają określonych kategorii gramatycznych (inaczej: nie mogą być elementami nadrzędnymi w żadnym związku w zdaniu).

Poniżej podano przymiotniki tej grupy (ze słownika systemu POLENG) zaczynające się w języku polskim na literę ‘a’:

abonamentowy: subscription, abonencki: subscriber, absolutny: absolute, abstrakcyjny: abstract, absurdalny: absurd, achromatyczny: achromatic, adaptacyjny: adaptive, addytywny: additive, administracyjny: administrative, adresowy: address, aerodynamiczny: aerodynamic, afrykański: African, akademicki: academic, akapitowy: paragraph, akcyjny: stock, aktualizacyjny: update, aktualny: current, aktualny: present, aktywny: active aktywny: on-line, akustyczny: acoustic, alarmowy: alarm, alfabetyczny: alphabetic, alfanumeryczny: alphanumeric, algebraiczny: algebraic, algorytmiczny: algorithmic, alkaliczny: alkaline, aluminiowy: aluminium, amatorski: amateurish, ambitny: ambitious, amerykański: American, analityczny: analytic, analogiczny: parallel, analogowo-cyfrowy: analog-digital, analogowy: analogue, anatomiczny: anatomical, angielski: English, angielskojęzyczny: English-language, anglojęzyczny: English-speaking, anglosaski: Anglo-Saxon, animacyjny: animation, animowany: animated, anomalny: anomalous, anonimowy: anonymous, antarktyczny: Antarctic, antenowy: aerial, antyczny: ancient, antyczny: antique, antymonopolowy: antitrust, antystatyczny: antistatic, antysymetryczny: antisymmetric antywirusowy: anti-virus, antywłamaniowy: security, apelacyjny: appellate, aplikacyjny: application, apostolski: apostolic, arabski: Arabic, arbitralny: arbitrary, archaiczny: archaic, archeologiczny: archeological, architektoniczny: architectural. archiwalny: archival, argentyński: Argentinian, armatni: cannon, artystyczny: artistic, arytmetyczny: arithmetic, ascetyczny: ascetic, asemblerowy: assembler, astronomiczny: astronomical, asymetryczny: asymmetric, asymptotyczny: asymptotic, asynchroniczny: asynchronous, atelierowy: atelier atmosferyczny: atmospheric, atomowy: nuclear, atomowy : atom, atramentowy: ink-jet, audiowizualny: audiovisual, australijski: Australian, autentyczny: authentic, autentyczny: true, automatyczny: automatic, autonomiczny: autonomous, autorski: authoring, autoryzowany: authorized, awangardowy: avant-guard.

Duża liczność tej klasy w porównaniu z innymi klasami wskazuje, że tylko pewna niewielka część przymiotników może w zdaniu pełnić rolę nadrzędną w stosunku do innych kategorii gramatycznych.

3. Wnioski

W pracy tej omówiono kolokacje przymiotników z kategoriami gramatycznymi w języku polskim i angielskim, w których przymiotniki pełnią rolę nadrzędną. Planowane jest utworzenie podobnego typu klasyfikacji dla innych części mowy (partykuło-przysłówek, spójników, przyimków, rzeczowników i czasowników). Wstępne badania wskazują, że klasyfikacje takie dla czasowników i rzeczowników będą bardziej złożone niż dla przymiotników. Wydaje się więc, że korzystne będzie przedstawienie takich klasyfikacji w postaci drzew. Pierwszym poziomem podziału dla czasowników może być na przykład semantyka podmiotu opisywanej czynności, drugim – kryterium przechodniości, a dopiero trzecim charakterystyka kolokacji (czyli w przypadku czasowników: dopełnień).

- [1] strona internetowa: <http://bobo.link.cs.cmu.edu> (opis gramatyki Link Grammar)
- [2] Polański K. (red.), *Słownik syntaktyczno-generatywny czasowników polskich*, Warszawa-Wrocław-Katowice-Gdańsk, 1980
- [3]. Jassem K. *Format wejścia w polsko-angielskim słowniku elektronicznym konstruowanym dla potrzeb tłumaczenia automatycznego* w: W. Jassem, C. Basztura, K. Jassem: *Speech and Language Technology, vol. 2*, str. 137 – 149, Poznań, Polskie Towarzystwo Fonetyczne, 1998
- [4] Łojasiewicz A. *Wasności składniowe polskich spójników*, Wydawnictwa Uniwersytetu Warszawskiego, Warszawa 1992