

Małgorzata Gębka-Wolak, Uniwersytet Mikołaja Kopernika (Polska, Toruń)  
Малгожата Гембка-Воляк, Университет им. Николоя Коперника  
(Польша, Торунь) mge@umk.pl

**Syntetyczny model stopniowania przymiotników we współczesnej  
polszczyźnie (uwagi na podstawie badania ilościowego)**

**The Synthetic Model of Comparison of Contemporary Polish Adjectives  
(Remarks Based on Quantitative Research)**

**Streszczenie**

W artykule omówiono przebieg i wyniki badań ilościowych dotyczących stopniowania przymiotników we współczesnej polszczyźnie. Badania przeprowadzono z myślą o zastosowaniu w nauczaniu języka polskiego jako obcego. Artykuł weryfikuje i uzupełnia wiedzę na temat wpływu czynników fonologicznych (liczba sylab) i morfologicznych (alternacje i morfemy słowotwórcze w temacie przymiotnika) na wybór syntetycznego modelu stopniowania. Otrzymane wyniki mogą być podstawą do formułowania uproszczonych reguł, które pomogą cudzoziemcom utworzyć taką formę stopnia wyższego, która będzie poprawna i statystycznie częstsza.

**Summary**

The paper describes the course and the results of a quantitative analysis concerning the comparison of adjectives in contemporary Polish. Our research was designed, taking into consideration the needs of teachers of Polish as a Foreign Language. The article verifies and completes the existing knowledge on the impact of phonological factors (i.e. the number of syllables) and morphological factors (i.e. the existence of alternations and word-forming morphemes in the adjective root) on the choice made by the speaker who decides to use a synthetic model of adjective comparison. Our results may serve as the basis for formulating simplified grammar rules, helping foreigners to choose between a synthetic and an analytic comparative form and allowing them to opt for a form which is correct and statistically more frequent.

słowa kluczowe: współczesny język polski; gramatyka opisowa; badanie ilościowe; stopniowanie; przymiotniki

keywords: contemporary Polish; descriptive grammar; quantitative research; comparison of adjectives

Код УДК 81-22

## 1. Wprowadzenie

Badania, których wyniki przedstawiono w artykule, zostały przeprowadzone z myślą o zastosowaniu w nauczaniu języka polskiego jako obcego. Ich zadaniem było sformułowanie wskazówek pomagających obcokrajowcom uczącym się języka polskiego typować przymiotniki, które powinni stopniować syntetycznie. Innymi słowy, chodziło o ułatwienie uczącym się polskiego dokonywania samodzielnego wyboru między wyrażeniami *grzeczniejszy* i *bardziej grzeczny*, *częstszy* i *bardziej częsty*, tj. wyboru bez posiłkowania się informacją pochodzącą ze słowników lub od rodzimych informatorów. Ze względu na założony cel dydaktyczny w badaniach przyjęto szereg ograniczeń i uproszczeń.

### 1.1. Przedmiot opisu

Przedmiotem opracowania jest zjawisko różnicowania ilości cech za pomocą operacji fleksyjnej, tj. z wykorzystaniem morfemów dodawanych do form przymiotnikowych, np. *stary*, *starszy*, *najstarszy*, *łatwy*, *łatwiejszy*, *najłatwiejszy*, bądź za pomocą operacji składniowej, polegającej na tworzeniu konstrukcji składającej się z leksemu wskazującego na większą lub mniejszą ilość cechy i bazowej formy przymiotnikowej, np. *bardziej siwy*, *najbardziej siwy*, *mniej zapracowany*, *najmniej zapracowany*. Pierwsza operacja tradycyjnie określana jest mianem stopniowania prostego (syntetycznego), druga – stopniowania opisowego (analitycznego)<sup>1</sup>. Należy przy tym dodać, że mimo że komponentami konstrukcji o funkcji gradacyjnej mogą być różne elementy leksykalne<sup>2</sup>, np. *niesamowicie ciekawy*, *okropnie ciekawy*, *niewiarygodnie ciekawy*, *piekielnie ciekawy*, *diabło ciekawy*, w opracowaniach przygotowywanych do celów dydaktycznych informacje o stopniowaniu opisowym koncentruje się wokół konstrukcji najbardziej regularnych, mianowicie tworzonych z wyrażeniami *bardziej*, *najbardziej*, *mniej*, *najmniej* (zob. np. Kaleta 1995: 246-250; Jędrzejko, Kita 2000: 158-160). Jako że i temu opracowaniu przyświeca cel dydaktyczny, zakres

---

<sup>1</sup> Przyjmujemy za Maciejem Grochowskim, że „stopniowanie jest operacją semantyczną, która polega na porównywaniu obiektów/stanów rzeczy, wykrywaniu różnic między nimi i komunikowaniu, w czym się te różnice ujawniają. Podstawą kategorii stopnia jest więc porównanie co najmniej dwóch obiektów lub stanów rzeczy. Dotyczy ono różnicy ilościowej między cechami obiektów lub między cechami stanów rzeczy i stanowi odpowiedź na pytanie, o ile dana cecha obiektu X (lub stanu rzeczy X) różni się od tej samej cechy innego, co najmniej jednego, obiektu Y (stanu rzeczy Y).” (Grochowski 2008: 61).

<sup>2</sup> Listę leksykalnych wykładników stopniowania synonimicznych względem *bardzo* przedstawia E. Janus (1977).

stopniowania opisowego ogranicza się do konstrukcji typu *bardziej zdrowy, najbardziej zdrowy*.

Perspektywa badawcza została zawężona do przymiotników, które podlegać mogą obu mechanizmom gradacji, czyli są stopniowane w sposób syntetyczny i analityczny. Tym samym, poza polem obserwacji znalazły się przymiotniki stopniowane tylko analitycznie. Będą to przede wszystkim genetyczne imiesłowy przymiotnikowe, np. *bardziej interesujący, bardziej wyspany, najbardziej skryty*, a także homonimy przymiotników tzw. relacyjnych, mające znaczenia przenośne, np. *bardziej bezpłciowy, bardziej drewniany, najbardziej polski*, zob. (1)-(3).

(1) *Po bezpłciowym ponurym urzędniku o nazwie, nomen omen, Bida, jeszcze bardziej bezpłciowy minister o nazwisku, nomen omen, Łopatka.* (NKJP)<sup>3</sup>

(2) *W miarę upływu czasu jego pogodny ton zaczął robić się coraz bardziej drewniany [...]* (NKJP)

(3) *Celtic – najbardziej polski zespół w Lidze Mistrzów – jechał do Kopenhagi pewny udziału w fazie pucharowej rozgrywek.* (NKJP)

Ponadto niewielką grupę przymiotników stopniuje się tylko opisowo ze względu na czynniki fonetyczne, a mianowicie możliwy po dodaniu morfemu stopnia wyższego zbieg spółgłosek tych samych, por. *gorzki* i *\*gorzszy*, lub mało fonetycznie zróżnicowanych, por. *śliski* i *\*ślisszy*, *łasy* i *\*łasszy*<sup>4</sup>. Stopniowanie opisowe niektórych przymiotników, jak np. *cierpki, sypki, szorstki*, jest motywowane także czynnikami zwyczajowymi (zob. Buttler, Kurkowska, Satkiewicz 1986: 256).

## 1.2. Założenia i cel badań

Ze względu na planowane zastosowanie wyników badań i stosowaną metodę konieczne jest przyjęcie kilku założeń opisu.

Uznajemy, że każdy przymiotnik, który można stopniować syntetycznie, podlega także stopniowaniu analitycznemu. Nie ma bowiem systemowych ograniczeń w tworzeniu konstrukcji *bardziej głośny, bardziej nudny, bardziej ciekawy, bardziej ciężki, bardziej dokładny, bardziej długi, bardziej miły*. W praktyce zakres użycia modelu analitycznego ogranicza norma językowa,

<sup>3</sup> Przykłady opatrzone skrótem NKJP pochodzą z Narodowego Korpusu Języka Polskiego, udostępnionego na stronie internetowej <http://www.nkjp.pl/>. Zawartość i możliwości wyszukiwawcze korpusu są opisane w pracy Przepiórkowski et al. 2012.

<sup>4</sup> Wymienione ograniczenia stopniowania syntetycznego są opisane w literaturze, zob. np. Dick 1976, Klebanowska, Kochański, Markowski 1985: 134, Buttler, Kurkowska, Satkiewicz 1986: 254-255, Jadacka 2005: 79, Marcisz 2009.

skodyfikowana, tj. zapisana w kompendiach poprawnościowych, lub naturalna, czyli odpowiadająca poczuciu językowemu użytkowników języka<sup>5</sup>. Tym samym, jeśli mówimy, że dany przymiotnik stopniuje się tylko / przeważnie syntetycznie, mamy na uwadze ograniczenia normatywne, a nie systemowe.

Należy przy tym podkreślić, że w kompendiach poprawnościowych – mimo odnotowywania upowszechniania się modelu opisowego – pierwszeństwo przyznaje się modelowi prostemu, czego wyrazem jest zalecenie posługiwania się formami syntetycznymi „w tych wszystkich wypadkach, w których ich tworzenie jest możliwe” (Buttler, Kurkowska, Satkiewicz 1986: 257), gdy „można korzystać z tradycyjnych, regularnych form stopniowania syntetycznego” (Jadacka 2005: 80). W istocie jednak norma wyboru modelu stopniowania jest nieco bardziej skomplikowana. Na podstawie komentarzy normatywnych zawartych w NSPP, podstawowym źródle informacji o normie skodyfikowanej, zbiór przymiotników stopniowalnych można podzielić na 3 podzbiory: 1) przymiotniki, które zaleca się stopniować tylko syntetycznie (opatrzone komentarzem „nie: *bardziej* + forma przymiotnika”), np. *głośny, głęboki, bliski, brzydki, dobry, korzystny, obfity, potworny, trwały*, 2) przymiotniki, dla których dopuszcza się dwa modele stopniowania, z których pierwszeństwo przyznaje się modelowi syntetycznemu (opatrzone komentarzem „forma syntetyczna *a. bardziej* + forma przymiotnikowa”), np. *bolesny, charakterystyczny, dynamiczny, gwałtowny, negatywny, serdeczny, zieleńszy*, 3) przymiotniki, dla których dopuszcza się dwa modele stopniowania, z których pierwszeństwo przyznaje się modelowi analitycznemu (opatrzone komentarzem „*bardziej* + forma przymiotnikowa *a. forma syntetyczna*”), np. *agresywny, dyspozycyjny, jednoznaczny, sumienny, wyborny*. Dodajmy, że podzbiory pierwszy i drugi są w NSPP zdecydowanie liczniejsze niż trzeci, co zgadza się z ogólnym zaleceniem posługiwania się przede wszystkim stopniowaniem syntetycznym<sup>6</sup>.

Jak pokazuje przeprowadzona przez autorkę tekstu analiza danych z NKJP, w odniesieniu do podgrupy pierwszej intuicja rodzimych użytkowników języka na ogół pokrywa się z normą skodyfikowaną w NSPP.

---

<sup>5</sup> Podział na normę naturalną i skodyfikowaną zaproponowała Danuta Buttler w artykule *Zróżnicowanie współczesnej normy językowej* (podaję za: Markowski 2005: 30).

<sup>6</sup> Jak podaje A. Marcisz (2009: 297), spośród 427 przymiotników odnotowanych w NSPP 140 podlega stopniowaniu tylko analitycznemu. Z kolei zalecenie posługiwania się stopniowaniem prostym jako jedynym lub podstawowym sposobem intensyfikacji cechy pojawia się przy 282 przymiotnikach, co stanowi 66,43% całego zbioru przymiotników.

Więcej rozbieżności między oceną normatywną a użyciem w tekstach notujemy w odniesieniu do podzbioru drugiego<sup>7</sup>. Jest jasne, że intuicyjny wybór, który rodzimemu użytkownikowi języka zasadniczo nie sprawia trudności, dla obcokrajowca bywa kłopotliwy. Obcokrajowca trzeba bowiem nauczyć – w programie maksymalnym – które przymiotniki musi / powinien stopniować syntetycznie, które analitycznie, a dla których równie dobre są oba sposoby stopniowania. Zadaniem minimum jest nauczenie, które przymiotniki trzeba stopniować syntetycznie. Najmniej efektywnym rozwiązaniem jest podanie odpowiednich list przymiotników. Zdecydowanie korzystniejsze byłoby odwołanie się do reguł czy chociażby wskazanie prawidłowości związanych z wyborem wzorca stopniowania. Tymczasem nie tylko w materiałach do nauczania jpjo, ale także w gramatykach języka polskiego o zasadach wyboru sposobu stopniowania mówi się niewiele (zob. Dick 1976, Bartnicka, Rybicka 1984). Stąd zrodził się pomysł, by przeprowadzić badania, które uzupełniłyby wiedzę z tego zakresu.

Celem badań było zweryfikowanie hipotezy o istnieniu korelacji między wybranymi cechami przymiotników, stanowiącymi w badaniu zmienne niezależne, a sposobem stopniowania danego leksemu przymiotnikowego, stanowiącym zmienną zależną. Wykorzystane zostały dane ilościowe, a w ich analizie posłużono się programem komputerowym IBM SPSS Statistics 22 (PS IMAGO), pozwalającym na ich obróbkę statystyczną i wyciąganie wniosków nt. tego, które zmienne należy uznać za relewantne.

Ze względu na ograniczoną objętość artykułu dokładne omówienie przeprowadzonych badań nie jest możliwe. Ograniczymy się więc do krótkiego przedstawienia etapów prac, a następnie zaprezentujemy wyniki dotyczące roli wybranych czynników formalnych.

## **2. Wstępny etap badań – ustalenia, trudności, wyniki**

Badanie zostało podzielone na 3 etapy: 1) ustalenie wyjściowej listy przymiotników objętych obserwacją, 2) ustalenie normy dotyczącej sposobu stopniowania przymiotników wytypowanych na pierwszym etapie prac i przygotowanie ostatecznej listy przymiotników wchodzących w skład próby badawczej, 3) weryfikacja hipotezy o istnieniu korelacji między określonymi zmiennymi a typem stopniowania.

Każdy z etapów był przeprowadzany z wykorzystaniem innych narzędzi badawczych.

---

<sup>7</sup> Istnienie rozbieżności między normą skodyfikowaną a uzusem pokazują także wyniki ankiety przeprowadzonej przez A. Marcisz (2009).

## 2.1. Zestawienie wyjściowej listy przymiotników objętych obserwacją

Informację o tym, które przymiotniki są stopniowalne, zawiera kilka słowników języka polskiego, np. ISJP, USJP, NSPP. O wyborze źródła materiału zdecydowało przede wszystkim kryterium techniczne, tj. łatwość automatycznego wyszukania materiału do badań i przeniesienia do przygotowywanej w programie EXCEL bazy danych. Najwygodniejszym źródłem materiału okazała się internetowa wersja *Słownika gramatycznego języka polskiego* (SGJP), opracowywanego przez Zygmunta Saloniego, Marcina Wolińskiego, Roberta Wołosza, Włodzimierza Gruszczyńskiego i Danutę Skowrońską<sup>8</sup>.

Autorzy SGJP syntetyczny stopień wyższy przymiotników traktują jako odrębny leksem, stanowiący osobne hasło w słowniku. Leksemy te daje się z kolei automatycznie wyszukać za pomocą wbudowanego do internetowej wersji SGJP narzędzia filtrowania materiału, a otrzymaną listę można skopiować. Automatycznie wyszukany materiał zawierał 658 przymiotników stopniowalnych syntetycznie. Trzeba jednak pamiętać, że część wybranych przymiotników odnosi się do znacznie większej liczby jednostek leksykalnych, gdyż przeprowadzone badania nie uwzględniły kwestii homonimii. Przykładowo, wyrażenie *świeży* na liście materiału wybranego do badania zanotowano tylko raz, mimo że w ISJP (2000: 2, 796) w hasle *świeży* opisano aż 13 jednostek leksykalnych fundowanych na tym kształcie, np. 1) 'to, co pojawiło się lub zostało zrobione niedawno, czasem zastępując coś starego', np. *świeży śnieg*, 2) 'takie, które przygotowano niedawno i które są teraz najlepsze, a później mogą być niesmaczne lub zepsute', np. *świeże jedzenie*, 3) 'świeże owoce, warzywa, mięso, ryby itp. pochodzą prosto od producenta lub plantatora i nie były gotowane, mrożone, wekowane ani w inny sposób konserwowane' itd. Z punktu widzenia prowadzonych badań ważne jest to, że homonimiczne jednostki notowane w poszczególnych hasłach mimo różnic semantycznych lub pragmatycznych nie tracą cechy stopniowalności. Tak więc nieuwzględnienie homonimii niewątpliwie upraszcza opis, ale istotnie nie wpływa na otrzymane wyniki.

Przyjęto ponadto założenie, że jeśli dla danego przymiotnika SGJP notuje formę stopnia wyższego, to przymiotnik ten tworzy też regularnie (formalnie i semantycznie) formę stopnia najwyższego.

---

<sup>8</sup> Korzystano z wersji roboczej, w czasie przeprowadzania badania dostępnej na stronie internetowej <http://sgjp.pl/kuznia/sownik/>. W momencie pisania artykułu internetowa wersja *Słownika gramatycznego języka polskiego* jest już gotowa i można z niej korzystać za pośrednictwem strony <http://sgjp.pl/>.

## 2.2. Ustalenie normy dotyczącej sposobu stopniowania wytypowanych przymiotników

Jak wspomniano we Wprowadzeniu, nie ma systemowych ograniczeń w analitycznym tworzeniu stopnia, zakres użycia stopniowania analitycznego ogranicza jednak norma. Drugi etap prac polegał więc na zebraniu informacji normatywnych. Jeśli chodzi o źródło normy, ze względu na to, że kształt kodyfikacji zaproponowany w NSPP w interesującym nas zakresie stopniowania budzi dyskusje (zob. Maciak 2009), zdecydowano, że materiał referencyjny będzie stanowił Narodowy Korpus Języka Polskiego. Tym samym omawiane badanie uwzględnia nie normę skodyfikowaną (postulowaną), ale naturalną, a zasadniczo szerzej – uzus.

Do wyszukania przykładów użycia form stopnia przymiotników zgromadzonych w bazie przygotowanej na pierwszym etapie prac zastosowana została wyszukiwarka korpusowa IPI PAN-u<sup>9</sup>. Przeszukiwano podkorpus zrównoważony, obejmujący 300 milionów segmentów. Na tej podstawie zebrane zostały dane liczbowe na temat użycia stopniowania syntetycznego i analitycznego. Ustalono frekwencję w korpusie syntetycznych i analitycznych form stopnia wszystkich wytypowanych do badania przymiotników, tj. 658 wyrażen. Dodatkowo obliczono: 1) łączną liczbę form stopnia wyższego odnotowanych w korpusie oraz 2) procentowy udział form stopnia syntetycznego.

Wnioski z obserwacji danych z NKJP zostały wykorzystane do ograniczenia materiału poddanego dalszym badaniom oraz do jego podziału.

Ograniczenie materiału wynikało z faktu, że w NKJP poświadczono zostało użycie form stopnia wyższego nie wszystkich badanych jednostek, a dla części jednostek odnalezione zostały tylko pojedyncze przykłady użycia. W NKJP nie znaleziono ani jednego przykładu użycia form stopnia wyższego 60 przymiotników, przypomnijmy, mimo że systemowo możliwy jest dla nich zarówno model syntetyczny, jak i analityczny. Liczba przymiotników, dla których znaleziono nie więcej niż 10 przykładów użycia form stopnia wyższego, łącznie syntetycznego i analitycznego, w próbie 300 milionów segmentów wyniosła 198. Jako że opracowanie ma służyć uczącym się języka polskiego jako obcego, te dwie podgrupy przymiotników, obejmujące 258 wyrażen, zostały wyłączone z próby będącej podstawą dalszych analiz. Poza polem obserwacji znalazły się takie wyrażenia, jak *bajeczniejszy*, *częstotliwszy*, *dbalszy*, *drętwiejszy*, *kapitalniejszy*, *kurczliwszy*, *lesistszy*,

---

<sup>9</sup> Wyszukiwanie odbywa się za pośrednictwem strony: <http://nkjp.pl/poliqarp/>.

*muzykalniejszy, nieomylniejszy, ofiarniejszy, okropniejszy, rumieńszy, sekretniejszego, tendencyjniejszego.* Liczebność badanej próby po wyłączeniach wyniosła 400 przymiotników.

### **2.3. Wyodrębnienie podklas przymiotników ze względu na kryterium procentowego użycia syntetycznych form stopnia wyższego**

Ważnym, aczkolwiek spodziewanym wnioskiem z badań korpusowych było to, że udział syntetycznych form przymiotnikowych wśród form stopnia wyższego jest mocno zróżnicowany i waha się od 100% dla przymiotników typu *głupi, brzydki, chudy, mały, niski, drogi, wczesny*, do 0% dla przymiotników typu *odważny, oporny, dziecinny, typowy, osobisty, pomysłowy*. To zróżnicowanie użycia form syntetycznych stało się podstawą do wyróżnienia w badanym materiale podklas wyrażań. Należy wyjaśnić, że przeprowadzony podział był niezbędny, by przeprowadzić zaplanowane na następnym etapie prac badania statystyczne z uwzględnieniem zmiennych niezależnych.

Podstawą wyróżniania podklas przymiotników było kryterium procentowego udziału w materiale NKJP syntetycznych form stopnia wyższego. Zasadniczy problem polegał na tym, ile podklas wyróżnić i o jakim zakresie. Żeby pokazać potwierdzone w materiale płynne przejście od przymiotników, dla których w materiale potwierdzono model stopniowania tylko syntetyczny, do przymiotników, dla których w materiale potwierdzono stopniowanie tylko wg modelu analitycznego, należałoby wyróżnić kilkanaście podklas. Mogłoby to być np. 11 podklas, o zmniejszającym się o 10 % zakresie użycia danego modelu stopniowania, tj. przedział 1. =100% użycie modelu syntetycznego, przedział 2. < 100% i ≥ 90%, przedział 3. < 90% i ≥ 80%, przedział 4. < 80% i ≥ 70%, przedział 5. < 70% i ≥ 60%, przedział 6. < 60% i ≥ 50%, przedział 7. < 50% i ≥ 40%, przedział 8. < 40% i ≥ 30%, przedział 9. < 30% i ≥ 20%, przedział 10. < 20% i ≥ 10%, przedział 11. < 10% i ≥ 0%. Nie ma jednak merytorycznego uzasadnienia dla zastosowania podziału wg kryterium 10-procentowego czy zresztą innego, co 5% czy co 20%. Niewątpliwie lepszy, bo uzasadniony logicznie, byłby podział dychotomiczny na 2 wzajemnie rozłączne podklasy, tj. podklasę 1.: przymiotniki, których 100% użycie notowanych w NKJP to użycia syntetyczne, i podklasę 2: przymiotniki, dla których nie odnotowano 100-procentowego użycia modelu syntetycznego. Z punktu widzenia zastosowania wyników badań w nauczaniu jppo podział poprawny logicznie ma jednak istotne ograniczenia. W klasie 1. znalazło się stosunkowo niewiele przymiotników: z 400 wyrażań jedynie 23 (np. *mały, niski, drogi, głupi,*



wczesny, późny, świetny, cienki, głodny, brzydki, chudy, prędko). Z kolei klasa 2. objęła nie tylko dużo materiału, ale także mocno zróżnicowanego. Obok przymiotników, dla których poświadczono w korpusie użycie modelu syntetycznego bliskie 100% (np. 99,9% dla przymiotników *wysoki, duży, szybki, krótki, ładny, lekki, gruby, łatwy*, 99,8% dla przymiotników *tani, dawny, dobry*, 99,7% dla przymiotników *ubogi, długi, częsty*, 99,6% dla przymiotników *drobny, słaby, mocny*), w klasie 2. znalazłyby się przymiotniki, dla których w NKJP poświadczono użycie jedynie modelu analitycznego (np. *typowy, charakterystyczny, pilny, obiektywny, obszerny, energiczny, odważny, odporny, chaotyczny, lojalny, pomysłowy, zasadny*). W tej sytuacji podjęta została decyzja o odstąpieniu od podziału logicznego i arbitralnym wyznaczeniu granicy między klasami pierwszą a drugą, zob. tab. 1.

**Tabela nr 1. Podklasy przymiotników ze względu na kryterium procentowego użycia syntetycznych form stopnia wyższego**

kategorie ze względu na % form syntetycznych					
		Częstość	Procent	Procent ważnych	Procent skumulowany
Ważne	100%-85%	150	37,5	37,5	37,5
	poniżej 85%	250	62,5	62,5	100,0
	Ogółem	400	100,0	100,0	

Podklasa pierwsza objęła przymiotniki, dla których stopniowanie syntetyczne jest jedynym lub dominującym modelem stopniowania (przedział od 100% do 85%), a więc polecanym obcokrajowcom do stosowania. Podklasa druga (przedział poniżej 85%) obejmuje natomiast przymiotniki, dla których model analityczny jest modelem, który z powodzeniem można polecać obcokrajowcom do stosowania.

### 3. Badanie korelacji między wybranymi zmiennymi a typem stopniowania

Następny etap badań polegał na weryfikacji hipotezy o istnieniu korelacji między przynależnością przymiotnika do podklasy 1. lub 2. a przyjętymi zmiennymi niezależnymi, ustalonymi na podstawie literatury i intuicji językowej autorki badań. Pod uwagę wzięto następujące czynniki:

- 1) alternacja w temacie formy stopnia wyższego (zob. Dick 1976),
- 2) obecność przyrostka słowotwórczego w temacie (zob. Dick 1976),
- 3) liczba sylab formy przymiotnikowej stopnia równego w mianowniku,
- 4) frekwencja przymiotnika, od którego tworzymy formę stopnia,

5) cechy semantyczne przymiotnika, takie jak: a) nazywanie cechy fizycznej (wielkość, ilość, kształt, wrażenia zmysłowe, relacje czasowe i przestrzenne) np. *mały, gładki, wysoki*), b) nazywanie cechy abstrakcyjnej przypisywanej głównie obiektom, nie ludziom, np. *abstrakcyjny, realny, ekonomiczny*, c) nazywanie cech przypisywanych przede wszystkim człowiekowi (charakter, stany fizyczne i umysłowe), np. *spokojny, życzliwy, mądry*, d) cechy zawierające element oceny, np. *głupi, pomyślny, słuszny*.

Pierwsze badania, podczas których testowano wagę poszczególnych czynników, doprowadziły do wniosku, że ostatni czynnik – cechy semantyczne przymiotnika – zasadniczo nie jest relewantny. Inaczej rzecz się ma z frekwencją przymiotnika, która okazała się czynnikiem skorelowanym z typem stopniowania. Przymiotniki często używane w polszczyźnie mają bowiem tendencję do stopniowania syntetycznego. Rodzimy użytkownik języka jest w stanie wypowiedzieć się na temat przybliżonej frekwencji wyrażeń z leksykonu i orzec, że przymiotniki typu *pytki, ubogi, grzeczny* mają niższą frekwencję niż *niski, szybki, krótki*. Takiej intuicji nie ma natomiast obcokrajowiec, stąd reguły typu *stopniuj syntetycznie te przymiotniki, które mają w polszczyźnie wysoką frekwencję*, aczkolwiek zgodne z wynikami badań, dla uczących się języka polskiego nie będą przydatne. Tym samym w dalszych badaniach wzięto pod uwagę rolę trzech pierwszych z wyżej podanych cech przymiotników.

### 3.1. Podział badanej próby ze względu na uwzględnione zmienne niezależne

Kryterium alternacji w temacie formy stopnia wyższego stało się podstawą do wyróżnienia 2 podklas przymiotników, jednej z alternacjami, do której należą przymiotniki typu *tęgi - tęższy, rzadki - rzadszy, daleki - dalszy, silny - silniejszy*, drugiej bez alternacji, np. *prosty - prostszy, ciekawy - ciekawszy, żywy - żywszy*. W badanej próbie dominuje typ pierwszy, co pokazują dane zawarte w tabeli 2.

**Tabela nr 2. Podział badanej próby ze względu na alternacje w temacie**

		Alternacje w temacie			
		Częstość	Procent	Procent ważnych	Procent skumulowany
Ważne	brak alternacji w temacie	77	19,3	19,3	19,3
	alternacja w temacie	323	80,8	80,8	100,0
	Ogółem	400	100,0	100,0	

Na korelację między obecnością w temacie bazowego przymiotnika morfemu słowotwórczego a wyborem modelu stopniowania zwrócił uwagę Dick (1976). Zdaniem badacza brak sufiksu słowotwórczego sprzyja stopniowaniu syntetycznemu, np. *blady, cichy, dobry, długi, duży, gruby, mały, miły, nowy, słaby, stary, zły*. Zdecydowana dominacja stopniowania syntetycznego jest ponadto związana z obecnością w temacie sufiksów -k, -ek, -ok, np. *bliski, niski, daleki, głęboki, szeroki*. Zasada ta nie działa jednak bezwzględnie, por. *obcy, szary*, stopniowane opisowo.

W omawianym badaniu kryterium obecności przyrostka słowotwórczego w temacie stało się podstawą do wyróżnienia 4 podklas przymiotników: 1) bez przyrostka, np. *cichy, twardy*, 2) z przyrostkami -k-, -ek-, -ok-, np. *bliski, daleki, szeroki*, 3) z przyrostkiem -n-, -ń-, np. *celny, pogodny, dostatni*, 4) z innym przyrostkiem, np. *mglisty, postępowy, dokuczliwy, formalny*, zob. tab. 3.

**Tabela 3. Podział badanej próby ze względu na obecność i typ przyrostka tematycznego**

		Przyrostek			
		Częstość	Procent	Procent ważnych	Procent skumulowany
Ważne	bez przyrostka słowotwórczego	109	27,3	27,3	27,3
	k/-ek/-ok	20	5,0	5,0	32,3
	-n/-ń	176	44,0	44,0	76,3
	inny przyrostek	95	23,8	23,8	100,0
	Ogółem	400	100,0	100,0	

Jeśli chodzi o liczbę sylab formy stopnia równego w mianowniku, to w badanej próbie znalazły się przymiotniki, które miały od 1 do 6 sylab, przy czym najczęściej zanotowano przymiotników dwu- i trzysylabowych, zob. tab. 4.

**Tabela 4. Podział badanej próby ze względu na liczbę sylab formy mianownikowej**

Liczba sylab

		Częstość	Procent	Procent ważnych	Procent skumulowany
Ważne	1	1	,3	,3	,3
	2	162	40,5	40,5	40,8
	3	161	40,3	40,3	81,0
	4	67	16,8	16,8	97,8
	5	7	1,8	1,8	99,5
	6	2	,5	,5	100,0
	Ogółem	400	100,0	100,0	

### 3.2. Wyniki klasyfikacji przeprowadzonej metodą drzewa decyzyjnego

Ostatni etap analizy miał na celu weryfikację hipotezy o istnieniu związku między wymienionymi w 3.1. zmiennymi niezależnymi a zmienną zależną, odnoszącą się do klasy przymiotnika ustalonej ze względu na kryterium procentowego udziału w NKJP syntetycznych form stopnia wyższego (zob. p. 2.3.). Innymi słowy, chodziło o zbadanie, jakie czynniki sprzyjają wyborowi stopniowania syntetycznego (podklasa 1), a jakie mają związek ze stopniowaniem opisowym (podklasa 2). Badanie zostało przeprowadzone za pomocą narzędzi wbudowanych w program komputerowy IBM SPSS STATISTIC 22. Zastosowano klasyfikację metodą drzewa decyzyjnego. Procedura ta polega na kojarzeniu obiektów na podstawie ich charakterystycznych cech (zmiennych opisujących) z odpowiednimi wartościami pewnej jakościowej zmiennej celu, którą nazywamy kategorią lub klasą. Metoda jest stosowana do kojarzenia obiektów na podstawie ich charakterystycznych cech, zmienne nie są ilościowe (nominalne, porządkowe), a z takimi mamy tu do czynienia. Wykorzystano jeden z popularniejszych algorytmów budowania drzewa, tj. algorytm CRT. Wyróżnia się on tym, że budowane drzewa są binarne, a w wyniku klasyfikacji otrzymujemy reguły decyzyjne o postaci *jeżeli..., to ...* Parametry zastosowane w modelu prezentuje poniżej tabela 5.

**Tabela 5. Podsumowanie modelu****Model - podsumowanie**

Specyfikacja	Algorytm budowy drzewa	CRT
	Zmienna zależna	kategorie ze względu na % form syntetycznych
	Zmienne niezależne	Liczba sylab, Przyrostek, Alternacje w temacie
	Walidacja	Losowy podział zbioru
	Maksymalna głębokość drzewa	5
	Minimalna ilość obserwacji w węźle nadrzędnym	20
	Minimalna ilość obserwacji w węźle podrzędnym	5
Wyniki	Uwzględnione zmienne niezależne	Liczba sylab, Przyrostek
	Liczba węzłów	3
	Liczba węzłów końcowych	2
	Głębokość	1

Należy podkreślić, że spośród trzech wziętych pod uwagę zmiennych, tj. liczby sylab, obecności i typu przyrostka i alternacji w temacie, tylko dwie zostały wzięte pod uwagę podczas klasyfikacji. Pomięta – jako mniej istotna dla podziału – została zmienna alternacje. Waga zmiennych nie jest jednakowa, co pokazuje tabela 6. Liczba sylab jest cechą donioślejszą niż typ przyrostka.

**Tabela 6. Ważność zmiennych niezależnych****Ważność zmiennej niezależnej**

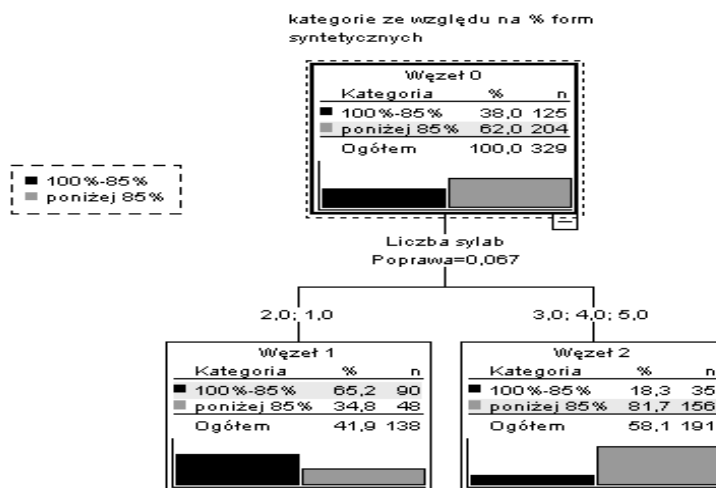
Zmienna niezależna	Ważność	Znormalizowana ważność
Liczba sylab	,067	100,0%
Przyrostek	,055	82,0%

Algorytm budowy drzewa: CRT

Zmienna zależna: kategorie ze względu na % form syntetycznych

W wyniku klasyfikacji otrzymujemy drzewo decyzyjne, zob. schemat 1., i reguły.

### Schemat 1. Drzewo decyzyjne



Węzły drzewa nie są czyste, a to oznacza, że zastosowane kryteria nie stanowią podstawy podziałów binarnych. Pokazują więc pewną tendencję, na której podstawie można sformułować jedynie wskazówkę, która w programie standardowo jest nazywana regułą. W wyniku klasyfikacji otrzymaliśmy dwie reguły decyzyjne, zob. Reguły 1, i 2.

**Reguła 1.** Odnosi się do węzła pierwszego.

IF (((Liczba sylab = 2 OR Liczba sylab = 1) OR (Liczba sylab != 3 AND Liczba sylab != 4 AND Liczba sylab != 5) AND (Przyrostek = "bez przyrostka słowotwórczego" OR Przyrostek = "k/-ek/-ok")))

THEN Node = 1 Prediction = 1 Probability = 0.600000

Jeżeli liczba sylab równa się 2 lub nie równa się 3 i nie równa się 3,4, 5 i kategoria przyrostek ma wartość 'bez przyrostka' lub '-k/-ek/-ok', to przymiotnik należy do kategorii pierwszej, tj. z wyraźną przewagą w tekstach współczesnej polszczyzny form syntetycznych. Takie przymiotniki zaleca się stopniować syntetycznie. Zwróćmy uwagę, że prawdopodobieństwo, że przymiotnik będzie miał formę stopnia syntetycznego wynosi 60%.

## Reguła 2. Odnosi się do węzła 2

IF (((Liczba sylab = 3 OR Liczba sylab = 4 OR Liczba sylab = 5) OR (Liczba sylab != 2 AND Liczba sylab != 1) AND (Przyrostek != "bez przyrostka słowotwórczego" AND Przyrostek != "k/-ek/-ok")))

THEN Node = 2 Prediction = 2 Probability = 0.782609

Jeżeli liczba sylab równa się 3 lub 4, 5 lub liczba sylab nie równa się 2, 1 i kategoria przyrostek nie ma wartości 'bez przyrostka' i '-k/-ek/-ok', wtedy z prawdopodobieństwem 78,26% mamy do czynienia z przymiotnikiem należącym do kategorii drugiej, to jest takiej, dla którego wzrasta procent użycia form analitycznych.

Program dostarcza nam też informacji na temat jakości naszego modelu. Świadczą o tym dane zawarte w dwóch tabelach, zob. Tab. 7. i 8.

### Tabela 7. Ocena ryzyka

Ryzyko	Ocena	Błąd standardowy
Próba		
Ucząca	,282	,053
Testowa	,252	,024

Algorytm budowy drzewa: CRT

Zmienna zależna: kategorie ze względu na % form syntetycznych

Ocena ryzyka dla próby testowej (omawianej) wynosi 0,252, co oznacza, że 25,2% przykładów w próbie testowej zostało źle zakwalifikowanych.

Macierz klasyfikacji pokazuje z kolei procent poprawnych kwalifikacji.

### Tabela 8. Macierz klasyfikacji

Macierz klasyfikacji		Przewidywana		Procent poprawnych
Próba	Obserwowana	100%-85%	poniżej 85%	
Ucząca	100%-85%	15	10	60,0%
	poniżej 85%	10	36	78,3%
	Procent ogółem	35,2%	64,8%	71,8%
Testowa	100%-85%	90	35	72,0%
	poniżej 85%	48	156	76,5%
	Procent ogółem	41,9%	58,1%	74,8%

Algorytm budowy drzewa: CRT

Zmienna zależna: kategorie ze względu na % form syntetycznych

Co ciekawe, procent poprawnych kwalifikacji jest nieco wyższy dla drugiej podklasy przymiotników, tj. tych, dla których to model analityczny statystycznie staje się donioślejszy i jako taki może być polecany obcokrajowcom do stosowania.

#### 4. Wnioski

Przeprowadzone badania zweryfikowały tezę Dicka (1976) o roli czynnika morfologicznego w wyborze modelu stopniowania przymiotników. Okazało się, że czynnik ten jest mniej istotny, niż sądził Dick. Potwierdzono natomiast, że w języku polskim przy wyborze modelu stopniowania ważny jest czynnik fonologiczny, mianowicie liczba sylab bazowego przymiotnika. Należy dodać, że zależność zjawisk morfologicznych od fonologicznych, która jest podstawową tezą kierunku badawczego określanego mianem morfologii prozodycznej (zob. Szpyra-Kozłowska 2000: 60)<sup>10</sup>, nie jest w językach czymś rzadkim. Potwierdzona dla języka polskiego korelacja liczby sylab z modelem stopniowania, która jest czynnikiem ważniejszym niż budowa morfologiczna, na tle innych języków nie jest więc niczym wyjątkowym.

Otrzymane wyniki uprawniają do sformułowania 2 uproszczonych reguł pomagających cudzoziemcom utworzyć taką formę stopnia wyższego, która będzie poprawna i statystycznie częstsza, niezależnie od tego, czy istnieje druga możliwość:

1) jeśli przymiotnik ma 1 sylabę lub 2 sylaby, stopnij go syntetycznie, np. *slaby – slabszy, młody – młodszy, cichy – cichszy, biały – bielszy*;

2) jeśli przymiotnik ma 3 sylaby lub więcej sylab, to stopnij go analitycznie, np. *uparty – bardziej uparty, różowy – bardziej różowy, nieszczęśliwy – bardziej nieszczęśliwy, niespokojny – bardziej niespokojny, wartościowy – bardziej wartościowy*.

Badanie ujawniło szereg trudności, z którymi trzeba się zmierzyć podczas analizy materiału językowego metodami ilościowymi, m.in. związane z doбором próby i zmiennych niezależnych, a także z ryzykiem błędu, jakim obarczone są otrzymane wyniki.

---

<sup>10</sup> Szpyra-Kozłowska (2000: 60) przytacza wyniki badań Szymanka, który badał związek między morfologią a fonologią, tj. związek między długością przymiotnika a wyborem sufiksu przymiotnikowego. Szymanek pokazał go na przykładzie derywacji przymiotników za pomocą sufiksu *-ski/-cki*, np. *prezydent > prezydencki, minister > ministerski*. Okazuje się, że model działa na podstawach więcej niż jednosylabowych. Do baz jednosylabowych dodaje się bowiem przyrostek *-owski*, np. *szpieg > szpiegowski*.



## Bibliografia

- Bałabaniak, D. Kategoria intensywności na tle wykładników stopnia // *Polonica*. – 2007. – № 28. – P. 13–21.
- Bartnicka, B., Rybicka, H. Analiza statystyczna form stopnia opisowego w języku polskim, czeskim i słowackim // *Zagadnienia kategorii stopnia w językach słowiańskich* – 1984. Vol. 4. – P. 137–150.
- Buttler, D., Kurkowska, H., Satkiewicz, H. *Kultura języka polskiego. Zagadnienia poprawności gramatycznej*. – Warszawa : Państwowe Wydawnictwo Naukowe, 1986.
- Dick, J.H. 1976, Stopniowanie przymiotników i przysłówków we współczesnej polszczyźnie kulturalnej // *Zagadnienia kategorii stopnia w językach słowiańskich*. – 1976. Vol. 1. – P. 43–52.
- Grochowski, M. Operatory metatekstowe o kształcie superlatiwu przysłówka // *Јужнословенски Филолог*. – 2008. – № 64. – P. 61–72.
- ISJP: Bańko, M. (Ed.) *Inny słownik języka polskiego*. Warszawa : Wydawnictwo Naukowe PWN, 2000.
- Jadacka, H. *Kultura języka polskiego. Fleksja, słowotwórstwo, składnia*. Warszawa : Wydawnictwo Naukowe PWN, 2005.
- Janus, E. Wykładniki intensywności cechy we współczesnej polszczyźnie // *Polonica*. – 1977. – № 3. – P. 47–57.
- Jędrzejko, E., Kita, M. *Gramatyka polska : podstawy wiedzy o budowie języka polskiego*. Warszawa : Ex Libris, 2000.
- Kaleta, Z. *Gramatyka języka polskiego dla cudzoziemców*. Kraków : Wydawnictwo Uniwersytetu Jagiellońskiego, 1995.
- Klebanowska, B., Kochański, W., Markowski, A. *O dobrej i z złej polszczyźnie*. Warszawa : Wiedza Powszechna, 1985.
- Marcisz, A. Normatywne założenia użycia form stopnia gramatycznego a faktyczny stan języka // *Prace Filologiczne*. – 2009. – № 56. – P. 293–301.
- Markowski, A. *Kultura języka polskiego. Teoria, zagadnienia leksykalne*. Warszawa : Wydawnictwo Naukowe PWN, 2005.
- NSPP: Markowski, A. (Ed.) *Nowy słownik poprawnej polszczyzny*. Warszawa : Wydawnictwo Naukowe PWN, 1999.
- Przepiórkowski A. et al. *Narodowy Korpus Języka Polskiego*. Warszawa : Wydawnictwo Naukowe PWN, 2012.
- Szpyra-Kozłowska, J. Słowotwórstwo bez morfemów // *Biuletyn PTJ* – 2000. – № 56. – P. 59–78.
- Saloni, Z. et al. *Słownik gramatyczny języka polskiego*. Url: <http://sgjp.pl> (data обращения: 26.09.2016).
- USJP: Dubisz, S. (Ed.) *Uniwersalny słownik języka polskiego*. Warszawa : Wydawnictwo Naukowe PWN, 2004.