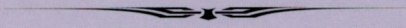


Prace językoznawcze  
dedykowane  
Profesor Jadwidze Sambor



## O ZALEŻNOŚCI MIĘDZY WŁAŚCIWOŚCIAMI SKŁADNIOWYMI I FREKWENCYJNYMI LICZEBNIKÓW GŁÓWNYCH W POLSZCZYŹNIE

### 0. Wstęp

Celem niniejszego artykułu jest przedyskutowanie wysuniętej przez Heinego (1997) tezy (zreferowanej niżej), iż różnice między zachowaniem składniowym liczebników głównych tak zwanych niższych (*jeden, dwa, trzy i cztery*) i wyższych (np. *tyśiąc*) wynikają z różnic w częstości ich występowania. W naszych rozważaniach wykorzystamy dane z badań frekwencyjnych nad polszczyzną — przede wszystkim ze *Słownika frekwencyjnego polszczyzny współczesnej* (Kurcz, Lewicki, Sambor, Szafran i Woronczak (1990), dalej — SFPW).

### 1. Model Heinego (1997)

Heine (1997) zauważa, że jedną z najczęstszych cech systemów liczebnikowych w różnych językach świata jest podział na liczebniki niższe i wyższe. Czasami podział ten przybiera szczególnie wyraźną formę: liczebniki wyższe niż *trzy* lub *cztery* po prostu nie istnieją, tak jak np. w językach Nowej Gwinei, por. Greenberg (1978). W innych wypadkach liczebniki obu grup różnią się budową morfologiczną. Liczebniki niższe są najczęściej etymologicznie nieprzejrzyste (muszą więc być uznane za elementy etymologicznie i strukturalnie prymarne). Dobrym przykładem jest język mamvu (nilo-saharyjski z centralnego Sudanu), w którym np. 'dwa' (*juè*) i 'cztery' (*jetò*) to niepodzielne morfologicznie słowa o nieznannej etymologii, podczas gdy 'siedem' (*el' qodè juè*, czyli dosłownie 'ręka łapie dwa') i 'osiem' (*jetò jetò*, czyli 'cztery cztery') są tworzone z prostszych elementów.

W językach słowiańskich podział ten znajduje odzwierciedlenie w różnicach morfosyntaktycznych. Wielu badaczy (w tym autor omawianej tu pracy, a także cytowani przez niego Greenberg 1978 i Corbett 1978) uważa, że liczebniki główne, wyodrębnione ze względów semantycznych jako odrębna część mowy w różnych językach, wykazują cechy morfologiczne i składniowe bądź rzeczowników, bądź przymiotników<sup>1</sup>. „Przymiotnikowość” liczebników głównych polega na tym, iż

<sup>1</sup> Rzadziej, zapewne, także przysłówków, tak jak w językach Azji Wschodniej — Romuald Huszcza, informacja ustna.

zajmują podrzędną pozycję w grupie nominalnej, najczęściej przed rzeczownikiem<sup>2</sup>. Ponadto mogą wchodzić w uzgodnienia w zakresie rodzaju i przypadku. Z kolei „rzeczownikowość” wyraża się tym, że to liczebnik stanowi nadrzędnik frazy nominalnej i narzuca przypadek rzeczownikowi, np. łac. *tres centi puerorum* dosł. ‘trzy setki chłopców’. Otóż w językach słowiańskich liczebniki „niższe” zachowują się raczej przymiotnikowo: *dwaj młodzi chłopcy, dwoma młodymi chłopcami, ... dwie młode dziewczyny, dwie ciekawe książki* itd. Liczebniki „wyższe” natomiast wykazują więcej cech rzeczownikowych. Skrajny przykład to klasa słów takich jak *tysiąc, milion* i wyższe, bezwyjątkowo rządzące rzeczownikiem i odmieniające się przez liczbę, np. *pięcioma tysiącami żołnierzy*. Liczebniki wyższe niż *cztery* a niższe niż *tysiąc* łączą w sobie, co pokażemy dalej (w podrozdziale 2.1), cechy przymiotnikowe i rzeczownikowe (por. też Rutkowski 2001).

Analizując odrębność syntaktyczną liczebników niższych w językach słowiańskich, Heine (1997) przytacza propozycję dwóch uniwersaliów sformułowanych przez Corbetta (1978):

1. Zachowanie syntaktyczne prostych liczebników głównych będzie zawsze czymś pośrednim między zachowaniami przymiotników i rzeczowników.
2. Jeśli proste liczebniki główne w danym języku różnią się zachowaniem syntaktycznym, to te, które bardziej przypominają rzeczowniki będą określać liczby wyższe niż te, które mniej przypominają rzeczowniki.

Zależność opisana w drugim z uniwersaliów występuje w wielu językach i ma wynikać z tego, że z diachronicznego punktu widzenia liczebniki wyższe są najczęściej po prostu rzeczownikami, które wraz z rozwojem liczenia były stopniowo włączane do systemu liczebnikowego. Przykładem ma być staro-cerkiewno-słowiańskie słowo *ѡма*, które od rzeczownika znaczącego ‘mnóstwo’ przeszło w liczebnik znaczący ‘dziesięć tysięcy’. Jest to, zdaniem Heinego, potwierdzenie ogólniejszej prawidłowości związanej z gramatyzacją: rzeczowniki konkretne stanowią najczęstszą podstawę do swoistego przesunięcia konceptualnego (ang. *conceptual transfer*), w wyniku którego powstają elementy zamkniętych klas leksykalnych (nazywane też często elementami funkcjonalnymi lub gramatycznymi — np. przyimki, łączniki, rodzajniki, liczebniki). Taki model rozwoju systemu gramatycznego języka zakłada „ugramatycznianie” elementów leksyki konkretnej poprzez nadawanie im znaczenia cech bardziej abstrakcyjnych. W myśl założenia Heinego nie ma powodu, by twierdzić, że tylko liczebniki wyższe wykształciły się z pierwotnych rzeczowników. Rzeczownikową genezę mogą mieć także liczebniki niższe. Ich przesunięciu od elementu konkretnego do abstrakcyjnego towarzyszyła jednak zmiana statusu kategoryalnego leksemu: z rzeczownikowego

<sup>2</sup> Nawet w językach, w których nienacechowany jest postpozycyjny szyk przydawki przymiotnej, jak np. w językach romańskich.

na przymiotnikowy. Potwierdzeniem zasadności powyższego modelu wydaje się Heinemu historyczny rozwój liczebników słowiańskich (zob. niżej, podrozdział 2.2).

Postulowane zmiany owocują w językach słowiańskich swego rodzaju kontinuum (z różnymi stopniami „przymiotnikowości”), nie zaś jednolitym przesunięciem wszystkich liczebników rzeczownikowych do klasy przymiotników. Według Heinego czynnikiem odpowiedzialnym za tę nierównomierność zmian językowych jest częstość użycia. Im niższy liczebnik, tym częściej jest używany. Teoria gramatyzacji zakłada zaś, że tylko elementy często używane mogą podlegać zmianom kategoryalnym.

Taka argumentacja wydaje się uzasadniona. Przeciętny użytkownik języka z pewnością częściej używa liczebnika *dwa* niż *bilion*. Dotyczy to w szczególności społeczności o niskim poziomie edukacji i techniki. Dlaczego jednak społeczeństwa, w których posługiwanie się liczbami i liczebnikami jest szeroko rozpowszechnione (takie jak np. społeczeństwa słowiańskojęzyczne i anglojęzyczne), nie uprzymiotnikowiły, czyli — w myśl założeń Heinego (1997) — nie uabstrakcyjniały liczebników w równym stopniu? W języku angielskim wszystkie liczebniki są dziś modyfikatorami rzeczownika zbliżonymi syntaktycznie do przymiotników. Natomiast w językach słowiańskich odrębność gramatyczna liczebników „niższych” i „wyższych” jest wyraźna. Przyczyną tego mogłaby być utrzymująca się różnica częstości użycia liczebników niższych i wyższych. W dalszej części tego artykułu postaramy się sprawdzić słuszność tej hipotezy w SFPW.

## 2. Cechy systemowe polskich liczebników głównych w świetle hipotezy Heinego

Szczegółowa konfrontacja omawianej tu tezy z danymi z SFPW wymaga bliższego przyjrzenia się cechom składniowym liczebników głównych w polszczyźnie, zasygnalizowanym powyżej. Postaramy się także zestawić model Heinego (1997) z faktami diachronicznymi dotyczącymi tej klasy leksemów.

### 2.1. Fakty synchroniczne

Nawet pobieżna analiza właściwości fleksyjnych i składniowych leksemów nazywanych w tradycji polonistycznej liczebnikami głównymi pokazuje, że jest to klasa wewnątrznie niejednolita. Za liczebniki główne uznaje się zazwyczaj następujące jednostki słownikowe (zob. np. Doroszewski 1980, Saloni 1977, Laskowski 1984 i 1991):

1. *jeden*
2. *dwa, oba, obydwa, trzy, cztery, pięć, sześć, siedem, osiem, dziewięć, dziesięć*;

3. *jedenaste, dwanaście, trzynaście, czternaście, piętnaście, szesnaście, siedemnaście, osiemnaście, dziewiętnaście*;
4. *dwadzieścia, trzydzieści, czterdzieści, pięćdziesiąt, sześćdziesiąt, siedemdziesiąt, osiemdziesiąt, dziewięćdziesiąt*;
5. *sto, dwieście, trzysta, czterysta, pięćset, sześćset, siedemset, osiemset, dziewięćset*;
6. *tysiąc, milion, miliard, bilion, biliard, trylion, tryliard etc.*<sup>3</sup>

Status niektórych wymienionych wyżej jednostek jest dyskusyjny. Przy powierzchniowo-morfologicznym kryterium wyodrębniania części mowy i uznawaniu, jak to robią między innymi Saloni (1974) i Laskowski (1984, 1991), że dla liczebników głównych definicyjne jest posiadanie klasyfikacyjnej kategorii liczby trudno uznać za liczebniki leksemy wymienione w punkcie 6. (np. *tysiąc* — *dwa tysiące*; zob. także wyżej). Autorzy SFPW traktują te jednostki jak rzeczowniki. Równie ciekawy jest przypadek jednostki *jeden*, dla którego liczba wydaje się być kategorią fleksyjną (np. *jeden chłopiec* — *jedni chłopcy*) — na tej podstawie Saloni (1974) wyklucza *jeden* spośród liczebników głównych. Istnieją jednak argumenty przemawiające za uznaniem, że w parze takiej reprezentowane są różne jednostki o wspólnej postaci hasłowej *jeden*, a mianowicie liczebnik i zaimek. Forma liczby mnogiej jest formą zaimka, forma liczby pojedynczej jest dwuznaczna. Na korzyść zaimkowego charakteru jednostki reprezentowanej przez *jedni* świadczy między innymi zablokowanie możliwości pojawienia się we frazie zaimka wskazującego, *\*ci jedni chłopcy*, do którego nie dochodzi ani we frazie z leksemem *jeden* w liczbie pojedynczej (*ten jeden chłopiec*), ani we frazie z niewątpliwym liczebnikiem *dwa* (*ci dwaj chłopcy, tych dwóch chłopców*)<sup>4</sup>.

Jak już wspominaliśmy, polskie liczebniki główne rozpadają się na grupy o odmiennych właściwościach składniowych. Leksemy *jeden, dwa, oba, obydwaj, trzy* i *cztery* wykazują wszystkie cechy składniowe przymiotników: uzgadniają wartości przypadka i rodzaju z kwantyfikowanym rzeczownikiem, np. *trzej młodzi chłopcy, trzy młode dziewczyny, dla trzech młodych chłopców//dziewczyn, z trzema młodymi chłopcami//dziewczynami* — dalej będziemy je nazywać liczebnikami typu A (za Rutkowskim 2000). Przymiotnikowy charakter tego typu liczebników cechuje wszystkie języki słowiańskie (por. Giusti, Leko 1995).

<sup>3</sup> Jako liczebniki główne są również traktowane liczebniki sygnalizujące ilość kwantyfikowanych przedmiotów w sposób przybliżony (nazywa się je liczebnikami nieokreślonymi — por. Laskowski 1984 i 1991), np. *kilka, kilkanaście, kilkadziesiąt, kilkaset, parę, paręnaście, parędziesiąt, paręset, ile, ileś, ilekolwiek, tyle, wiele*, oraz liczebniki niewłaściwe takie jak *mało, dużo, sporo, trochę*. Pomijamy je w tym artykule, gdyż leksemy tego typu nie zostały objęte opisem przez Heinego (1997).

<sup>4</sup> Brak łączliwości z zaimkiem jako kryterium zaimkowości zaczerpnęliśmy z prac Huszczy (1996, 2000), który stosuje je do wyodrębnienia zaimków osobowych.

Konstrukcje z liczebnikami wyższymi niż *cztery* (nazywanymi przez nas dalej liczebnikami typu Q) rządzą się zupełnie innymi prawami. Mimo że z semantycznego punktu widzenia liczebniki spełniają w nich funkcję członu zależnego, pozostają z członem głównym grupy nominalnej w stosunku wzajemnej determinacji gramatycznej. Rodzaj gramatyczny liczebnika zależny jest od liczonego rzeczownika, np. rzeczowniki o słownikowej kategorii rodzaju żeńskiego łączą się z liczebnikami o fleksyjnej kategorii rodzaju niemęskoosobowego (*pięć młodych dziewczyn//małych kotów//nowych stołów* w odróżnieniu od *pięciu młodych mężczyzn*). Z kolei przypadek rzeczownika jest w pewnych kontekstach zależny od liczebnika. Gdy fraza zawierająca liczebnik znajduje się w pozycji podmiotu, liczebnik przypisuje następującemu po nim rzeczownikowi dopełniacz (tzw. dopełniacz kwantyfikacji — por. Franks 1995), np. *pięć dziewczyn czyta//\*pięć dziewczyny czytają* (kwestię uzgodnienia formy czasownika pomijamy). Dopełniacz kwantyfikacji występuje też, gdy fraza zawierająca liczebnik zajmuje pozycję biernikową, czy to w funkcji dopełnienia bliższego, czy to po przyimku, np. *widzę pięć dziewczyn//\*pięć dziewczyny, przez pięć dziewczyn//\*przez pięć dziewczyny*. W pozostałych przypadkach między liczebnikiem a rzeczownikiem zachodzi składnia zgody, np. *ufam pięciu dziewczynom//chłopcom, opowiadam o pięciu dziewczynach//chłopcach*. Dystrybucja konstrukcji w składni rządu (z dopełniaczem kwantyfikacji) i w składni zgody — odpowiednio hetero- i homogenicznych w terminologii, którą proponuje Babby (1988) — wydaje się korelować z podziałem na przypadki strukturalne, inaczej gramatyczne, i leksykalne, inaczej konkretne (por. Heinz 1965, Chomsky 1986; bardziej szczegółowo na ten temat: Przepiórkowski 1996 i Rutkowski 2001).

Ograniczenie dystrybucji dopełniacza kwantyfikacji nie dotyczy, jak już wspomniano, liczebników *tysiąc* i wyższych, np. *ufam tysiącu dziewczyn//chłopców, opowiadam o tysiącu dziewczyn//chłopców*. Te właśnie liczebniki chcemy — ze względu na ich wspomniany już skrajnie rzeczownikowy charakter — nazywać liczebnikami typu N (por. Rutkowski, Szczegot 2001).

## 2.2. Fakty diachroniczne

Zdaniem Heinego (1997) to, że część leksemów liczebnikowych charakteryzuje się bardzo wieloma cechami rzeczownikowymi (liczebniki typu N), a część (liczebniki typu Q) ma w dzisiejszej polszczyźnie status pośredni, potwierdza hipotezę stopniowego przejmowania cech przymiotnikowych przez liczebniki. Można by stwierdzić, że polskie liczebniki typu Q znajdują się obecnie w połowie drogi od rzeczownikowości do przymiotnikowości (ta ostatnia cechuje już przymiotniki najniższe — typu A).

Dane diachroniczne zdają się potwierdzać tę tezę. Klemensiewicz, Lehr-Spławiński i Urbańczyk (1964) podają, że liczebniki od *pięć* wzwyż pierwotnie

wymagały od liczonego rzeczownika dopełniacza we wszystkich przypadkach. Występowały więc konstrukcje takie jak pokazane w poniższej tabeli (pola zacienione pokazują, w jakich kontekstach przypisywany jest dopełniacz kwantyfikacji):

	Dane diachroniczne	Wiek	Polszczyzna współczesna
Kontekst przypadków gramatycznych			
Nom	sześć grzywien	XV	sześć grzywien
	siedm grzechow	XVI	siedem grzechów
Acc	przez sześć świadkow	XV	przez sześciu świadków
	przez trzydzieści lat	XVI	przez trzydzieści lat
Kontekst przypadków konkretnych			
Dat	ku czterem stom mężow	XV	ku czterystu mężom
	przeciw sześćdziesiąt dział	XVII	przeciw sześćdziesięciu działom
Instr	siedmią ran	XVII	siedmioma ranami
	pod czterdzieści chorągwi	XVI	pod czterdziestoma chorągwiami
Loc	w ośmi lat	XV	w ośmiu latach
	we stu koni	XVII	w stu koniach

Tabela 1. Historyczna składnia liczebników typu Q

Liczebniki takie jak *pięć* łączyło z rzeczownikami nie tylko podobieństwo składniowe. Ich odmiana odpowiadała odmianie rzeczowników żeńskich z dawnym tematem na \*-i, np. *pięć*. Zmiany syntaktyczne zaczęły pojawiać się w tym systemie co prawda już w piętnastym wieku (konstrukcje typu *sześcią groszami*, *we dwudziestu dniach*), ale dzisiejszy stan rzeczy staje się powszechny dopiero w wieku dziewiętnastym. Od szesnastego wieku rozpoczynają się także zmiany fleksyjne upodabniające omawiane liczebniki do liczebnika przymiotnikowego *dwa*. Najważniejsza z nich to pojawienie się końcówki -u (charakterystycznej dla formy *dwu*), np. *pięciu ran*, która wypiera rzeczownikową końcówkę -i.

### 2.3. Podsumowanie

Obserwujemy zatem tendencję postulowaną przez Heinego: w wyniku zmian językowych liczebniki stają się coraz bliższe przymiotnikom. Im niższe liczebniki, tym proces jest aktywniejszy (dlatego zmiany nie obejmują leksemu *tysiąc*). Co więcej, wskazać można także argumenty słowotwórcze przemawiające za

tym, że zmiany zachodzą szybciej w liczebnikach niższych. Z diachronicznego punktu widzenia dzisiejsze leksemy typu *dwanaście* są kontynuantami prasłowiańskich konstrukcji składniowych takich jak *dva na desęte*. Uproszczenie, ujednoczenie i, jak nazwałby to Heine, uabstrakcyjnienie tej struktury nastąpiło szybciej niż odpowiadającej wyższemu liczebnikowi struktury *tri sęta*, której kontynuacja do dziś jest odbierana jako w pewnym sensie morfologicznie złożona (o czym świadczy chociażby spotykana jeszcze odmiana typu *trzechset*).

Niemniej można w polszczyźnie znaleźć fakty niezgodne z tezą Heinego. W świetle zarysowanej tendencji niezrozumiałe wydaje się zachowanie leksemów *dwa*, *oba*, *obydwa*, *trzy* i *cztery* (czyli liczebników typu A), które we współczesnej polszczyźnie mogą w połączeniu z rzeczownikami męskoosobowymi przyjmować w pozycji podmiotu dwie formy *trzej chłopcy idą/trzech chłopców idzie*. Wielu polskich lingwistów (np. Saloni 1977, Laskowski 1984) uznaje w takich przykładach *trzech* za alternatywną formę mianownika<sup>5</sup>. Oznaczałoby to koegzystencję dwóch form, z których pierwsza odpowiada schematowi składni liczebników typu A, druga zaś — schematowi składni liczebników typu Q. W świetle modelu Heinego ta ostatnia powinna być formą starszą, która stopniowo się uprzymiotnikowiła, co zaowocowało innowacją *trzej*. Jednak dane diachroniczne wskazują, że rozwój leksemów takich jak *trzy* był odwrotny<sup>6</sup>. Forma *trzej* to powstała w piętnastym wieku kontynuacja starej formy *trze* (z *j* będącym wynikiem asynchronicznej wymowy ścieśnionego *e*). Forma *trzech* w funkcji mianownika pojawia się natomiast dopiero w siedemnastym wieku. Model Heinego (1997) nie uwzględnia zaś możliwości wtórnego „odprzymiotnikowienia” liczebnika<sup>7</sup>.

Najpoważniejszym zarzutem, który trzeba postawić Heinemu, jest chyba niemożność sfalsyfikowania jego tezy. To, że rzeczywiste zmiany diachroniczne są zbliżone do założeń referowanego modelu, nie dowodzi wcale, że ich powodem były różnice frekwencyjne. Nie dysponując odpowiednimi danymi, nie możemy jednoznacznie stwierdzić, jakie częstości charakteryzowały poszczególne leksemy kilkaset lat temu. Możemy natomiast szukać wskazówek w danych synchronicznych. Te jednak, o czym będzie mowa w następnym rozdziale, nie tylko nie wykazują zbieżności z analizą Heinego (1997), ale wręcz odkrywają jej braki.

<sup>5</sup> Inny pogląd przeważa w pracach generatywnych, gdzie formy takie uznaje się za podmiot biernikowy (por. Franks 1995, Rutkowski 2000).

<sup>6</sup> Zwraca na to uwagę także Szczegot (2001).

<sup>7</sup> Utrata cech przymiotnikowych przez liczebniki niższe wcale nie jest tak rzadka. W większości języków romańskich (poza np. portugalskim), liczebnik ‘dwa’ nie zachował odmiany przez rodzaj, która istniała w łacinie, podobnie jak odmiana przez rodzaj liczebnika ‘trzy’.



### 3. Dane frekwencyjne

Jedynym kompleksowym opracowaniem frekwencyjnym współczesnego języka polskiego jest SFPW. Podaje ono informacje o częstościach leksemów oraz ich słowoform. Korpus, na podstawie którego stworzono SFPW, liczy 500 000 słów i dzieli się na pięć równych części reprezentujących pięć stylów funkcjonalnych: teksty popularnonaukowe, drobne wiadomości prasowe, publicystyka, proza artystyczna i dramat artystyczny. Teksty, które go tworzą, pochodzą z lat 1963–1967. Materiałem porównawczym są dane o występowaniu hasel z SFPW w korpusie polszczyzny mówionej przygotowane na podstawie słownika Zgółkowej (1983).

W SFPW za liczebniki uznano leksemy odmieniające się przez przypadek, nie odmieniające się przez liczbę i oznaczające stosunki liczbowe (jest to więc definicja zbliżona do tych proponowanych przez Saloniego 1974 i Laskowskiego 1984). W skład liczebników (klasy hasel oznaczonych w słowniku skrótem *Num*) wchodzi następujące leksemy (w nawiasach podana jest ogólna frekwencja każdego z nich):

*dwa* (936), *trzy* (568), *cztery* (373), *pięć* (431), *sześć* (240), *siedem* (164), *osiem* (221), *dziewięć* (135), *dziesięć* (202);

*jedenaste* (42), *dwanaście* (87), *trzyaste* (40), *czternaście* (46), *piętnaście* (109), *szesnaście* (39), *siedemnaście* (42), *osiemnaście* (38), *dziewiętnaście* (16);

*dwadzieścia* (400), *trzydzieści* (266), *czterdzieści* (185), *pięćdziesiąt* (262), *sześćdziesiąt* (229), *siedemdziesiąt* (123), *osiemdziesiąt* (124), *dziewięćdziesiąt* (79); *sto* (330), *dwieście* (161), *trzysta* (101), *czterysta* (65), *pięćset* (102), *sześćset* (67), *siedemset* (51), *osiemset* (85), *dziewięćset* (572);

*oba* (170), *obydwa* (23);

*dwoje* (19), *oboje* (15), *troje* (8), *czworo* (6);

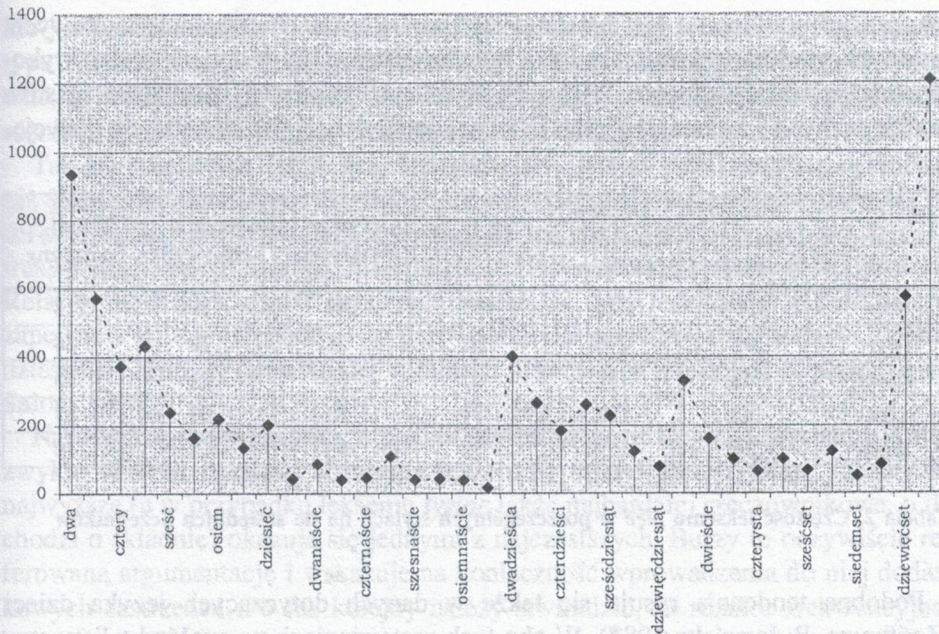
*pół* (84), *półtora* (22);

*kilka* (400), *kilkanaście* (39), *kilkadziesiąt* (37), *kilkaset* (22), *parę* (96);

*ile* (51), *ileż* (4), *tyle* (83), *wiele* (424), *niewiele* (14)

Jak była mowa powyżej, autorzy SFPW nie uznają za liczebniki leksemów takich jak *tysiąc* (frekwencja: 1207).

Ogólna frekwencja to wartość wskaźnika F, czyli częstości absolutnej (wynik zsumowania wszystkich wystąpień hasła w korpusie). Już te podstawowe dane pokazują, że analizy Heinego (1997) fakty synchroniczne w zasadzie nie potwierdzają. Częstości wystąpień liczebników we współczesnej polszczyźnie nie wydają się w jednoznaczny sposób zależne od wielkości liczb naturalnych, do których odnoszą się liczebniki. Ilustruje to poniższy wykres (wartości wskaźnika F połączone są linią przerywaną, aby uwydatnić różnice frekwencyjne między poszczególnymi leksemami).



Rycina 1. Częstotliwość absolutna leksemów liczebnikowych (oraz leksemu *tysiąc*)

Liczebniki typu A (*dwa*, *trzy* i *cztery*) są co prawda o wiele częstsze niż np. *siedem* lub *dziewięć*, ale bardzo wysoką częstością wystąpień cechuje się także liczebnik *pięć*, którego cechy składniowe nie różnią się niczym od cech liczebników występujących zdecydowanie rzadziej. Obserwacja ta nie ogranicza się do leksemu *pięć*. W danych z SFPW prawidłowością wydaje się wysoka częstość wystąpień wszystkich leksemów liczebnikowych związanych z liczbą pięć. Prawidłowość ta powoduje swego rodzaju „załamanie” gradacji frekwencyjności (która odpowiadałaby założeniom Heinego 1997). W każdej z trzech podstawowych serii liczebników prostych (od *dwa* do *dziewięć*, od *jedenaste* do *dziewiętnaste* i od *dziesięć* do *dziewięćdziesiąt*) leksem zbudowany na bazie liczebnika *pięć* jest częstszy od tych, których podstawą słowotwórczą są *cztery* i *sześć*. Wy tłumaczeniem takiego stanu rzeczy jest oczywiście powszechne postrzeganie liczby pięć jako podstawowej w systemie dziesiętnym (obok liczby dziesięć) i „okrągłej” (por. Ifrah 1981), co powoduje jej częste występowanie w różnego rodzaju przybliżeniach.

Oznacza to, że analiza Heinego (1997), jeśli uznać ją za właściwą, musi odnosić się do kształtowania języka w czasach poprzedzających rozpowszechnienie systemu liczenia „premiującego” frekwencyjnie leksemy oparte o podstawę słowotwórczą *pięć*. Pewnym potwierdzeniem takiego założenia jest np. porównanie danych z różnych stylów. Jak pokazuje tabela 2, wysoka frekwencja

leksemu *pięć* widoczna jest przede wszystkim w tekstach popularnonaukowych, drobnych wiadomościach prasowych i publicystyce, czyli w stylach często wykorzystujących dane liczbowe. W danych z bliższych językowi potocznemu tekstów dramatycznych i z samego języka potocznego *pięć* traci swą wyjątkową pozycję.

Leksem	Teksty popularnonaukowe	Drobne wiadomości prasowe	Publicystyka	Proza artystyczna	Dramat artystyczny	Styl potoczny
dwa	199	322	157	147	111	218
trzy	118	240	82	58	70	84
cztery	65	135	67	33	23	65
pięć	83	187	97	25	39	57
sześć	34	129	52	13	12	19

Tabela 2. Częstość leksemu *pięć* w poszczególnych stylach na tle sąsiednich liczebników

Podobna tendencja rysuje się także w danych dotyczących języka dzieci (Zgółkova, Bułczyńska 1987). W obu tych zestawieniach na czołówkę listy rangowej składają się liczebniki najniższe (typu A) oraz leksem *pięć*, jednak różnice między tym ostatnim a leksemem *cztery* są znacznie mniejsze niż w danych zbiorczych z SFPW.

Leksem	Częstość absolutna w języku dzieci
dwa	207
trzy	90
cztery	47
pięć	50

Tabela 3. Wskaźnik F dla najczęstszych leksemów liczebnikowych w wynikach badań języka dzieci

Można uznać, że właśnie język potoczny i język dzieci są w największym stopniu niezależne od czynników cywilizacyjnych (takich jak powszechna edukacja i masowy przepływ informacji podawanych w liczbach). Co więcej, przy założeniu, że rozwój języka u dzieci odzwierciedla w pewien sposób historyczny rozwój języka jako takiego (ontogeneza odzwierciedla filogenezę) można by argumentować, że takie same właściwości frekwencyjne charakteryzowały najniższe leksemy liczebnikowe przed rozwojem systemów pieniężnych, kalendarzy oraz masowej edukacji (która rozpowszechniła znajomość liczb i wymusiła częstsze używanie liczebników wyższych). Niemniej nawet w danych dotyczących stylu potocznego

lub języka dzieci nie pojawia się wyraźna cezura między frekwencją liczebników typu A (*dwa, trzy, cztery*) i liczebnika *pięć*. Brak tym samym podstaw, by twierdzić, że kiedykolwiek istniała, a tylko ona mogłaby tłumaczyć istniejące do dziś różnice syntaktyczne między omawianymi liczebnikami.

Tak jak przewiduje Heine, liczebniki od *jedenaste* do *dziewiętnaste* są rzadsze niż liczebniki określające liczby jednocyfrowe (zgodnie z zasadą mówiącą, że im wyższy liczebnik, tym rzadsze użycie). Jednak charakteryzujące się wysoką frekwencją leksemy takie jak *dwadzieścia* i *sto* wyraźnie tej prawidłowości przeczą. Relatywnie niską częstość wystąpień leksemów typu *jedenaste* wiązać należy zapewne z ich ograniczoną — w porównaniu do leksemów oznaczających pełne dziesiątki i setki — łączliwością w ramach liczebników złożonych (Gruszczynski, Saloni 1978).

Największą niespodzianką z punktu widzenia teorii Heinego jest jednak niezwykle wysoka częstość absolutna leksemów *dziewięćset* i *tysiąc*. Liczebniki najwyższe (a w przypadku leksemu *tysiąc* także najbardziej rzeczownikowe, o ile chodzi o składnię) okazują się jednymi z najczęstszych. Burzy to oczywiście referowaną argumentację i wskazuje na konieczność wprowadzenia do niej dodatkowych zastrzeżeń. Po raz kolejny należy stwierdzić, że różnice frekwencyjne, które mogły wpłynąć na wykształcenie odmiennych właściwości morfosyntaktycznych liczebników typu A i N, dotyczą stanu języka sprzed upowszechnienia umiejętności liczenia. Powszechne liczenie pozwala bowiem użytkownikom języka na wprowadzanie do niego na dużą skalę elementów motywowanych zewnętrznie (por. Pisarek 1972). Tak na przykład w realiach Polski lat sześćdziesiątych mogły to być słowa *tysiąc* i *dziewięćset* używane przy określaniu dat lub *tysiące*, w których liczono pieniądze). Mimo niezwykle częstego występowania, takie elementy nie podlegają jednak uprzymiotnikowaniu, co oznacza, że trudno uznać procesy opisane przez Heinego (1997) za aktywne synchronicznie.

Oprócz wskaźnika częstości absolutnej F w SFPW podawane są dla każdego leksemu wartości dwóch innych wskaźników statystycznych: U (częstość względna) i D (wskaźnik równomierności rozkładu w stylach — dyspersja złożona). Ten ostatni przyjmuje wartości od 0 (maksymalne skoncentrowanie — wystąpienia tylko w jednym stylu) do 100 (całkowicie równomierny rozkład we wszystkich stylach). Autorzy SFPW wyróżniają na podstawie wskaźników statystycznych F, U i D cztery strefy leksyki:

1. słownictwo gramatyczne (największe częstości i równomierność rozkładu powyżej 80)
2. słownictwo podstawowe (wysoka częstość i równomierność rozkładu powyżej 50)
3. słownictwo charakterystyczne (hasła częste, ale nie we wszystkich stylach — dyspersja złożona poniżej 50, częstość absolutna powyżej 10)
4. słownictwo rzadkie (niska częstość absolutna)

Ze względu na to kryterium jedynie leksem *dwa* można zaliczyć do słownictwa gramatycznego. W ramach modelu Heinego cecha taka wydaje się oczywistą konsekwencją jego przymiotnikowości. Jak była mowa w rozdziale 1, nabywanie cech przymiotnika ma być efektem gramatyzacji i uabstrakcyjnienia danego leksemu. Jednak większość pozostałych leksemów liczebnikowych należy w klasyfikacji opartej na kryteriach SFPW uznać za elementy słownictwa podstawowego, co w niczym nie tłumaczy różnic w ich zachowaniu syntaktycznym. Elementami podstawowego zasobu leksykalnego są bowiem liczebniki wszystkich typów: A (np. *cztery*,  $D = 61,23$ ), Q (np. *osiem*,  $D = 61,3$ ) i N (np. *tysiąc*,  $D = 64,5$ ). Ilustruje to poniższa tabela.

Leksem	Wskaźnik D
dwa	80,5
dziesięć	72,16
trzy	70,81
piętnaście	66,94
pięć	66,92
pięćset	65,18
dwadzieścia	65,15
tysiąc	64,5
dwanaście	62,68
trzydzieści	62,09
osiem	61,3
cztery	61,23

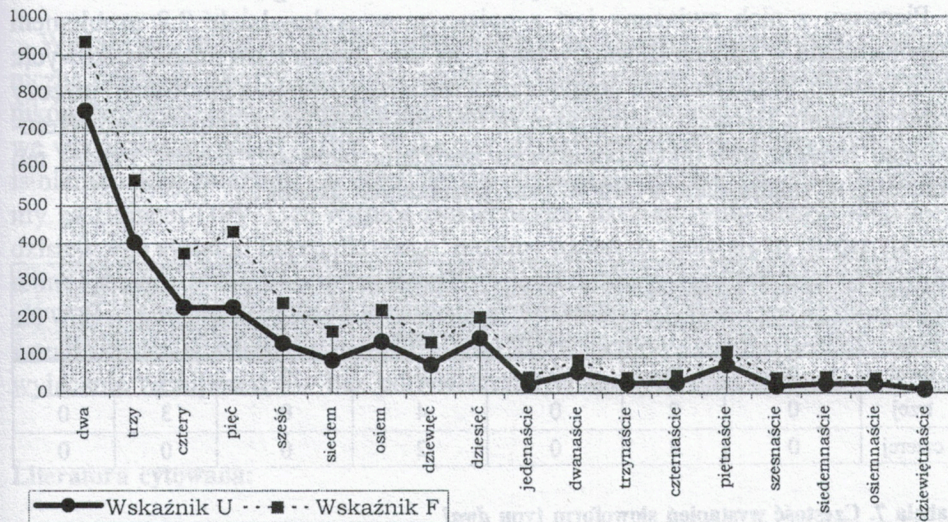
Tabela 4. Wskaźnik równomierności rozkładu leksemów liczebnikowych w poszczególnych stylach (najwyższe pozycje listy rangowej)

Aż pięć leksemów liczebnikowych typu Q jest natomiast według kryteriów zastosowanych w SFPW elementami słownictwa charakterystycznego (które nie powinno raczej podlegać procesowi gramatyzacji).

Leksem	Wskaźnik D	Wskaźnik F
dziewięćdziesiąt	48,96	79
czterysta	45,66	65
szesnaście	44,65	39
osiemnaście	41,88	38
osiemset	39,24	51

Tabela 5. Wskaźnik równomierności rozkładu leksemów liczebnikowych w poszczególnych stylach (najniższe pozycje listy rangowej)

Wskaźnik U (nazywany też współczynnikiem powszechności wyrazu — por. Sambor 1972), będąc wynikiem przemnożenia wskaźników F i D, relatywizuje dane o częstości. Wydaje się przez to w większym stopniu niż F odpowiadać tezom Heinego (1997). Użycie wskaźnika U niweluje np. w rezultatach badania różnice frekwencyjne między leksemami *pięć* ( $U = 228,43$ ) i *cztery* ( $U = 228,38$ ). Zestawione wartości wskaźników U i F dla niższych leksemów liczebnikowych przedstawia rycina 2.



Rycina 2. Częstości relacyjna i absolutna leksemów liczebnikowych od *dwa* do *dziewiętnaście*

Różnice częstości względnej między kolejnymi leksemami są mniejsze niż różnice częstości absolutnej, co powoduje, że bardziej widoczna jest ogólna tendencja zmian wysokości wskaźnika. W przypadku liczebników takich jak *sto* lub *tysiąc* zamiana wskaźnika F na U nie wnosi jednak niczego nowego — wartość U jest nadal zbyt wysoka, by można było ją uznać za wpisującą się w model Heinego:

Leksem	Wskaźnik U
dwadzieścia	260
sto	200,63
tysiąc	778,52

Tabela 6. Częstość relacyjna wyższych leksemów liczebnikowych

Poza badaniem cech frekwencyjnych całych leksemów, SFPW ułatwia także analizę poszczególnych słowoform. Dla każdej słowoformy określona jest wartość przypadku. Odstępstwem od tej reguły są jedynie nieodmienne słowoformy występujące w nazwach liczb wielocyfrowych (np. *tysiąc* i *dziewięćset* w sekwencji od *tysiąc dziewięćset piętego roku*) oraz nieodmienne przydawki ilościowe (np. *zwyciężyli Belgowie w stosunku pięć-cztery*). Można zatem dosyć dokładnie prześledzić częstości form liczebników typu Q w strukturach homogenicznych i heterogenicznych (zob. wyżej — podrozdział 2.1). Taka analiza dostarcza szczególnie istotnych argumentów przeciw tezom Heinego.

Pierwszy z nich związany jest z opisanym w podrozdziale 2.2 problemem współlistnienia w ramach leksemów takich jak *dwa* dwóch form mianownikowych: *dwaj* i *dwóch*. Pojawienie się drugiej z nich (będącej późniejszą innowacją) przeczy logice systemu Heinego. Jednak jeszcze bardziej niespójny z założeniami referowanego modelu jest fakt stopniowego zanikania pierwszej z nich. Proces ten pokazują wyraźnie dane z SFPW oraz pozostałych źródeł.

Słowo-forma	Styl potoczny	Język dzieci	Teksty popularno-naukowe	Drobne wiadomości prasowe	Publicystyka	Proza	Dramat
dwaj	0	2	2	9	0	4	1
trzej	0	2	0	4	0	3	0
czterej	0	3	0	2	0	0	0

Tabela 7. Częstość wystąpień słowoform typu *dwaj*

We wszystkich stylach częstość wystąpień słowoform typu *dwaj* jest niewielka. Szczególnie symptomatyczne są dane ze stylu potocznego, w których formy te nie występują wcale. Oznacza to, że mianownik rodzaju męskoosobowego leksemów *dwa*, *trzy* i *cztery* uległ tendencji odwrotnej niż postulowana przez Heinego: stracił własności przymiotnikowe na korzyść rzeczownikowych.

Innym poważnym problemem związanym z frekwencją poszczególnych słowoform, którego zdaje się nie rozwiązywać opis Heinego, jest wspomniana już wcześniej różnica między przypadkami gramatycznymi i konkretnymi. Jeśli przyjąć, że właściwości liczebników typu Q we współczesnej polszczyźnie są rezultatem procesów, które zmieniają ich charakter rzeczownikowy w przymiotnikowy, należy zapytać, dlaczego zmiany te rozpoczęły się od przypadków konkretnych. Jeśli bowiem gramatyzacja i nabieranie cech przymiotnikowych wynikałyby z częstszego występowania, formy przypadków konkretnych musiałyby występować częściej niż formy przypadków gramatycznych. Teza ta nie znajduje jednak potwierdzenia w danych z SFPW. W ogólnej liczbie 431 wystąpień leksemu *pięć* forma narzędnika *pięcioma* występuje tylko 4 razy, forma narzędnika *pięciu* także

4 razy, a forma miejscownika *pięciu* 13 razy. Dla porównania — biernikowa forma *pięć* (łącząca się z rzeczownikami o słownikowej kategorii rodzaju innego niż męski osobowy) pojawia się aż 149 razy.

Nasuwa się zatem wniosek, iż omówiony tu model nie jest wystarczającym narzędziem do opisania i wyjaśnienia niezwykle skomplikowanych właściwości systemu liczebnikowego w polszczyźnie.

#### 4. Zakończenie

Wydaje się, że dane z SFPW jednoznacznie wykluczają jakąkolwiek synchroniczną zależność między częstością użycia a właściwościami syntaktycznymi liczebników głównych, uniemożliwiają więc uznanie modelu Heinego (1997) za aktywny we współczesnej polszczyźnie. Wskazują też na niewielkie prawdopodobieństwo istnienia takiej zależności w stosunkowo niedalekiej przeszłości. Jeśli przyjmiemy za Heinem (1997), że różnice frekwencyjne ukształtowały kiedyś podstawy dzisiejszego systemu liczebnikowego, musimy uznać, że stało się to bardzo dawno temu i jedynie w wyniku konserwatywności języka efekty tego procesu zachowały się w systemie gramatycznym współczesnej polszczyzny. Historyczne uwarunkowanie składni liczebników przez czynniki frekwencyjne musi być jednak uznane wyłącznie za hipotezę — bez potwierdzenia w znanych nam danych.

#### Literatura cytowana:

- Babby L. H., 1988, „Noun Phrase internal case agreement in Russian”, (w:) M. Barlow, C. A. Ferguson (red.), *Agreement in natural language. Approaches, theories, descriptions*, Stanford, Center for the Study of Language and Information, Stanford University, s. 287–304.
- Chomsky N., 1986, *Knowledge of Language: Its Nature, Origin and Use*, Nowy Jork, Praeger.
- Corbett G. G., 1978, „Universals in the syntax of cardinal numerals”, *Lingua* 46, s. 355–368.
- Doroszewski W. (red.), 1980, *Słownik poprawnej polszczyzny*, wyd. XIV, Warszawa, Państwowe Wydawnictwo Naukowe.
- Franks S., 1995, *Parameters of Slavic Morphosyntax*, Nowy Jork–Oxford, Oxford University Press.
- Giusti G., Leko N., 1995, „Definite and indefinite quantity expressions in Bosnian”, (w:) R. Benacchio, F. Fici, L. Gebert (red.), *Determinatezza e indeterminatezza nelle lingue slave*, Padwa, Unipress, s. 127–145.
- Greenberg J. H., 1978, „Generalizations about numeral systems”, (w:) J. H. Greenberg (red.), *Universals of human language*, Vol. 3, Word structure, Stanford, Stanford University Press.
- Gruszczyński W., Saloni Z., 1978, „Składnia grup liczebnikowych we współczesnym języku polskim”, (w:) Z. Topolińska, R. Laskowski (red.), *Studia gramatyczne II*, Wrocław, Ossolineum, s. 17–42.
- Heine B., 1997, *Cognitive Foundations of Grammar*, Oxford, Oxford University Press.



- Heinz A., 1965, *System przypadkowy języka polskiego*, Kraków, Uniwersytet Jagielloński.
- Huszczka R., 1996, *Honoryfikatywność. Gramatyka. Pragmatyka. Typologia*, Warszawa, Wydawnictwo Akademickie Dialog.
- Huszczka R., 2000, „Jeszcze o honoryfikatywności polskich zaimków osobowych”, *Poradnik Językowy* 6, s. 21–30.
- Ifrah G., 1981, *Histoire universelle des chiffres*, Paris, Editions Seghers.
- Kurcz I., Lewicki A., Sambor J., Szafran K., Woronczak J., 1990, *Słownik frekwencyjny polszczyzny współczesnej*, Kraków, Polska Akademia Nauk, Instytut Języka Polskiego.
- Klemensiewicz Z., Lehr-Splawiński T., Urbańczyk S., 1964, *Gramatyka historyczna języka polskiego*, wydanie II, Warszawa, Państwowe Wydawnictwo Naukowe.
- Laskowski R., 1984, „Liczebnik”, (w:) R. Grzegorzczkova, R. Laskowski, H. Wróbel (red.), *Gramatyka współczesnego języka polskiego. Morfologia*, Warszawa, Państwowe Wydawnictwo Naukowe, s. 283–293.
- Laskowski R., 1991, hasło „Liczebnik”, (w:) S. Urbańczyk (red.), *Encyklopedia języka polskiego*, Wrocław, Ossolineum, s. 186.
- Pisarek W., 1972, *Frekwencja wyrazów w prasie*, Kraków, Ośrodek Badań Prasoznawczych RSW „Prasa”.
- Przepiórkowski A., 1996, „Case assignment in Polish: Towards an HPSG analysis”, (w:) C. Grover, E. Vallduví (red.), *Edinburgh Working Papers in Cognitive Science. Vol. 12: Studies in HPSG*, Edinburgh, University of Edinburgh, s. 191–228.
- Rutkowski P., 2000, „Składnia polskich grup liczebnikowych: próba opisu formalnego”, *Poradnik Językowy* 8, s. 10–28.
- Rutkowski P., 2001, „Numeral Phrases in Polish and Estonian”, (w:) A. Holmer, J.-O. Svantesson, A. Viberg (red.), *Proceedings of the 18th Scandinavian Conference of Linguistics, Vol. 2, Travaux de l'Institut de Linguistique de Lund*, Lund, Lund University Press, s. 181–190.
- Rutkowski P., Szczegot K., 2001, „On the Syntax of Functional Elements: Numerals, Pronouns and Expressions Indicating Approximation”, (w:) A. Przepiórkowski, P. Bański (red.), *Generative Linguistics in Poland: Syntax and Morphosyntax*, Warszawa, IPIPAN, s. 187–196.
- Saloni Z., 1974, „Klasyfikacja gramatyczna leksemów polskich”, *Język Polski* LIV, nr 1, s. 3–13 i nr 2, s. 93–101.
- Saloni Z., 1977, „Kategorie gramatyczne liczebników we współczesnym języku polskim”, (w:) Z. Topolińska (red.), *Studia gramatyczne I*, Wrocław, Ossolineum, s. 145–173.
- Sambor J., 1972, *Słowa i liczby. Zagadnienia językoznawstwa statystycznego*, Wrocław, Ossolineum.
- Szczegot K., 2001, „Dlaczego liczebniki niższe nie są elementami funkcjonalnymi?”, (w:) S. Ostrowska, P. Rutkowski (red.), *Szkice Humanistyczne II*, Warszawa, CI i KFNrD.
- Szober S., 1928, „Trzy piękne córki było nas u matki czyli Formy podmiotu i orzeczenia w zdaniach z podmiotem logicznym, określonym przydawką liczebnikową”, *Język Polski* XIII, s. 97–106.
- Szober S., 1957, *Gramatyka języka polskiego* (wyd. IV), Warszawa, Państwowe Wydawnictwo Naukowe.
- Zgólkowa H., 1983, *Słownictwo współczesnej polszczyzny mówionej. Lista frekwencyjna i rangowa*, Poznań, UAM.
- Zgólkowa H., Bułczyńska K., 1987, *Słownictwo dzieci w wieku przedszkolnym*, Poznań, UAM.