

Wyrazy n-fonemowe i n-sylabowe w mowie dorosłych adresowanej do niemowląt

STANISŁAW MILEWSKI

Zakład Logopedii
Uniwersytet Gdański
Gdańsk

STRESZCZENIE

Z dotychczasowych badań fonostatystycznych wynika, że istnieje korelacja między rozkładem częstości wyrazów o określonej długości fonemowej i sylabowej a przynależnością tekstu do określonej odmiany lub pododmiany stylistycznej języka polskiego. Na fakt występowania większej liczby wyrazów krótszych lub dłuższych w zależności od rodzaju analizowanych tekstów językoznawcy zwracali uwagę już na początku XX wieku, chociaż najczęściej opinie te nie były jeszcze poparte odpowiednimi danymi ilościowymi. Takowe dane zaczęto publikować dopiero w latach siedemdziesiątych minionego stulecia.

W swoim artykule prezentuję wyniki dotyczące rozkładu częstości wyrazów n-fonemowych i n-sylabowych w specyficznym rejestrze uproszczonym, jakim jest mowa dorosłych adresowana do niemowląt (infant-directed speech – IDS). Porównanie częstości występowania wyrazów o określonej długości fonemowej i sylabowej w IDS z danymi uzyskanymi dla innych prób tekstów, m.in. polszczyzny dzieci w wieku przedszkolnym i tekstów współczesnego języka polskiego, pozwala stwierdzić, iż długość wyrazów może być jednym z ważniejszych parametrów pozwalających na potwierdzenie „inności” mowy kierowanej do małych dzieci.

Analizowany materiał językowy (nagrania magnetofonowe) pochodzi od 102 opiekunów niemowląt (matek, ojców i babć). Został

on jednoosobowo odsłuchany, zapisany fonologicznie i wprowadzony do pamięci komputera. W skład wszystkich analizowanych tekstów weszło 74 534 wyrazów (308 918 fonemów).

Słowa kluczowe: fonem, niemowlę, mowa adresowana do niemowląt, język

WPROWADZENIE

Językowe zachowanie człowieka w każdej konkretnej sytuacji komunikacyjnej zależy od wielu parametrów, m.in. nadawcy i odbiorcy wypowiedzi, sytuacji komunikacyjnej i celu, w jakim komunikacja została podjęta (zob. np.: van Dijk (red.), 2001; Bartmiński, 2001; Grabias, 2001). W kontakcie dorosłego z dzieckiem ogniwem wpływającym w istotny sposób na formę komunikacji jest bez wątpienia samo dziecko. Kontaktując się z małym dzieckiem o niepełnej kompetencji językowej, dorosły użytkownik danego języka próbuje uprościć swoje wypowiedzi, uczynić je bardziej zrozumiałymi dla „rozmówcy”. Uproszczenia te widoczne są w poszczególnych podsystemach języka, przede wszystkim zaś w płaszczyźnie fonetycznej, syntaktycznej i leksykalnej. Z powodu różnych poziomów kompetencji językowej relacje między dorosłym a dzieckiem są nierównorzędne¹. Jest to jedna z przyczyn, dla których dorośli porozumiewają się z małymi dziećmi (przede wszystkim w wieku niemowlęcym i wczesnym ponimowlęcym) w sposób różniący

się od komunikacji między osobami dorosłymi. Ponieważ taka pododmiana języka wydaje się uproszczoną wersją języka dorosłych, określa się ją często jako rejestr² uproszczonej (ang. *simplified register*, ros. *упрощенный регистр*) (zob. Bouvet, 1996, s. 99; Ferguson, 1977; van Dijk, 2001, s. 153, 157; Berko Gleason, Bernstein Ratner, 2005, s. 518)³. R. Jakobson (1962) określał nawet tzw. język nianiek jako rodzaj pidżynu, w którym uczestnicy rozmowy przystosowują wzajemnie swoje kody językowe, w wyniku czego powstaje nowy (uproszczony) kod służący porozumiewaniu się. Paralela z pidżynem w odniesieniu do komunikacji dorosły – małe dziecko wydaje się jednak niezbyt fortunna. Przy tworzeniu pidżynu obydwu osobom biorącym udział w akcie językowym zależy na powstaniu wspólnego kodu (znalezieniu „wspólnego języka”), w komunikacji z niekompetentnym komunikacyjnie małym dzieckiem rejestrem uproszczonym posługuje się zaś tylko jego dorosły opiekun (zob. Dil, 1971). Posługujący się pidżynem dążą do optymalnej komunikacji z nastawieniem na funkcję informacyjną języka, która nie jest głównym celem w tzw. języku nianiek. Dorosły rozmawia z dzieckiem/mówi do dziecka nie tylko dlatego, że chce, aby ono go rozumiało. Opiekun małego dziecka dąży przede wszystkim do nawiązania z nim kontaktu, mówi „za dwóch”. Nie wymaga od niego prawidłowego/pełnego rozumienia i adekwatnej reakcji, ponieważ każda reakcja dziecka (lub jej brak) jest odpowiednio interpretowana przez dorosłego rozmówcę.

Istotny jest fakt, iż omawiany typ komunikacji pojawia się wyłącznie w sytuacjach emocjonalnie pozytywnych podczas wzrokowego kontaktu opiekuna z jego małym podopiecznym. Wykorzystywanie w czasie komunikowania się kontaktu wzrokowego pozwala na stosowanie niewerbalnych sposobów porozumiewania się, np. mimiki i gestów. Językowy kontakt z małym dzieckiem w najwcześniejszym okresie jego rozwoju spełnia przede wszystkim funkcję fatyczną⁴. Służy nie do wymiany informacji, lecz do samego obcowania z dzieckiem, podtrzymywania z nim kontaktu, budzenia w dziecku postawy komunika-

cyjnej, czyli chęci bycia odbiorcą i nadawcą wypowiedzi (zob. Szuman, 1961; Ročławski, 1991). Funkcja nawiązywania i podtrzymywania kontaktu szczególnie uwidacznia się w różnych rodzajach tekstów adresowanych do małych dzieci, które można umownie nazwać jako gruchanie (*воркование*) i tekst-komentarz do działania (*текст-комментарий к действию*). W tekstach mówionych w imieniu dziecka (*текст от имени ребенка*), tekstach – „pokaż” (*текст-„покажи”*), oraz tekstach-nakłaniania (*текст-побуждение*) dominuje funkcja informacyjna (zob. Gawriłowa, 2007).

Istnieje dość duże rozchwianie terminologiczne w nazywaniu tego specyficznego typu komunikacji. W poszczególnych językach przypisuje się jej wiele określeń (zob. np. pol. język/mowa nianiek/piastunek/matek, język dziecienny/dziecięcy, mowa pieszczotliwa; ang. *adultese*, *baby talk* (BT), *caregiver/caretaker speech*, *infant-directed speech* (IDS), *child-directed speech* (CDS), *input language*, *motherese*, *parentese*; ros. *воркование*, *материнская речь*, *регистр общения с детьми* (РОД), *стюканье*, *язык нянь*).

Obecnie najczęściej ze względu na przejrzystość używany jest termin *child-directed speech* (CDS). Jako adresaci mieszczą się w nim niemowlęta i starsze dzieci, nie ogranicza on także „rozmówców” dziecka, którymi mogą być rodzice, starsze rodzeństwo, dziadkowie, inni dorośli itp. Termin „mowa skierowana do dziecka” jest akceptowany przez współczesnych polskich psycholingwistów rozwojowych (zob. Kielar-Turska, 2001). Mniej „pojemny” i z tego względu też rzadziej używany jest termin *infant-directed speech*. „Specyficzna” mowa jest bowiem skierowana nie tylko do niemowląt, ale także używana w kontaktach ze starszymi dziećmi. W niektórych pracach termin IDS jest jednak z powodzeniem stosowany (zob. Sundberg, 1998). W swoim artykule będę przede wszystkim posługiwał się terminem *infant-directed speech* (IDS). W tym samym znaczeniu będę czasami stosował także inne terminy, które wyżej wymieniłem.

Zaprezentowanie odpowiednich danych na temat występowania wyrazów n-fonemowych i n-sylabowych w jakimkolwiek korpusie tekstowym⁵ wymaga odpowiedniego (żmudnego i czasochłonnego) przygotowania materiału badawczego na komputerowym nośniku informacji i jego odpowiedniej obróbki statystycznej. W polskich badaniach nad CDS brakuje prac, które w szerszym zakresie przedstawiałyby kwantytatywne (ilościowe) aspekty tej szczególnej odmiany języka⁶. Wydaje się zaś rzeczą oczywistą, że podejście statystyczne do omawianego rodzaju tekstów może w istotnym stopniu przyczynić się do obiektywizacji przedstawianych faktów, pokazania ich takimi, jakie są one naprawdę, niezależnie od przyjętej przez danego badacza teorii naukowej czy też uprzedzeń do wyników badań osiągniętych przez innych. Z takimi badaniami można się już spotkać w literaturze przedmiotu. Znamy już m.in. takie parametry tekstów CDS, jak: intonacja (Garnica, 1977; Li, 2000; Gawriłowa, 2001), długość frazy (Freed, 1981; Gawriłowa, Fiodorowa, 2004; Gawriłowa, 2007), tempo mowy (Gawriłowa, Fiodorowa,

2004), indeks leksykalnej różnorodności (ros. *индекс лексического разнообразия*)⁷, indeks leksykalnej trywialności (*индекс лексической тривиальности*)⁸ (Gawriłowa, Fiodorowa, 2004).

BADANIA WŁASNE

Analiza fonemowa

W swoim artykule chciałbym przedstawić wyniki badań dotyczących długości fonemowej i sylabowej wyrazów występujących w tekstach CDS⁹. Porównanie uzyskanych przeze mnie danych z odpowiednimi wynikami dla innych odmian polszczyzny (zob. tabela 1) pozwoli na stwierdzenie, czy krótkość lub długość wyrazów może być jednym z parametrów pozwalających zaliczyć mowę adresowaną do małych dzieci (niemowląt) do rejestrów uproszczonych. Długość wyrazów można mierzyć w fonemach, sylabach lub literach. Ponieważ w swoim artykule zajmuję się tekstami mówionymi, przedstawię zagadnienie długości wyrazów mierzonych liczbą występujących w nich fonemów i sylab.

Tabela 1. Porównywane w artykule odmiany stylistyczne współczesnego języka polskiego

Korpusy tekstowe	Liczba wyrazów	Liczba fonemów
CDS (<i>child-directed speech</i>)	74 534	308 918
JDP* (język dzieci przedszkolnych)	101722	409 663
WJP** (współczesny język polski)	69 151	363 856
T-OMK** (język ogólny potoczny/mówiony/konwersacyjno-monologowy)	29 610	142 904
T-OPA** (język ogólny literacki/pisany/artystyczny)	15 085	74 247
T-OPND** (język ogólny literacki/pisany/normatywno-dydaktyczny)	16 531	95 999
T-OPN** (język ogólny literacki/pisany/naukowy)	7925	50 706

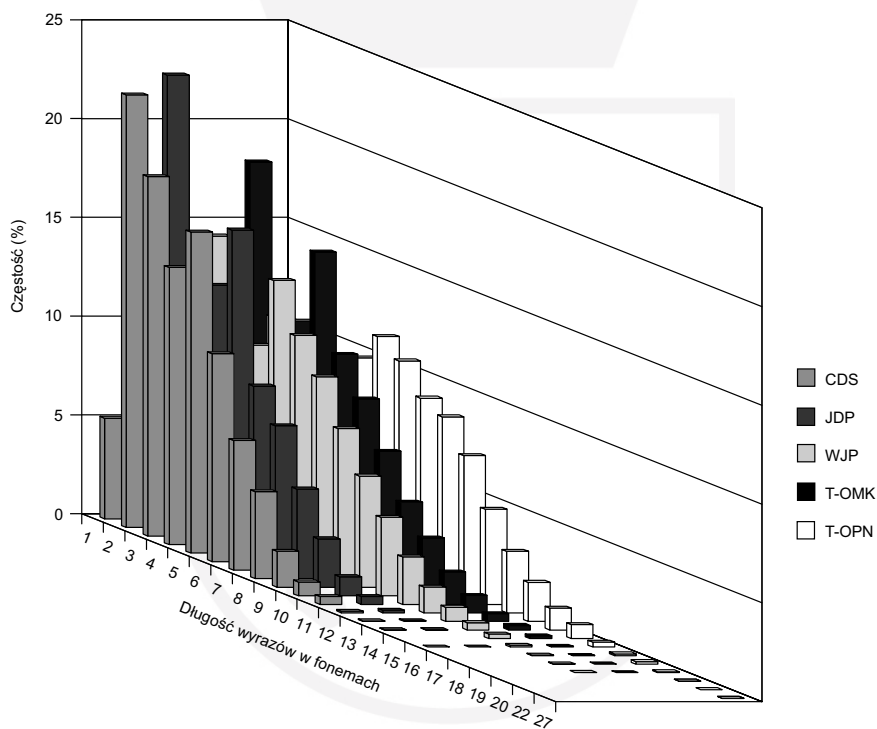
* Dane sporządzone na podstawie słownika K. Zgółkowej, K. Bułczyńskiej (1987)

** Dane z pracy B. Rocławskiego (1981)

Z dotychczasowych badań poświęconych badaniom fonostatystycznym wynika, że istnieje korelacja między rozkładem częstości wyrazów o określonej długości fonologicznej a przynależnością tekstu do określonej odmiany lub poddminy stylistycznej języka polskiego. (zob. Rocławski, 1981, s. 51–71). Na fakt występowania większej liczby wyrazów krótszych lub dłuższych w zależności od rodzaju analizowanych tekstów w pracach językoznawczych zwracano już uwagę wcześniej, chociaż najczęściej opinie te nie były jeszcze poparte odpowiednimi danymi ilościowymi. Warta przytoczenia jest obserwacja N.S. Tru-

bieckiego (1970), który, pisząc o statystyce fonologicznej w swojej klasycznej już dziś pracy *Podstawy fonologii*, stwierdził: „Mowa uczona, nastawiona na wyższy poziom intelektualny słuchaczy, charakteryzuje się długimi wyrazami, podczas gdy opowiadanie proste, nastawione na dość niski poziom słuchaczy, woli wyrazy krótkie” (s. 212). Liczba wchodzących w skład danego wyrazu fonemów bez wątpienia wpływa bowiem na jego „łatwość” bądź „trudność”.

Częstość występowania wyrazów o określonej długości fonologicznej w różnych próbach tekstów ilustruje wykres 1.

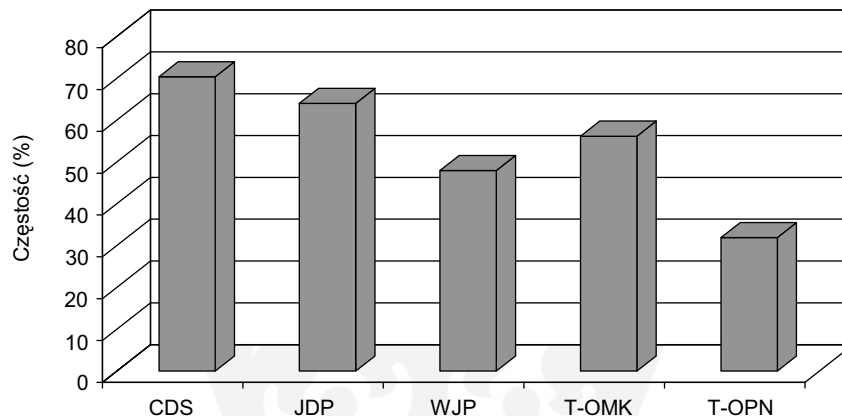


Wykres 1. Frekwencja wyrazów n-fonemowych w tekstach IDS

Wynika z niego, iż w tekstach CDS najwyższą frekwencją charakteryzują się wyrazy dwu-, trój-, pięcio- i czterofonemowe. Łączny odsetek ich wystąpień wynosi 70,35% wszystkich wyrazów n-fonemowych w tej próbie tekstów.

Suma frekwencji tych struktur w innych odmianach polszczyzny jest niższa: w tekstach

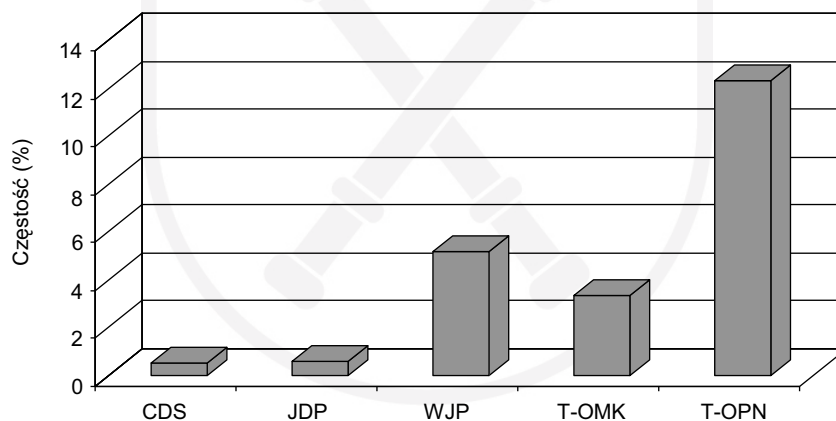
dzieci przedszkolnych (JDP) wyniosła 64,01%, we współczesnym języku polskim (WJP) – 47,94%, tekstach konwersacyjno-monologowych (T-OMK) – 56,13%, tekstach pisanych naukowych (T-OPN) – 31,89% (zob. wykres 2).



Wykres 2. Frekwencja wyrazów dwu-, trój-, cztero-, pięcioletnich w różnych próbach tekstów

Wyrazy dłuższe niż dziesięcioletnie bardzo rzadko pojawiały się w tekstach adresowanych do dzieci. Ich łączna frekwencja wyniosła tylko 0,52% wszystkich odnotowanych wyrazów. W innych odmianach polszczyzny występowanie struktur jedenasto- i więcej fonemowych jest wyższe. W wypadku tekstów dzieci przedszkolnych odpowiedni odsetek

wynosi 0,60%. Dla innych próbek tekstów jest on już znacznie wyższy i wynosi: 5,18% dla WJP, 3,34% dla T-OMK i 12,29% dla T-OPN. W porównaniu z tekstami CDS w tekstach naukowych wyrazy najdłuższe występują więc aż 23,63 razy częściej. Przedstawione wyżej dane ilustruje wykres 3.



Wykres 3. Frekwencja wyrazów jedenastofonemowych i dłuższych w różnych odmianach polszczyzny

Należy podkreślić, iż w tekstach IDS niższa niż w innych odmianach polszczyzny jest także frekwencja wyrazów siedmio-, ośmio-, dziewięcio- i dziesięcioletnich. W świetle tych danych nie znajduje potwierdzenia teza M. Seemana (1962, s. 28), iż w mowie kierowanej do dzieci jest dużo wyrazów długich,

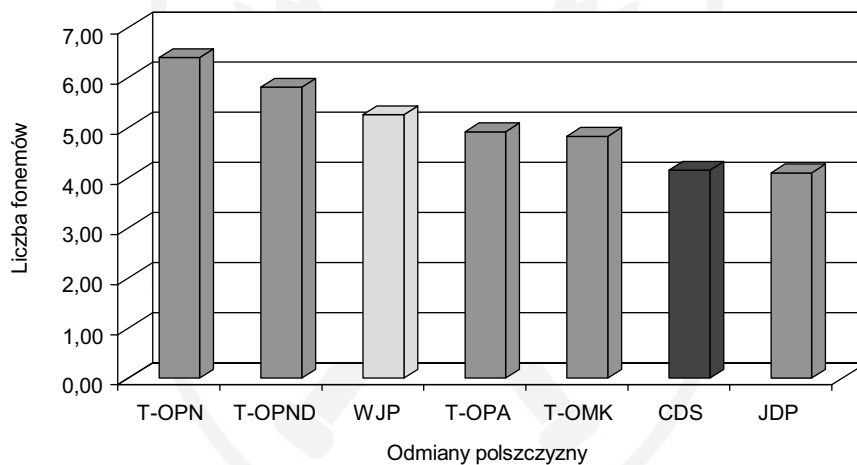
ponieważ wyrazy zdrobniałe, które często w niej występują, są dłuższe niż ich neutralne odpowiedniki¹⁰. Bez wątpienia formacje deminutywno-hipokorystyczne są dłuższe niż odpowiadające im neutralne emocjonalnie odpowiedniki. Jednak występująca w omawianym rodzaju tekstów duża liczba wyrazów

krótkich powoduje, że odsetek wyrazów długich jest stosunkowo niewielki.

Dysponując danymi dotyczącymi liczby wyrazów i fonemów w analizowanych tekstach, można obliczyć średnią długość fonemową wyrazów. Tym parametrem posłużył się B. Rocławski w badaniach nad systemem fonostatystycznym współczesnej polszczyzny. Na podstawie analizy tekstów czterech odmian stylistycznych języka polskiego (T-OPN, T-OPND, T-OPA, T-OMK), stwierdził on, że średnia długość wyrazów we współczesnym języku polskim wynosi 5,26 fonemu/wyraz (Rocławski, 1981, s. 40). Średnie dla innych badanych przez niego odmian polszczyzny wynoszą odpowiednio: T-OPN – 6,40, T-OPND – 5,81, T-OPA – 4,92, T-OMK – 4,83 (zob. Rocławski, 1981, tab. 1.1–1.4). Na podstawie tych danych dochodzi on do konkluzji, iż można „wnioskować o ścisłej korela-

cji między tą średnią a przynależnością tekstu do określonej odmiany stylistycznej” (*ibidem*, s. 40).

W tekstach IDS średnia długość fonemowa wyrazów wyniosła 4,15 fonemu/wyraz, co z pewnością wynika z wysokiej w nich frekwencji wyrazów krótkich (dwo-, trójfonemowych) i niewielkiego udziału w ich budowie wyrazów długich. Właściwość tę w odniesieniu do tekstów mówionych podkreślał we wspomnianych już badaniach B. Rocławski. Z jego wyliczeń wynika, iż wszystkie analizowane przez niego teksty żywych rozmów mają najmniejszą wartość średniej długości wyrazów, która wynosi około 4,4 fonemu/wyraz (zob. Rocławski, 1981, s. 41). W tekstach dzieci przedszkolnych wartość tego parametru jest jeszcze niższa i wynosi 4,03 fonemu/wyraz (zob. wykres 4).



Wykres 4. Średnia długość fonemowa wyrazów w różnych próbach tekstów

Jest rzeczą charakterystyczną, iż w tekstach IDS odnotowano najmniejszy odsetek (5,10%) wyrazów jednofonemowych spośród wszystkich porównywanych odmian polszczyzny (zob. wykres 1). Zdecydowanie więcej wystąpiło wśród nich wyrazów będących samogłoskami (84,61% wszystkich wyrazów jednogłoskowych) niż spółgłoskami (15,39%). Najwięcej jednofonemowych struktur wystąpiło w tekstach dziecięcych (11,24%) i teks-

tach naukowych (8,97%). Ponieważ takimi strukturami są najczęściej spójniki i przyimki (np. *w*, *z*) można wnioskować, iż w tekstach adresowanych do małych dzieci jest ich stosunkowo niewiele.

Najczęściej występującym wyrazem samogłoskowym był fonem /a/, który z częstotnością wystąpień $f = 1400$ stanowi w tekstach CDS 1,89% sumy wszystkich wyrazów w tej odmianie polszczyzny. Częściej występował on

jedynie w języku dzieci przedszkolnych, gdzie jego frekwencja wyniosła 2,54%. W innych odmianach polszczyzny częstość jego występowania wahała się od 0,30% w tekstach drobnych wiadomości prasowych do 1,20% we współczesnej polszczyźnie mówionej. Dane te potwierdzałyby wcześniejsze obserwacje B. Rocławskiego, że *a* pojawia się najczęściej w tekstach języka mówionego, jednak „w miarę im tekst staje się bardziej specjalistyczny, bardziej odległy od tekstów języka potocznego, wartość częstości *a* maleje” (Rocławski, 1981, s. 57–58). We wszystkich odmianach polszczyzny wyraz *a* najczęściej pełni funkcję spójnika. W roli wykrzyknika w tekstach dziecięcych wystąpił tylko w 0,01% przypadków, w *Słowniku frekwencyjnym polszczyzny współczesnej* (Kurcz i in., 1990) odnotowano go w tej funkcji łącznie dla pięciu omawianych tam odmian polszczyzny tylko 21 razy (0,004%). W tekstach adresowanych do dzieci wyraz *a* obok podstawowej roli spójnika pełni jeszcze inne funkcje, np. partykuły wzmacniającej, którą poprzedza się dany wyraz, aby go szczególnie zaakcentować. W analizowanym materiale szczególnie często odnotowywałem konstrukcje typu: *a co...* (110), *a gdzie...* (89), *a teraz...* (83), *a kto...* (65), *a powiedz...* (47).

Stosunkowo wysoka częstość występowania fonemu /e/ jako samodzielnego wyrazu w tekstach adresowanych do niemowląt wynika z jego specyficznej roli, jaką w nich odgrywa. Nie służy on bowiem do wyrażania stanów emocjonalnych nadawcy wypowiedzi, lecz do naśladowania wokalizacji dziecka i zachęcania go do załatwienia potrzeb fizjologicznych.

Głoska /e/, używana przede wszystkim jako zachęta do załatwienia określonej potrzeby fizjologicznej, ma często w języku opiekunów dziecka artykulację (bardziej tylną), która nie jest używana podczas „normalnego” komunikowania się. Jest to już swego rodzaju dźwięk pozasystemowy. Na używanie takich dźwięków podczas kontaktowania się dorosłego z niemowlęciem zwrócił już uwagę w swoich dziennikach Baudouin de Courtenay (1974, s. 25). Znaczenie tzw. dźwięków prymitywnych w kształtowaniu postawy komunikacyjnej małego dziecka podkreśla także

w swoich obserwacjach B. Rocławski (1991, zob. także: Stopa, 1966).

Często także w roli samodzielnego wyrazu w tekstach adresowanych do dzieci występuje fonem /o/. Stanowi on w tym korpusie badawczym 1,19% wszystkich wyrazów, z czego 0,64% przypada na wykrzykniki. W tekstach przedszkolaków jego frekwencja wyniosła 0,63% (0,40% przyimki i 0,23% wykrzykniki). Rzadziej notowano go we współczesnej polszczyźnie mówionej (0,48%), tekstach popularnonaukowych (0,50%) oraz drobnych wiadomościach prasowych (0,60%). W dwóch ostatnich wymienionych odmianach polszczyzny fonem /o/ w roli wykrzyknika wystąpił tylko po jednym razie.

Jako samodzielny wyraz fonem /o/ najczęściej pełni w tekstach funkcję przyimka. W roli wykrzyknika stosunkowo często odnotowywano go w tekstach dziecięcych (0,23%). Najczęściej w funkcji wykrzyknika (częściej niż w roli przyimka) występował on w tekstach skierowanych do małych dzieci i służył przede wszystkim do wyrażania podziwu, zachwytu dla wyglądu lub zachowania dziecka oraz jego najbliższego otoczenia. Często w tym znaczeniu wyodrębniono go z połączeń typu: *o, jak...!* (93); *o, jaki...!* (21); *o, jaka...* (17); *o, jakie...!* (10).

W wielu wypadkach pozytywne emocje nadawcy były wyrażane tylko w formie wykrzyknień, np.: *O, jak ślicznie!* *O, jak ładnie!* *O, jak pięknie!*

Często fonem /o/ pojawiał się w połączeniu z wyrazem *tak* – *O tak!* (62). Wykrzyknik ten najczęściej oznaczał w najszerszym sensie akceptację tego, co zrobiło dziecko, podkreślenie pozytywnego zakończenia wykonania pewnej czynności z nim związanej, np. *No pochlap się w wodzie, Kasiunia. O tak!*

Fonem /i/ jako samodzielny wyraz w tekstach IDS występuje najrzadziej ze wszystkich odmian polszczyzny. Z frekwencją $f = 800$ stanowi ono 1,07% wszystkich wyrazów w analizowanym materiale. Jest to ponad pięciokrotnie mniej niż w polszczyźnie przedszkolaków (5,49%) i ponad dwukrotnie mniej niż we współczesnej polszczyźnie mówionej (2,53%). Słowniki frekwencyjne odnotowują

/i/ wyłącznie w funkcji spójnikowej. W tekstach CDS /i/ obok podstawowej funkcji spójnika pełni także funkcję wzmacniającą, podkreślającą wypowiedź. Szczególnie często notowałem konstrukcje: *no i co* (61), *i co* (16), *i cio!* (5), *i jeszcze* (28), *i teraz* (25), *i już* (18), np.: *No i co powiesz? I co jeszcze niunia powie? I cio 'co' telaś 'teraz' będzie?*

Stosunkowo niska frekwencja spójnika *i* w tekstach adresowanych do dzieci może świadczyć m.in. o rzadkim występowaniu w nich zdań współrzędnych łącznych. Potwierdzałoby to obserwacje innych badaczy, którzy twierdzą, iż w mowie dorosłych podczas obcowania z małymi dziećmi brak lub bardzo niska frekwencja zdań złożonych podrzędnie i współrzędnie jest swoistą prawidłowością (zob. Cejtin, 2000, s. 29).

Bardzo wysoka częstość występowania /i/ w roli spójnika w tekstach dziecięcych może wynikać z faktu, iż materiałem uwzględnionym w analizie najczęściej były opowiadania, w których dzieci, często nieporadnie, „sklejały” swoje wypowiedzi właśnie tym spójnikiem.

Najrzadziej występującym jednofonemowym wyrazem samogłoskowym w tekstach adresowanych do dzieci był fonem /u/ (0,05% wszystkich wyrazów), który pełnił w nich wyłącznie funkcję przyimka, np.: *A u kogo niunia jest na rączkach?*

Podsumowując, można stwierdzić, że na tle innych odmian polszczyzny wyrazy jednofonemowe spółgłoskowe w tekstach CDS występują bardzo rzadko. Trudno jednoznacznie wskazać na przyczyny takiego stanu rzeczy. Z pewnością na omawianym etapie komunikowania się dorosłych z niemowlęciem nie istnieje potrzeba używania dużej liczby wyrazów synsemantycznych (pomocniczych), do których zaliczamy przyimki, szczególną rolę zaś odgrywają wyrazy autosemantyczne, przede wszystkim rzeczowniki i czasowniki. Z badań T.O. Gawriłowej wynika, że teksty ROD we wczesnej fazie rozwoju dziecka (okres niemowlęcy i początkowy ponimowlęcy) są nominocentryczne (ros.

номинацентрические)¹¹, w późniejszym zaś – werbocentryczne¹² (ros. *вербоцентрические*) (Gawriłowa, 2007). Dysponując odpowiednimi danymi, można ustalić dla analizowanych tekstów tzw. indeks substancywności¹³ (ros. *индекс субстантивности*) i indeks werbalności¹⁴ (ros. *индекс глагольности*).

Odsetek wyrazów dwufonemowych występujących w tekstach adresowanych do dzieci (21,89%) był zbliżony do częstości występowania tychże struktur w tekstach dzieci przedszkolnych (22,90%). Stosunkowo wysoki procent omawianych układów wystąpił także w T-OMK (18,49%). Najmniejszy odsetek ich wystąpień odnotowano w tekstach naukowych (8,64%).

Najczęstsze dwufonemowe słowoformy wyekscerpowane z tekstów IDS to: /no/ (2898), /co/ (1765), /ńe/ (1668), /şe/ (1355), /ty/ (802), /to/ (777), /tu/ (609), /do/ (542), /na/ (534), /ma/ (524). Ich łączna frekwencja wyniosła 70,32% wszystkich wystąpień wyrazów dwufonemowych w omawianym rodzaju tekstów.

Najczęściej występujący wyraz dwufonemowy – partykułę *no* – najwięcej razy odnotowano w połączeniach: *no co* (338), *no no* (219), *no powiedz* (172), *no tak* (155), *no już* (92), *no chodź* (83), *no jeszcze* (78), *no ślicznie* (62), *no proszę* (50). W wypadku zaimka *co* oraz partykuły *nie* takimi konstrukcjami są: *co to* (238), *co ty* (138), *co tam* (101), *co się* (89), *co tu* (45), *co tak* (37), *co jeszcze* (29), *co powiesz* (26), *co teraz* (21); *nie ma* (180), *nie chcesz* (83), *nie wolno* (70), *nie chce* (62), *nie będzie* (21).

Wyrazów trójfonemowych w tekstach CDS odnotowano najwięcej (18,19% wszystkich wyodrębnionych). Zbliżony odsetek ich wystąpień stwierdzono w tekstach dziecięcych (11,64%) oraz T-OMK (11,10%). Najmniej takich struktur odnotowano w tekstach naukowych (5,53%).

W obrębie dziesięciu najczęściej występujących w tekstach CDS struktur trójfonemowych znalazły się: *tak* (3385), *jak* (730), *już* (697), *gdzie* (608), *chodź* (556), *kto* (487), *daj*

(422), *jej* (385), *ale* (384), *masz* (281). Suma ich wystąpień stanowi 58,52% wszystkich układów trójfonemowych w analizowanym materiale.

Wyrazy czterofonemowe należą także do struktur najczęściej występujących w tekstach adresowanych do dzieci (14,02%). Tylko o 0,90% rzadziej występują one w tekstach dziecięcych i aż o 6,18% w tekstach naukowych. Wśród dziesięciu najczęściej odnotowywanych słowoform w analizowanym materiale znalazły się: *jest* (653), *chcesz* (270), *mama* (242), *gili* (224), *jaka* (216), *moje* (200), *taki* (185), *robi* (183), *kosi* (179), *było* (171). Ich łączna frekwencja wyniosła 24,14% wszystkich odnotowanych tego typu struktur.

Odsetek wystąpień struktur pięcioletnich w tekstach skierowanych do dzieci oraz tekstach dziecięcych jest prawie identyczny oraz wynosi odpowiednio 16,25% i 16,35%. Jest on także stosunkowo wysoki w T-OMK (15,23%) i, podobnie jak w poprzednio omówionych strukturach, najniższy w tekstach naukowych (9,88%).

Do najczęściej odnotowywanych w tekstach CDS pięcioletnich struktur można zaliczyć: *jeszcze* (575), *teraz* (501), *ładnie* (408), *zobacz* (354), *będzie* (247), *proszę* (242), *dawaj* (241), *tutaj* (224), *jejku* (215), *dobrze* (186). Ich łączna frekwencja wynosi 26,58% wszystkich wyrazów omawianej długości fonemowej.

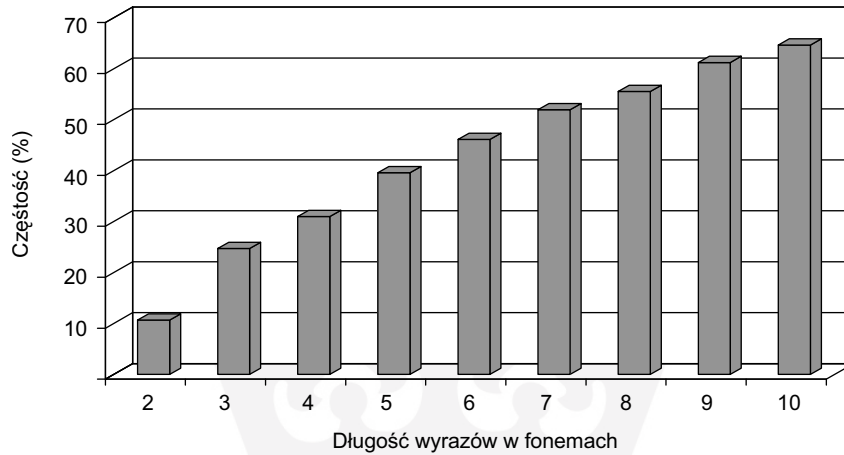
W porównaniu z częstością wystąpień struktur pięcioletnich frekwencja układów sześćfioletnich jest znacznie niższa i wynosi 10,52% wszystkich wyekscerpowanych wyrazów z tekstów CDS. Rzadziej niż w tej próbie tekstów omawiane połączenia występują tylko w polszczyźnie przedszkola-

ków (8,88%). Prawie identyczny odsetek ich wystąpień (10,48%) odnotowano w T-OMK, częściej zaś ich obecność stwierdzano w tekstach naukowych (11,39%) i współczesnym języku polskim (11,44%).

Wśród najczęściej ekscerpowanych sześćfioletnich słowoform w analizowanym materiale znalazły się: *powiedz* (592), *ślicznie* (469), *mamusi* (379), *mamusia* (294), *jestes* (276), *rączki* (141), *jestem* (136), *ojajku* (107), *kochana* (98), *będiesz* (94). Ich łączna frekwencja wyniosła 32,98% wszystkich odnotowanych układów o omawianej długości fonemowej.

Począwszy od wyrazów siedmiofioletnich, frekwencja wszystkich wyrazów n-fonemowych w tekstach CDS jest niższa niż w innych odmianach polszczyzny.

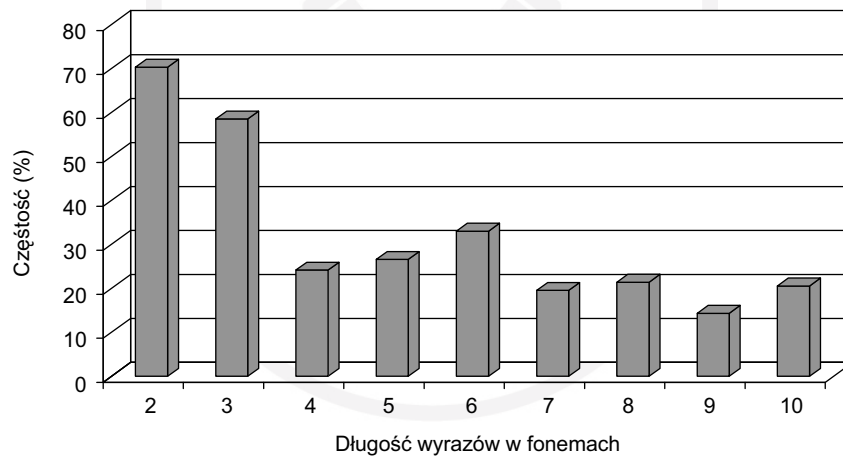
Wśród odnotowywanych w obrębie poszczególnych struktur n-fonemowych słowoform były takie, które występowały wielokrotnie (nawet kilka tysięcy razy), oraz takie, które odnotowywano tylko jeden raz. Zauważyłem pewną prawidłowość między długością fonemową wyrazów a liczbą wystąpień słowoform o frekwencji $f = 1$ (zob. wykres 5). Otóż liczba wystąpień tych struktur jest wprost proporcjonalna do długości fonemowej wyrazów, tzn. im dana grupa wyrazów zbudowana jest z większej liczby elementów, tym więcej „singli” funkcjonuje w ich obrębie. Na przykład struktur o frekwencji $f = 1$ w grupie wyrazów dwufonemowych odnotowano tylko 10,69%, w obrębie wyrazów dziesięcioletnich ich odsetek wynosił już 64,68%. Łącznie w obrębie wszystkich słowoform dwu- do dziesięcioletnich odnotowano 45,84% struktur, które pojawiły się tylko jeden raz (zob. wykres 5).



Wykres 5. Częstość występowania słowoform z frekwencją $f = 1$ w wyrazach n-fonemowych

Takiej jednoznacznej zależności nie widać pomiędzy frekwencją dziesięciu najczęściej występujących słowoform a długością fonemową wyrazów. Największy odsetek wszystkich odnotowanych w danym typie wyrazów n-fonemowych stanowi pierwszych dziesięć

słowoform z listy wyrazów dwufonemowych (70,32%) oraz trójfonemowych (58,52%). W pozostałych wyrazach odsetek ten waha się od 13,33% dla wyrazów dziewięcioletnich do 33,01% dla wyrazów sześciofonemowych (zob. wykres 6).



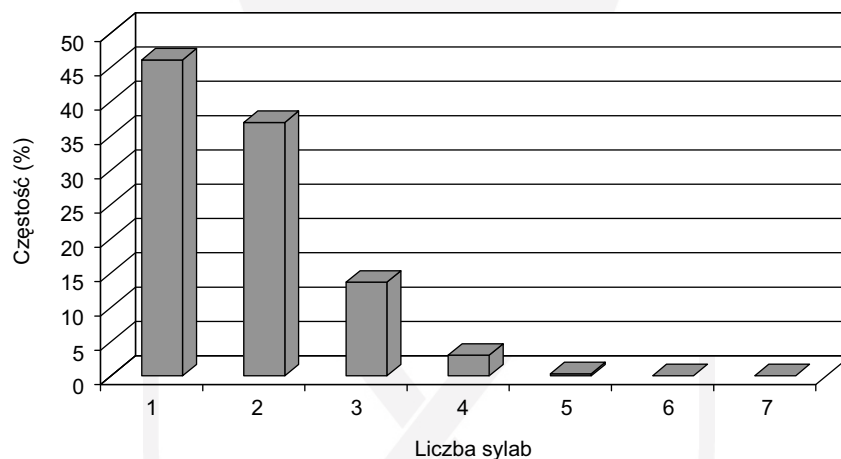
Wykres 6. Częstość występowania dziesięciu najczęstszych słowoform w obrębie poszczególnych struktur n-fonemowych

ANALIZA SYLABOWA

Długość wyrazów można także mierzyć liczbą sylab, które wchodzą w skład określonej jednostki leksykalnej. Dysponując liczbą wyrazów oraz liczbą sylab w danym korpusie tekstowym, można zaś podać wartość średniej długości sylabicznej wyrazów. Uzyskanie tej wartości jest zabiegiem nieskomplikowanym technicznie. Ponieważ w języku polskim ośrodkiem każdej sylaby jest samogłoska, liczba samogłosek odnotowanych w tekstach określa jednocześnie liczbę sylab w nich występujących. W wypadku badanych tekstów

osób dorosłych adresowanych do dzieci wynosi ona $f = 128\ 779$.

W analizowanych tekstach CDS najczęściej pojawiały się struktury jednosylabowe, stanowiące prawie połowę (46,06%) wszystkich odnotowanych wyrazów. Struktury jedno- i dwusylabowe stanowiły już 82,99% wszystkich wyrazów występujących w tej próbie tekstów. Po dodaniu do nich odsetka wystąpień układów trójsylabowych ich ogólny procent wystąpień wynosi aż 96,65%. Częstość wystąpień wszystkich odnotowanych wyrazów n-sylabowych w tekstach CDS ilustruje wykres 7.



Wykres 7. Frekwencja wyrazów n-sylabowych w tekstach CDS

W dotychczasowych badaniach stylistycznych, w których wykorzystywano średnią długość jednostek językowych, sylaba była bardzo często jednostką służącą do mierzenia ich długości. Z badań przeprowadzonych na tekstach języka pisanego artystycznego wynika, że średnia długość sylabowa wyrazów waha się w nich od 1,842 do 2,064 sylaby/wyraz (zob. Rocławski, 1981). Wartość tej średniej dla tekstów CDS wynosi 1,747 sylaby/wyraz. W tekstach dziecięcych zaś parametr ten przybiera wartość 1,706. W obu wypadkach średnie te świadczą o stosunkowo dużej liczbie występujących w nich wyrazów krótkich.

Wyrazy zbudowane z mniejszej liczby sylab są bez wątpienia łatwiejsze do przyswojenia przez dziecko. Interesujące spostrzeżenia na ten temat poczynił J.T. Kania (1982, s. 183; zob. także: Markowa, 1969; Safonika, 2003).

Zaprezentowane dane dotyczące wyrazów n-fonemowych i n-sylabowych w tekstach CDS wyraźnie wskazują, że długość wyrazów może być jednym z ważniejszych parametrów, pozwalających na stwierdzenie, że mowa dorosłych adresowana do niemowlęcia jest swoistym rodzajem rejestru uproszczonego.

PRZYPISY

¹ Stanisław Grabias rozpatruje omawiany typ komunikacji w kontekście odgrywania ról społecznych, polegających „na doborze środków językowych stosowanych do możliwości odbiorcy oraz do funkcji, jaką pełni on w społeczeństwie” (Grabias, 2001, s. 268).

² Przez pojęcie rejestru można rozumieć „wariant językowy używany w pewnych kontekstach społecznych i w odniesieniu do określonej grupy słuchaczy, np. mowa kierowana do dziecka lub obcokrajowca” (Berko Gleason, Bernstein Ratner, 2005, s. 518; zob. także: Grabias, 2001, s. 120).

³ Tego typu rejestr używany jest także w kontakcie z osobami, które, zdaniem nadawcy, z różnych powodów mają ograniczone zdolności rozumienia „normalnej” mowy dorosłych, np. w rozmowie z osobami upośledzonymi umysłowo, ciężko chorymi pacjentami, niedosłyszącymi, obcokrajowcami niedostatecznie znającymi język rozmówcy (ang. *foreigner talk*, ros. *регистр общения с иностранцами – РОИ*). Niektóre elementy tego rejestru używane są także w rozmowach zakochanych (*lover talk*) i podczas „rozmów” ze zwierzętami domowymi (zob. Ferguson, 1977; Jermakowa, 1988, Serpell, 1999; Fiedorowa, 2001, 2002; Worsowicz, 2003) itp.

⁴ Funkcja fatyczna (gr. *phatós* ‘powiedziany’; ang. ‘*phatic function*’ of language, ros. ф. *контакто-установливающая*) jest funkcją języka nastawioną na kontakt. Według R. Jakobsona jest to „pierwsza funkcja słowna, jaka opanowują dzieci: zdradzają one skłonność do komunikowania się, zanim są zdolne nadawać lub odbierać komunikat zawierający informację” (Jakobson, 1989, t. 1, s. 85). Przechodzenie od fatycznej komunikacji do informacyjnej jawi się jako główny czynnik wpływający na zmianę struktury tekstów CDS (*child-directed speech*).

⁵ Przez pojęcie korpusu rozumie się „zestaw tekstów językowych zebrany w celu badania jego systemu lub podsystemu. [...] Ogromny rozwój korpusów tekstowych był związany z zastosowaniem do badań językowych metod statystycznych [...]. Obecnie przez korpus do badań językowych rozumie się prawie wyłącznie zbiór tekstów udostępnionych na komputerowym nośniku informacji (Polański (red.), 1999)”.

⁶ W nurcie lingwistyki kwantytatywnej mieści się monografia S. Milewskiego (2004), w której zawarta jest część przedstawionych w tym artykule wyników badań.

⁷ Mierzy się go stosunkiem liczby haseł występujących w danym korpusie tekstowym do liczby występujących w nim słowoform. Indeks ten pokazuje, na ile „bogata” lub „uboga” jest leksyka w danym typie tekstów. Im wyższa wartość indeksu, tym bardziej różnorodny leksykalnie jest analizowany tekst. Dla rosyjskiego ROD wskaźnik ten wynosi 0,33, dla języka mówionego (dialog) dorosłych odpowiednia wartość wynosi 0,54 (zob. Gawriłowa, Fiedorowa, 2004). Niska wartość tego indeksu w ROD może być m.in. wynikiem spotykanych w tym rodzaju tekstów licznych powtórzeń.

⁸ Indeks leksykalnej trywialności (*индекс лексической тривиальности*) oblicza się jako stosunek sumy frekwencji słowoform w tekście (użyć wyrazów) do ich ogólnej liczby wystąpień. Indeks ten dla tekstów ROD jest dość niski (157), co świadczy o tym, że leksyki używanej w tym typie tekstów nie można zaliczyć do „pospolitej”. Paradoks ten nie ukazuje oczywiście bogactwa leksyki mowy adresowanej do małych dzieci, lecz specyfikę pojawiających się w nim leksemów, przede wszystkim zdrobnień i spieszceń. Na pozorne bogactwo leksykalne w ROD ma wpływ także szczególnie rodzaj językowej gry, w czasie której dorosły tworzy różnorodne formy (od asemantycznych połączeń głosek po okazjonalne neologizmy) nieakceptowane przez normę językową.

⁹ Analizuję dane językowe uzyskane od 102 opiekunów dzieci w pierwszym roku życia. Materiał językowy został nagrany na taśmę magnetofonową, jednoosobowo odsłuchany i zapisany fonologicznie. Łącznie cały korpus tekstowy składał się z 74 534 słowoform (308 918 fonemów). Dokładne dane na temat gromadzenia i analizy materiału znajdzie Czytelnik w mojej pracy (Milewski, 2004).

¹⁰ Podobną opinię zaprezentował W. Mańczak: „[...] nie można bezkarnie nadużywać zdrobnień, gdyż na skutek takiego postępowania tekst się niepotrzebnie wzdłuża” (1980).

¹¹ *Nomen* – tradycyjnie (od starożytności) – leksem odmieniający się przez przypadki, tj. rzeczownik, przymiotnik, zaimek i liczebnik; łac. ‘imię’ (zob. Polański (red.), 1999).

¹² *Verbum* – łac. ‘czasownik’.

¹³ Jego wartość jest stosunkiem liczby występujących w danym korpusie tekstowym liczby haseł rzeczowników do ogólnej liczby wszystkich części mowy występujących w tekstach. Dla rosyjskich tekstów

ROD waha się on w granicach 0,26–0,29, podczas gdy w języku mówionym wynosi 0,16 dla tekstów monologowych i 0,13 dla tekstów dialogowych (zob. Gawriłowa, 2007).

¹⁴ Jego wartość jest stosunkiem liczby występujących w danym korpusie tekstowym liczby haseł czasowników do ogólnej liczby wszystkich części mowy występujących w tekstach. Dla gruchania, komentarzy i tekstów – „pokaż” jego wartość równa się 0,15, dla tekstów – w imieniu dziecka i nakłaniania – 0,21 (zob. Gawriłowa, 2007).

BIBLIOGRAFIA

- Bartmiński J. (2001), Styl potoczny [w:] J. Bartmiński (red.), *Współczesny język polski*, 115–134. Lublin: Wydawnictwo UMCS.
- Baudouin de Courtenay J.N., (1974), *Spostrzeżenia nad językiem dziecka* (wybór i opracowanie M. Chmura-Klekotowa). Wrocław–Warszawa–Kraków–Gdańsk: Ossolineum.
- Berko Gleason J., Bernstein Ratner N.B. (2005), *Psycholingwistyka*. Gdańsk: Gdańskie Wydawnictwo Oświatowe.
- Bouvet D. (1996), *Mowa dziecka. Wychowanie dwujęzyczne dziecka niesłyszącego*, tłum. z jęz. francuskiego R. Gałkowski. Warszawa: WSiP.
- Cejtlin S.N. (2000), *Język i riebionok. Lingwistyka dietskiej rieczki*. Moskwa: Gumanitarnyj Izdatielskij Centr „WŁADOS”.
- Dijk van T.A. (red.), (2001), *Dyskurs jako struktura i proces*, tłum. G. Grochowski. Warszawa: Wydawnictwo Naukowe PWN.
- Dil A. (1971), Bengali Baby Talk. *Word*, 27, 11–27.
- Ferguson C.A. (1977), Baby talk as a simplified register [w:] C.E. Snow, C.A. Ferguson (red.), *Talking to children: language input and acquisition*, 209–235. Cambridge: Cambridge University Press.
- Fiedorowa K.S. (2001), Riecziewoje powiedienije nositiela ruskogo języka w sytuacji obszczenija s inostrancem: strategii organizacii dyskursu. *Wiesticnik molodych uczionych*, wyp. 4, 21–25.
- Fiedorowa K.S. (2002), Leksika w ruskom registrie obszczenija s inostrancami. *Antropologia. Folkloristika. Lingwistyka*, wyp. 2, 409–433, Sankt-Petersburg: Izdatielstwo Jewropejskiego uniwiersitieta.
- Freed B.F. (1981), Foreigner talk, baby talk, native talk. *International Journal of the Sociology Language*, 28, 19–39.
- Garnica O.K. (1977), Some prosodic and paradigmatic features of speech to young children [w:] C.E. Snow, C.A. Ferguson (red.), *Talking to children: language input and acquisition*, 63–88. Cambridge: Cambridge University Press.
- Gawriłowa T.O. (2001), Registr obszczenija s dietmi (baby talk): niekatoryje osobiennosti intonacii. *Antropologia. Folkloristika. Lingwistyka*, wyp. 1, 227–238, Sankt-Petersburg: Izdatielstwo Jewropejskiego uniwiersitieta.
- Gawriłowa T.O. (2007), Typy tiekstow w registrie obszczenija z dietmi [w:] S.N. Cejtin (red.), *Siemantyczeskije kategorii w dietskiej rieczki*, 384–402. Sankt-Petersburg: Nestor-Istoria.
- Gawriłowa T.O., Fiedorowa K.S. (2004), Mnimoje schodstwo i schodnaja mnimost: registr obszczenija z innostrancami vs. registr obszczenija s dietmi w ruskom jazykie [w:] *Trudy fakultieta etnologii*, wyp. 2, 203–220. Sankt-Petersburg, Jewropejskij uniwiersitet w Sankt-Pietierburgie.
- Grabias S. (2001), *Język w zachowaniach społecznych*. Lublin: Wydawnictwo UMCS.
- Jakobson R. (1962), Why ‘Mama’ and ‘Papa’? [w:] Jakobson R., *Selected writings*, t. 1, 538–545, Hague: Mouton.
- Jakobson R. (1989), Poetyka w świetle językoznawstwa [w:] R. Jakobson, *W poszukiwaniu istoty języka*, t. II, wyb. i red. M. R. Mayenowa, 77–124. Warszawa: PIW.
- Jermakowa O.P. (1988), Razgowory z ziwotnymi (linowo-psichologiczeskije zamietki) [w:] D.N. Szmielew, J.A. Ziemska (red.), *Raznowidnosti gorodskoj ustnoj rieczki. Sbornik naucznych trudow*, 240–247. Moskwa: Nauka.
- Kania J.T. (1982), *Szkice logopedyczne*. Warszawa: WSiP.

- Kielar-Turska M. (2001), Psychologiczne i psycholingwistyczne badania nad mową dziecka. Retrospekcja i obszary aktualnych badań [w:] S. Grabias (red.), *Zaburzenia mowy*, t. 1, 60–83. Lublin, Wydawnictwo UMCS.
- Kurcz I., Lewicki A., Sambor J., Szafran K., Woronczak J. (1990), *Słownik frekwencyjny polszczyzny współczesnej*, t. 1–2. Kraków: PAN Instytut Języka Polskiego.
- Li M.N. (2000), Wzajemnej stronie materii i riebionka. Intonacyjny aspekt (wozrast 1,3–1,8) [w:] *Psycholingwistyka i problemy dijetskiej rieczzi; Materiały Rossijskiej naučnoj konfierencji*, 50–52. Czerepowiec: Izdatielstwo Czieriepowieckiego uniwersitieta.
- Mańczak W. (1980), Ze zjawisk współczesnego języka polskiego – moda na zdrobienia, *Język Polski*, 1, 69–71.
- Markowa A.K. (1969), Owładzenie słogowym sossawom słowa w ranniem wozrastie, *Woprosy Psichologii*, 5, 118–125.
- Milewski S. (2004), *Mowa dorosłych kierowana do niemowląt. Studium fonostatystyczno-fonotaktyczne*. Gdańsk: Wydawnictwo UG.
- Polański K. (red.), (1999), *Encyklopedia językoznawstwa ogólnego*. Wyd. II poprawione i uzupełnione. Wrocław–Warszawa–Kraków: Ossolineum.
- Rocławski B. (1981), *System fonostatystyczny współczesnego języka polskiego*. Wrocław–Warszawa–Kraków–Gdańsk–Łódź: Ossolineum.
- Rocławski B. (1991), Kształtowanie komunikacyjnych postaw u dzieci w wieku niemowlęcym [w:] B. Rocławski (red.), *Opieka logopedyczna od poczęcia*, 85–89. Gdańsk: Zakład Logopedii UG.
- Safonika N.J., (2003), *Formiowanie priedposyłok stanowlenija słogowej struktury słowa u doszkolnikow piatogo goda žizni s naruszenijami rieczewogo razwitia* (rozprawa doktorska). Moskwa: Moskowskij piedadogiczeskij gosudarstwiennyj uniwersytet.
- Seeman M. (1962), *Rasstrojstwa rieczzi w dijetskom wozrastie*. Moskwa: Medicinskaja Literatura.
- Serpell J. (1999), *W towarzystwie zwierząt. Analiza związków ludzie – zwierzęta*, tłum. A. Alichnewicz, A. Szczęsna. Warszawa: Państwowy Instytut Wydawniczy.
- Stopa R. (1966), Mowa nianiek w Europie i Afryce (Studium historyczno-porównawcze), *Rozprawy Komisji Językowej WTN*, VI, 47–56.
- Sundberg U. (1998), *Mother Tongue – Phonetic aspects of infant-directed speech*. Stockholm: Department of Linguistics Stockholm University.
- Szuman S. (1961), Kształtowanie się kontaktów porozumienia i współdziałania między matką a niemowlęciem. *Pamiętnik XII Zjazdu Pediatrów*, 13–22. Warszawa: PZWL.
- Trubiecki N.S. (1970), *Podstawy fonologii*. Warszawa: PWN.
- Worsowicz M. (2003), Jak rozmawiają właściciele zwierząt w poradni weterynaryjnej [w:] M. Kita (red.), *Porozmawiajmy o rozmowie*, 113–120. Katowice: Wydawnictwo UŚ.
- Zgółkowa H., Bułczyńska K. (1987), *Słownictwo dzieci w wieku przedszkolnym. Listy frekwencyjne*. Poznań: Wydawnictwo Naukowe UAM.