

Ocena zasobu słownictwa u dzieci w wieku przedszkolnym – nowe narzędzie testowe¹

EWA HAMAN

Wydział Psychologii
Uniwersytet Warszawski
Warszawa

KRZYSZTOF FRONCZYK

Wyższa Szkoła Finansów i Zarządzania
Warszawa

ANETA MIĘKISZ

Wydział Psychologii
Uniwersytet Warszawski
Warszawa

STRESZCZENIE

W artykule tym przedstawiamy wyniki pierwszego etapu prac nad Obrazkowym Testem Słownikowym (część I – Rozumienie: OTS-R). Słownik umysłowy leży u podstaw kompetencji językowej i komunikacyjnej (Michnick-Golinkoff i in., 2000; Hall i Waxman, 2004). Zasób słownictwa istotnie wpływa na inne aspekty rozwoju językowego – w tym na rozwój gramatyki (Dale i in., 2000; Dionne i in., 2003). Prezentowane narzędzie służy do oceny rozumienia słów przez polskojęzyczne dzieci w wieku przedszkolnym. W opisywanym etapie badań wzięło udział 351 dzieci w wieku 2–6 lat. Rozumienie słów sprawdzane było poprzez wybór jednego z czterech obrazków na planszy jako reakcję na pytanie o słowo kluczowe (rzeczownik, czasownik lub przymiotnik). Obrazki dystraktory reprezentowały słowa dobrane do słów kluczowych według podobieństwa fonetycznego, semantycznego i tematycznego. Wyniki wskazują na wzrost zasobu słownictwa z wiekiem, lepsze rozumienie czasowników niż rzeczowników i przymiotników, zwiększoną w stosunku do

innych typów błędów proporcję błędów semantycznych. Badanie pozwoliło na przygotowanie nowej wersji narzędzia, która stanowić będzie podstawę badań normalizacyjnych.

Słowa kluczowe: słownictwo, metoda oceniania, wiek przedszkolny

WPROWADZENIE

Słownik umysłowy to dynamiczna struktura lokująca się w pamięci długotrwałej, zawierająca słowa oraz ich charakterystykę semantyczną, gramatyczną i fonetyczną (Aitchison, 2003; Kurcz, 2005). Rozwój słownika umysłowego jest jednym z kluczowych aspektów rozwoju językowego (Clark, 2003; Dionne i in., 2003). W początkach badań psycholingwistycznych nie uważano słów za interesujący przedmiot badań, w przeciwieństwie do zdań, których tworzenie wymaga znajomości reguł gramatycznych, co miało stanowić o istocie języka (Chomsky, 1957). Nie interesowano się nimi, a przecież to właśnie słowa są konieczne do powstania zdań. Obecnie jest już oczywiste,

że kwestia uczenia się słów jest złożonym procesem, którego wyjaśnienie jest niezbędnym elementem teorii psycholingwistycznej (Bloom, 2000; Clark, 1993; Gleitman i Landau, 1994; Michnick-Golinkoff i in., 2000).

Hościowe i jakościowe badania wczesnego rozwoju leksykalnego

Można wyróżnić dwa nurty badań nad rozwojem słownika: dotyczący przede wszystkim zmian ilościowych oraz skoncentrowany na zmianach jakościowych. W pierwszym z nich badacze starają się odpowiedzieć na pytania związane z zasobem słownictwa (*vocabulary size*) na różnych etapach rozwoju. W ramach drugiego – poszukują mechanizmów wyjaśniających przyswajanie znaczeń, tworzenie kategorii leksykalnych, a także sprawdzają złożoność (głębnię – *depth*) wiedzy o słowach (np. umiejętność definiowania znaczeń). Badanie zmian ilościowych w rozwoju leksykalnym to szacowanie wielkości słownika, ustalanie norm wiekowych, ale także określanie tempa przyrostu słownictwa na danym etapie rozwoju. Badania jakościowe wskazują na mechanizmy leżące u podstaw efektywnego uczenia się słów, łączenia ich w kategorie składające się na strukturę całego słownika. Najczęściej badanym przedziałem wiekowym jest okres od wypowiedzania pierwszych słów (ok. pierwszych urodzin) do końca trzeciego roku życia. Koncentracja na tym okresie jest uzasadniona kluczowymi zmianami, jakie zachodzą w tym czasie w języku dziecka: 1) z niemowlęcia staje się ono kompetentnym (choć nie oznacza to kompetencji pełnej – równoważnej kompetencji dorosłych) użytkownikiem języka; 2) od wypowiedzania pojedynczych słów przechodzi do tworzenia złożonych zdań adekwatnych w danym kontekście społecznym i komunikacyjnym. Już w tym okresie można i należy diagnozować ewentualne zagrożenia w rozwoju lub ryzyko ich wystąpienia w przyszłości (SLI: Leonard, 1998/2006; dysleksja: Lytinen i in., 2006). W badaniach ilościowych wczesnego rozwoju leksykalnego przeważają analizy oparte

na kwestionariuszach (inwentarzach) wypełnianych przez rodziców, którzy na gotowej liście słów zaznaczają te znane ich dzieciom. Rozpowszechnienie MacArthur-Bates CDI (Communicative Development Inventory; Fenson i in., 2007) – zarówno w badaniach dzieci angielskojęzycznych, jak i przyswajających inne języki (Caselli i in., 1995, 1999; Bornstein i in., 2004) – umożliwia stosunkowo łatwe przebadanie dużych grup, jak i porównania międzyjęzykowe. W badaniach koncentrujących się na mechanizmach odpowiedzialnych za przyswajanie słów przeważają metody eksperymentalne, dzięki którym w ściśle kontrolowanych warunkach można pokazać, jakie czynniki – zarówno wewnętrzne (np. poznawcze), jak i zewnętrzne (np. właściwości języka, środowisko) – wpływają na efektywne uczenie się słów (np. Markman, 1989; Bloom, 2000; Tomasello, 2003). Badania złożoności wiedzy o słowach i ich znaczeniach często dotyczą dzieci starszych (w wieku szkolnym) i dzięki temu możliwe jest wykorzystanie w tych badaniach umiejętności czytania i pisanie osób badanych, co w przypadku dzieci w wieku przedszkolnym jest niemożliwe.

Rozwój słownika a rozwój gramatyki

Początkowo badania nad rozwojem słownika i rozwojem gramatyki toczyły się względnie niezależnie – zgodnie z tradycyjnym podziałem językoznawczym na te dwa aspekty języka (Berko Gleason i Bernstein Ratner, 1998/2005). Jednakże od dłuższego już czasu zarówno badania oparte na metodach ilościowych (Marchman i Bates, 1994; Dale i in., 2000; Dionne i in. 2003), jak i testujące teorie dotyczące mechanizmów odpowiedzialnych za rozwój leksykalny (Naigles, 1990, 1996) wskazują na silną współzależność obu poziomów analizy (Clark, 1995/2007; Bloom, 1996/2007; Dixon i Marchman, 2007). Wyniki szeroko zakrojonych badań bliźniąt prowadzone w Wielkiej Brytanii skłoniły ich autorów do sformułowania hipotezy zakotwiczenia leksykalnego (*lexical bootstrapping hypothesis*), która rozwój słow-

nika uznaje za niezbywalną podstawę pełnego rozwoju gramatyki (Dale i in., 2000; Dionne i in., 2003).

Związek rozwoju słownika z innymi sprawnościami poznawczymi

Wychodząc poza krąg badań dotyczących ściśle języka, ważne wydają się także prace wskazujące na związki zasobu słownika z rozwojem poznawczym, m.in. zdolnością do kategoryzacji, postrzegania zależności przyczynowo-skutkowych czy naiwną teorią umysłu. Dotyczą one bardzo wczesnych etapów rozwoju (Gopnik, Choi i Baumberger, 1996), ale także pokazują długoterminowe związki pomiędzy wczesnymi etapami rozwojowymi i późniejszymi osiągnięciami dzieci (Watson, Painter i Bornstein, 2001; Marchman i Fernald, 2008). Wiele badań wskazuje na bardzo silny związek zasobu słownictwa z rozwojem umiejętności czytania, a w szczególności czytania ze zrozumieniem (Snow, Griffin i Burns, 2005; Wagner, Muse i Tannenbaum, 2006). Odpowiedni zasób słownictwa jest więc warunkiem powodzenia w nauce szkolnej, opartej w przeważającej mierze na umiejętności czytania. Z tego względu wczesna diagnoza zasobu słownictwa i jego obniżonego poziomu może stanowić ważną wskazówkę do podejmowania wczesnej interwencji i wyrównywania szans jeszcze przed podjęciem nauki szkolnej.

Narzędzia wykorzystywane w badaniach nad rozwojem słownika

Większość przytaczanych powyżej badań była możliwa do przeprowadzenia dzięki istnieniu standaryzowanych technik i narzędzi do pomiaru zasobu słownictwa (np. MacArthur-Bates CDI; Peabody Picture Vocabulary Test – PPVT). Prowadzone były one w pierwszym rzędzie z dziećmi angielskojęzycznymi, ale także z udziałem dzieci przyswajających rodzimy język włoski, hiszpański, koreański, niemiecki, przy zastosowaniu kulturowych adaptacji narzędzi pierwotnie stworzonych dla języka angielskiego. MacArthur-Bates CDI może być stosowany do wieku 30–36 miesięcy (w zależności od języka); kwestionariusz ten ma polską

adaptację (Smoczyńska, 1999). W wieku ok. 3 lat dzieci znają już na ogół zbyt dużo słów, żeby można było posługiwać się efektywnie metodą pośrednią, odwołującą się do wiedzy rodziców na ten temat. Dlatego też niezwykle popularnym sposobem badania zasobu słownictwa u dzieci starszych jest PPVT – obrazkowy test rozumienia słów (PPVT-RIII, Dunn i Dunn, 1997; najnowsze wydanie: PPVT-4, 2007). Szybkość przeprowadzania badania i możliwość odniesienia wyników do norm sprawia, że jest to jedno z najczęściej stosowanych narzędzi do pomiaru zasobu słownictwa biernego. Niestety w badaniach nad rozwojem języka polskiego ciągle nie zostały opracowane standaryzowane narzędzia do pomiaru zasobu słownictwa u dzieci w całym przedziale wieku od 2 do 6 lat (dotyczy to również innych aspektów rozwoju językowego). Kilka istniejących narzędzi umożliwia takie badanie dla krótszych lub innych przedziałów wieku, jednak porównywanie tych wyników w celu jednorodnej obserwacji zmian rozwojowych w zakresie zasobu słownictwa w całym tym okresie jest problematyczne lub wręcz niemożliwe. Inwentarz Rozwoju Mowy i Komunikacji (Smoczyńska, 1999) to polska wersja MacArthur-Bates CDI (Fenson i in., 2007), posiada (nieopublikowane) normy dla wieku od 9 miesięcy do 3 lat. Jego stosowanie w odniesieniu do dzieci starszych jest niemożliwe zarówno ze względu na metodę badawczą (inwentarz jest wypełniany przez rodziców), jak i efekt sufitowy (maksymalne wyniki) u najstarszych dzieci (3-letnich). Istnieje również znormalizowane narzędzie do diagnozy ogólnego rozwoju niemowląt i małych dzieci (od 2 miesięcy do 3 lat): Dziecięca Skala Rozwojowa (Matczak i in., 2007). To znormalizowane narzędzie, które zawiera szereg podskal, w tym podskalę 'słownik', ze względu na cele stawiane przed całym narzędziem również nie może być stosowane wobec dzieci starszych (o typowym rozwoju). Inne narzędzia służące do diagnozy inteligencji i posiadające podskale werbalne lub operujące materiałem werbalnym mają normy tylko dla dzieci starszych (DMI-2 – Matczak, 2001 – od lat 6; WISC-R – Matczak, Piotrowska i Ciarkowska, 1998 – od lat 6). Natomiast skale do oceny rozwoju intelektual-

nego dla dzieci w wieku przedszkolnym mają charakter całkowicie niewerbalny (Skala Leitera – Jaworowska, Matczak i Szustrowa, 1996 – od lat 3 do 15; Skala Dojrzałości Umysłowej Columbia – Burgemeister, Blum i Lorge; wersja polska: Ciechanowicz, 1992 – od 4 do 10 lat). Jedyne znormalizowane narzędzie bezpośrednio służące ocenie zasobu słownictwa – Test Językowy Leksykon – jest testem pisemnym, przeznaczonym dla młodzieży (od 12 do 17 lat; Jurkowski, 1997). Stan ten jest tym bardziej niepokojący, że narzędzia oceniające zasób słownictwa dzieci w wieku przedszkolnym poza badaniami podstawowymi mogą być wykorzystywane w diagnostyce zaburzeń rozwoju językowego, typowaniu grup ryzyka i wczesnej interwencji (Leonard, 1998/2006).

KONSTRUOWANIE NARZĘDZIA DO OCENY ROZUMIENIA SŁÓW W JĘZYKU POLSKIM

Zadanie stworzenia nowego testu oceniającego rozumienie słów w języku polskim zostało pomyślane jako próba wypełnienia luki w dostępnych narzędziach do oceny rozwoju dzieci w wieku przedszkolnym. Prezentujemy przyjętą metodę tworzenia testu oraz wyniki pierwszego etapu badań.

Proces przygotowania testu podzielony został na następujące części:

- A. Ustalenie zasad konstrukcji testu (sposób wyboru słów kluczowych; zasady doboru dystraktorów, określenie stylu ilustracji itd.).
- B. Opracowanie listy słów kluczowych i listy słów dystraktorów.
- C. Przygotowanie materiału ilustracyjnego i plansz do badań.
- D. Przeprowadzenie badań z wykorzystaniem jak największej liczby pozycji testowych.
- E. Opracowanie wersji do normalizacji.

A. Zasady konstrukcji testu rozumienia słów

Ilościowe badanie rozumienia słów (ocena zasobu słownictwa biernego) wymaga bardzo starannego zaprojektowania konstrukcji testu.

Test powinien uwzględniać słowa o różnej charakterystyce, tak by ocena była możliwie najpełniejsza. Dotyczy to zarówno potencjalnej dostępności słów w języku otaczającym dziecko, którą można szacować na podstawie frekwencji słów; kontekstu sytuacyjno-słownego, w jakim się zwykle pojawiają; ich przynależności do domeny semantycznej, części mowy; jak i formy fonetycznej. Wieloaspektowa organizacja słownika umysłowego powoduje, że słowa są powiązane ze sobą według różnych kryteriów i powiązania te wpływają na rozumienie pojedynczych słów (Kurcz, 2005; Reeves, Hirsch-Pasek i Golinkoff, 1998/2005). Test rozumienia słów powinien zatem te kryteria i czynniki uwzględniać. Stąd w konstrukcji testu postanowiliśmy uwzględnić zarówno frekwencję słowa, jak i jego przynależność do kategorii części mowy. Zdecydowaliśmy, że w teście sprawdzane będzie rozumienie słów należących do trzech podstawowych kategorii części mowy, zawierających słowa znaczące (*content words*): rzeczowniki, czasowniki i przymiotniki. Założyliśmy, że pula słów kluczowych powinna być maksymalnie zróżnicowana semantycznie, a słowa dystraktory powinny być dobrane celowo, tak by uwzględnione zostały różne typy relacji, jakie zachodzą między słowami w strukturze słownika umysłowego. Ostatecznie uwzględniony został aspekt fonetyczny (podobieństwo fonetyczne słów), semantyczny (przynależność do tej samej szerszej kategorii semantycznej) oraz kontekstualny (relacja tematyczna pomiędzy słowem kluczowym a dystraktorem).

Test rozumienia słów wymaga zaprezentowania osobom badanym przykładowych obiektów, do jakich słowa się odnoszą. Najprostszą formą takiej prezentacji jest ilustracja. Przyjęta forma prezentacji narzuca z konieczności poważne ograniczenia, bowiem część słów jest niemożliwa do zilustrowania, lub przynajmniej do zaprezentowania na obrazku w sposób dosłowny – dotyczy to przede wszystkim słów abstrakcyjnych. Świadomi tego ograniczenia, staraliśmy się jednak zminimalizować jego wpływ na dobór słów kluczowych i tworzenie pozycji testowych na kilka sposobów, a mianowicie poprzez: (1) dobór słów kluczowych przez

losowanie; (2) ocenę ‘rysowalności’ słów przez sędziów kompetentnych; (3) sformułowanie pytania testowego, tak by mogło odnosić się nie tylko do ilustracji dosłownych. Decyzje te zostaną szczegółowo opisane poniżej.

B. Dobór słów kluczowych i słów dystraktorów

W celu uzyskania puli słów do badania testowego zdecydowano się wykorzystać zlematyzowaną listę frekwencyjną Korpusu Języka Polskiego PWN (2004). Zawiera ona prawie 500 tysięcy różnych słów; ponad 64 miliony użyc. Wykonano losowanie warstwowe (11 warstw, pierwsza złożona z słów najczęstszych, 10 kolejnych wyrównanych pod względem sumy użyc). Zdecydowano się uwzględniać wyłącznie słowa o frekwencji powyżej 44, gdyż mimo starannego przygotowania listy słowa o niższej frekwencji często są niesklasyfikowanymi formami zależnymi bądź zawierają błędy literowe. Z każdej warstwy wylosowano 150 słów (łącznie 1750 słów). Słowa te stanowiły wstępną pulę losową o kontrolowanej frekwencji. 4 sędziów kompetentnych sklasyfikowało wszystkie słowa w puli ze względu na przynależność do kategorii części mowy (rzeczownik, czasownik, przymiotnik, inne – odrzucano wszystkie słowa sklasyfikowane niezgodnie przez sędziów oraz sklasyfikowane jako inne) i specjalistyczność (wyrazy ocenione zgodnie jako specjalistyczne były następnie odrzucane). W ten sposób uzyskano pulę 1112 słów – rzeczowników, czasowników i przymiotników. Następnie dwóch sędziów kompetentnych (zawodowych rysowników) oceniało te słowa pod względem „rysowalności”, tj. możliwości przedstawienia znaczenia danego słowa na rysunku. 601 słów zostało ocenionych zgodnie jako rysowalne, kolejnych 250 słów jeden z sędziów uznał za rysowalne, a pozostałe 261 obaj sędziowie uznali za nierysowalne. Pula 601 słów zgodnie rysowalnych stanowiła podstawę listy słów, dla których dobierane były słowa dystraktory. Do każdego słowa dobrane zostały trzy inne – podobne brzmieniowo, semantycznie i tematycznie. Ograniczenie liczby dystraktorów do trzech pozwoliło na uwzględnienie podstawowych relacji zachodzących między

słowa w strukturze słownika umysłowego, przy minimalizacji obciążenia pamięci roboczej (w procesie percepcji materiału ilustracyjnego) i innych czynników poznawczych, które mogłyby wpływać na odpowiedzi badanych. Dystraktory były dobierane przez pierwszą autorkę artykułu z użyciem różnego rodzaju słowników (Słownik Synonimów, Słownik Języka Polskiego, Indeks a tergo) i listy frekwencyjnej, która w wersji elektronicznej pozwoliła na wyszukiwanie wyrazów różniących się określoną liczbą głosek (w celu określenia możliwych dystraktorów bliskich fonetycznie). Propozycje zestawu dystraktorów były wielokrotnie konsultowane i dopracowywane zarówno w grupie bezpośrednio zaangażowanej w przygotowanie materiałów do badania (seminarium magisterskie), jak i oceniane przez badaczy rozwoju języka dziecka. Założono, że każdy z trzech typów dystraktorów dla danego słowa kluczowego powinien różnić się od niego wyłącznie pod jednym względem (tzn. dystraktor semantyczny nie powinien jednocześnie być podobny brzmieniowo do słów kluczowych itd.) – miało to umożliwić jednoznaczność odpowiedzi poprawnej oraz identyfikację popełnianego błędu. W trakcie realizacji zadania doboru dystraktorów postanowiono, że słowa bliskie brzmieniowo słowom kluczowym same również będą stanowiły słowa kluczowe (a więc dobrano do nich słowa bliskie semantycznie i tematycznie). Pozwoliło to w pewnym stopniu zneutralizować fakt, że wiele par słów podobnie brzmiących z konieczności różniło się między sobą innymi parametrami (np. frekwencją). Ostatecznie ustalono listę 512 zestawów czterech słów: słowo kluczowe plus trzy dystraktory (fonetyczny, semantyczny i tematyczny), i dla nich przygotowywano kolorowe ilustracje.

C. Przygotowanie materiału ilustracyjnego i 4-obrazkowych plansz dla pozycji testowych

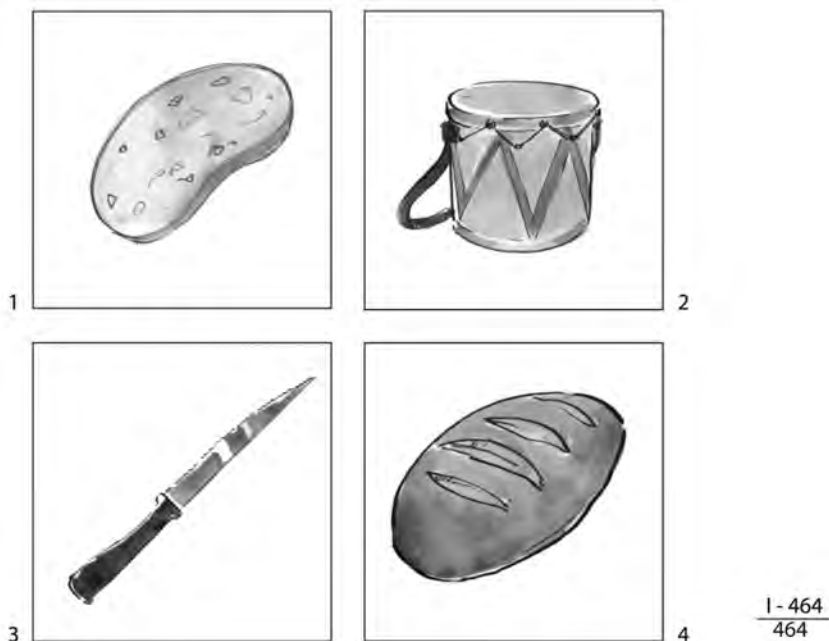
Wszystkie ilustracje były przygotowane przez jedną osobę – artystkę z dużym doświadczeniem w dziedzinie ilustracji dla dzieci oraz z wcześniejszym doświadczeniem w przygotowywaniu materiałów do badań naukowych. Autorka znała zasadę konstrukcji testu (zależność między

poszczególnymi elementami każdej pozycji testowej) i otrzymała listy słów do zilustrowania według zestawów. Była szczegółowo poinformowana o konieczności jednoznacznego i neutralnego przedstawiania znaczeń słów. Jednocześnie podjęto decyzję, aby przygotowywane obrazki zachowywały charakter wysokiej jakości ilustracji książkowych dla dzieci, co powinno sprzyjać zainteresowaniu nimi i ułatwiać utrzymanie uwagi na wykonywanym zadaniu. Zestawy ilustracji były oceniane pod względem jednoznaczności, wyrazistości i adekwatności przez zespół przygotowujący materiał badawczy, wspierany, również na tym etapie, przez badaczy rozwoju języka. W razie negatywnej oceny lub nawet wątpliwości, autorka ilustracji była proszona o przygotowanie wersji poprawionej według podanych wskazówek. Każdy obrazek był malowany techniką akwarelową, a następnie skanowany. Dzięki temu wszystkie obrazki zostały zarchiwizowane w postaci elektronicznej, co znacznie ułatwiło przygotowanie plansz do badania testowego. Ostatecznie do

utworzenia pierwszej wersji zestawu plansz do badania testowego użyto 1413 obrazków (poza powtarzającymi się obrazkami dla słów kluczowych i blisko brzmiących, nieliczne obrazki dla pozostałych dwóch kategorii też mogły się powtarzać na dwóch różnych planszach).

Każda plansza testowa (w rozmiarze A4 i układzie poziomym) zawierała 4 obrazki (odpowiadające słowu kluczowemu i 3 dystryktorom) o rozmiarach 10 × 10 cm rozłożone symetrycznie w odstępach 1 cm. Każdy obrazek miał czarną ramkę dla jednoznacznego wyodrębnienia go od pozostałych. Obrazki na planszy były numerowane (zgodnie z kierunkiem pisma). Każda plansza miała także swój numer umieszczony w prawym dolnym rogu. Przykładowa plansza przedstawiona jest na rycinie 1.

Skład kompletu plansz został przygotowany w programie InDesign i zapisany w formacie PDF. Na każdej 4-obrazkowej planszy obrazki zostały ułożone w kolejności losowej. Ostatecznie w badaniu testowym wykorzystano 497 plansz.



Rycina 1. Przykładowa plansza testowa dla słowa kluczowego ‘bochenek’; dystryktory: fonetyczny – ‘bębenek’, semantyczny – ‘kromka’, tematyczny – ‘nóż’

D. Liczba i układ pozycji testowych w pierwszym etapie badań

Aby test spełniał ostatecznie warunek szybkości i łatwego badania, liczba pozycji testowych powinna być niewielka. Jednocześnie zbyt mała liczba pozycji pogarsza różnicowanie badanych. Ustalenie optymalnej liczby pozycji jest trudne i wymaga badań wstępnych (które tutaj są właśnie opisywane). Docelowo test ma zawierać pozycje ułożone według skali trudności, tak by w zależności od wieku dziecka możliwe było rozpoczynanie badania w odpowiednim miejscu oraz kończenie go w momencie, gdy badane dziecko zacznie popełniać dużą liczbę błędów. Konkretnie kryterium określające moment przerwania badania na podstawie liczby popełnianych błędów powinno być ustalone na podstawie badań wstępnych.

Mimo że wybrane do badania słowa kluczowe były kontrolowane pod względem frekwencji, to jednak uznano, że nie będzie ona w pełni miarodajnym wskaźnikiem trudności pozycji testowych. Lista Frekwencyjna Korpusu Języka Polskiego PWN (2004) określa jedynie częstość występowania słów w języku dorosłych, w przeważającej mierze w tekstach pisanych. W momencie tworzenia testu była to największa dostępna lista frekwencyjna języka ogólnego i to zadecydowało o jej użyciu do losowania słów kluczowych. Specyfika mowy dziecka oraz mowy dorosłych skierowanej do dzieci sprawia, że częstość poszczególnych słów w tych rejestrach może znacznie odbiegać od frekwencji ogólnej, dlatego jednym z celów prezentowanego badania było oszacowanie trudności w stosowaniu słów kluczowych/ pozycji testowych w poszczególnych grupach wiekowych. Z tego względu w opisywanym badaniu wstępnym zdecydowano, że kolejność pozycji testowych będzie losowa. W celu uniknięcia ukrytego efektu porządku przygotowano 4 zestawy o różnych porządkach: (1) układ losowy, (2) odwrócony układ 1. (od końca), (3) układ 1. podzielony na dwie części zamienione miejscami (druga połowa zamieniona kolejnością z pierwszą), (4) układ 2. podzielony na dwie części zamienione miejscami. Ten sposób ustalenia różnych porządków plansz zagwarantował, że dana pozycja będzie pojawiała się

w różnych momentach badania (na początku, na końcu i w środku), a więc uzyskany dla niej wynik sumaryczny nie powinien zależeć od miejsca w całej badanej puli. Miało to istotne znaczenie przede wszystkim ze względu na ogólną liczbę pozycji w całym komplecie i przewidywane zmęczenie badanych dzieci.

Organizacja badania – podział według wieku badanych

Ze względu na bardzo dużą liczbę pozycji testowych (497) zdecydowano, że badanie zostanie podzielone na dwie części – w pierwszej części wzięły udział dzieci w wieku 4–6 lat, które odpowiadały na wszystkie 497 pozycji (opis badanej próby i procedury poniżej). Po przeprowadzeniu tej części badania zanalizowano jej wyniki i przygotowano zmniejszoną pulę pozycji dla dzieci młodszych (2–3-letnich). Z pełnej puli 497 plansz wybrano do badań dla dzieci młodszych (2–3-letnich) pozycje, które miały ogólną sumę poprawnych odpowiedzi >90% (pozycje bardzo proste), oraz pozycje, które spełniały następujące warunki: (1) miały dodatnią moc dyskryminacyjną (obliczaną dla wszystkich wyników łącznie); (2) miały rozkład poprawnych odpowiedzi niemalejący z wiekiem i monotoniczny; (3) nie miały nieciągłości w rozkładzie mocy dyskryminacyjnych obliczonych dla rocznych przedziałów wieku dzieci (tzn. nie miały ujemnej mocy dla środkowego przedziału wiekowego, przy dodatnich w pozostałych grupach); (4) miały nieujemną moc dyskryminacyjną dla 4-latków; (5) miały rozkład błędów zgodny z oczekiwanym opartym na średnim rzeczywistym lub na losowym. Wykluczone zostały natomiast pozycje o ogólnej liczbie poprawnych odpowiedzi <30% – jako potencjalnie bardzo trudne już dla starszych dzieci (lub wadliwie skonstruowane). W wyniku tej procedury selekcyjnej wyłoniono pulę 264 pozycji, które były wykorzystywane do badań dzieci młodszych.

E. Opracowanie wersji do normalizacji

Uczestnicy badania

Badanie przeprowadzono w żłobkach i przedszkolach w Warszawie i okolicach, Gdańsku,

Ciechanowie i Serocku. W analizach uwzględniono wyniki 351 dzieci w wieku 2–6 lat, które ukończyły całe badanie. W każdym rocznym przedziale wiekowym kilkanaścioro dzieci nie ukończyło badania i ich wyniki nie były analizowane. Przyczyną nieukończenia badania była najczęściej nieobecność dziecka przez kolejne 14 dni od pierwszej sesji (badanie było dzielone na kilka sesji, które mogły odbywać się w odstępie do 14 dni od pierwszej sesji). Dziecko miało też prawo odmówić dokończenia badania. W kilku przypadkach ingerencje personelu placówki, w której było przeprowadzane badanie, na tyle zaburzyły jego przebieg, że badania nie kończono. Łącznie przebadano około 420 dzieci.

Szczegółowy rozkład płci w poszczególnych grupach wiekowych przedstawiony został w tabeli 1. Dzieci młodsze (2- i 3-letnie) podzielono na kwartalne przedziały wiekowe. Wyniki dzieci 4-letnich analizowano w podziale na przedziały półroczne. Wyniki 5- i 6-latków były analizowane w rocznych przedziałach. Dzięki temu możliwe było uchwycenie zmian rozwojowych, które w przypadku młodszych dzieci są znacznie bardziej dynamiczne.

Procedura

Przed przeprowadzeniem badania uzyskano zgodę Komisji Rady Wydziału Psychologii UW ds. Etyki Badań Naukowych na jego przeprowadzenie. Rodzice wszystkich badanych dzieci wyrazili pisemną zgodę na udział dziecka w badaniu. Badanie przeprowadzane było na terenie placówki oświatowej, do której uczęszczało dziecko, w odrębnym pomieszczeniu (bez udziału osób trzecich lub w obecności personelu przedszkola/żłobka, poinformowanego wcześniej o konieczności nieingerowania w przebieg badania).

Procedura badawcza była identyczna w przypadku dzieci młodszych i starszych, z zastrzeżeniem, że dzieci młodsze otrzymywały okrojona pulę 264 pozycji.

Dzieci biorące udział w badaniu miały okazję poznać osoby przeprowadzające badanie w grupie przedszkolnej/żłobkowej w trakcie swobodnej zabawy. Badanie miało charakter indywidualny i trwało łącznie ok. 1,5 godziny (dla dzieci starszych) i ok. 40–60 minut (dla dzieci młodszych). W celu dostosowania warunków badania do indywidualnych potrzeb badanych

Tabela 1. Rozkład płci i liczebność grup w poszczególnych przedziałach wiekowych

Przedział wieku	Płeć		Razem
	Chłopcy	Dziewczynki	
2 lata I kwartał	7	6	13
2 lata II kwartał	11	10	21
2 lata III kwartał	12	18	30
2 lata IV kwartał	13	21	34
3 lata I kwartał	9	13	22
3 lata II kwartał	12	10	22
3 lata III kwartał	8	15	23
3 lata IV kwartał	14	20	34
4 lata I półrocze	15	11	26
4 lata II półrocze	12	16	28
5 lat	32	26	58
6 lat	19	21	40
Razem	164	187	351

dzieci było ono podzielone na kilka krótszych sesji. W sytuacji, gdy badane dziecko nie ujawniało oznak zmęczenia, osoba badająca proponowała krótką przerwę po obejrzeniu pierwszych 200 obrazków zestawu (dzieci starsze) / 132 (dzieci młodsze). W takich sytuacjach badanie kontynuowane było po przerwie tego samego dnia lub też w ciągu 14 dni od pierwszej sesji. W razie wcześniejszego okazania zmęczenia lub dekoncentracji prowadzący badanie proponował dziecku przerwę w dowolnym momencie badania. Niektóre (choć nieliczne) dzieci stanowczo odmawiały zrobienia jakiegokolwiek przerwy i wtedy decydowano się na przeprowadzenie badania w jednej sesji. Średnio badanie było przeprowadzane w 2 sesjach (odchylenie standardowe = 1,2), z rozrzutem od 1 do 10 sesji (rekord 10 sesji należał do jednego dziecka 4-letniego, kolejna trójka 2-latków zaliczyła całość w 7 sesjach).

Na wstępie każde dziecko było informowane ogólnie, że w trakcie oglądania wielu kartek z obrazkami jego zadanie za każdym razem będzie polegało na wyborze jednego obrazka, który najlepiej pasuje do tego, co powie eksperymentator. Następnie dziecko rozpoczynało oglądanie plansz testowych. W trakcie badania każdemu dziecku prezentowano zestaw 497/264 (w zależności od wieku) plansz z obrazkami. Przy każdej z plansz osoba badająca pytała dziecko o słowo kluczowe, a jego zadaniem było wskazanie tego obrazka, odpowiadającego słowu, o które pytano. Forma zadawanych pytań zależała od kategorii gramatycznej słowa kluczowego. Gdy właściwym słowem był rzeczownik, zadawano pytanie: „Gdzie jest... (np. kot)?”. O czasowniki pytano: „Kto... (np. pije)?” lub „Co... (np. leży)?”, a o przymiotniki: „Który jest... (np. bogaty)?” lub „Co jest... (np. górnicze)?”. Wszystkie pytania w dokładnym brzmieniu wydrukowane były na arkuszu odpowiedzi. Odpowiedzi dziecka (numer obrazka wskazanego na planszy) były zapisywane przez osobę prowadzącą badanie na tym samym arkuszu.

Dziecko miało możliwość przerywania badania w każdym momencie (było o tym informowane na wstępie), przy każdej sesji było pytane o chęć kontynuacji. Dzieci, które odmawiały kontynuowania, nie były badane do końca.

WYNIKI BADAŃ

Ze względu na to, że badanie było podzielone na dwie części (dzieci starsze i dzieci młodsze) oraz że materiał badawczy w części drugiej zawierał celowo dobraną, mniejszą liczbę pozycji testowych, wyniki obu części będą zaprezentowane oddzielnie – łączna analiza statystyczna wyników albo byłaby obciążona błędem nieuwzględnienia pozycji, które występowały wyłącznie w pierwszej części (gdyby wziąć w analizach pod uwagę jedynie wspólną pulę 264 pozycji), albo musiałaby się opierać na innej liczbie pozycji w grupach młodszych i starszych, co uniemożliwiłoby przeprowadzenie poprawnej analizy wariancji. A zatem wyniki zaprezentowane zostaną w dwóch częściach, według wieku badanych: najpierw wyniki dzieci młodszych (2–3-latków), a następnie starszych (4–6-latków).

Jednym z badanych czynników była przynależność słowa do kategorii części mowy, a liczba pozycji dla każdej z uwzględnianych w badaniu części mowy (rzeczowniki, czasowniki, przymiotniki) nie była jednakowa (zgodnie z pierwotnym losowaniem słów odzwierciedlała ona proporcje występowania rzeczowników, czasowników i przymiotników w polszczyźnie), aby zatem możliwe było porównanie pomiędzy częściami mowy, wszystkie analizy statystyczne przeprowadzone były na danych procentowych.

Analizy dotyczyły zarówno poprawnych odpowiedzi, jak i rodzajów popełnianych błędów. Wszystkie analizy statystyczne zostały przeprowadzone w programie Statistica 8.0.

Analizy poprawności odpowiedzi oraz typów popełnianych błędów dla dzieci młodszych (2–3-latków)

Analiza poprawności odpowiedzi dla dzieci młodszych (przedziały kwartalne) – 264 pozycje

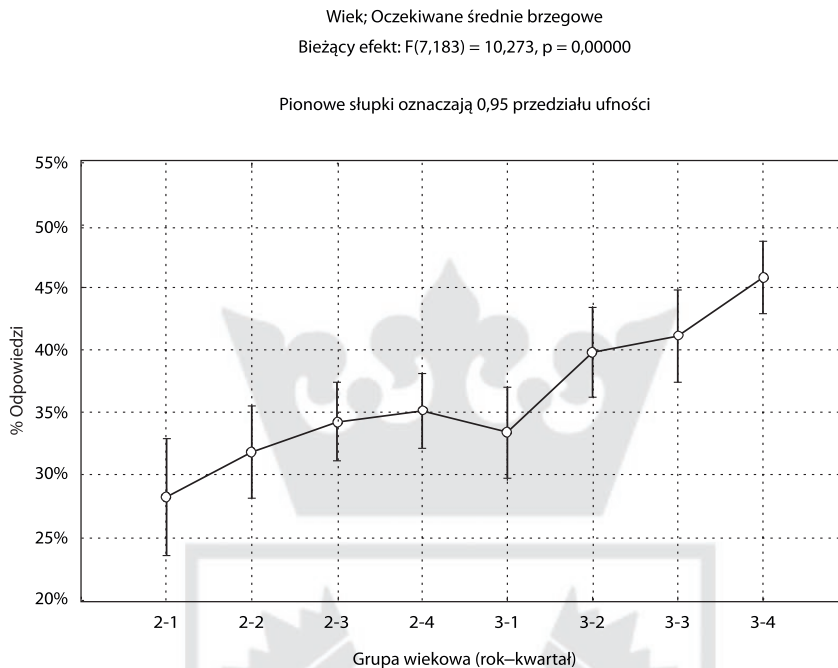
Analizę wariancji przeprowadzono w schemacie $8 \times 2 \times 3$ (wiek \times płeć \times część mowy), z powtarzanym pomiarem (dla części mowy) dla danych procentowych (procent poprawnych odpowiedzi w danej kategorii części mowy).

Tabela 2 przedstawia statystyki opisowe: średnie sumy poprawnych odpowiedzi i odchylenia standardowe dla poszczególnych przedziałów wiekowych i płci. Istotne okazały się efekty główne wieku ($F(7,183) = 10,27$, $p < 0,001$) i części mowy ($F(2,366) = 3,36$, $p < 0,04$), nie było żadnych interakcji. Zmienna 'płeć' nie wpływała w żaden sposób na wyniki – brak efektu głównego ($F(1,183) = 1,09$, $p > 0,29$) i interakcji. Wykres 1 obrazuje wpływ wieku na poprawność odpowiedzi, a wykres 2 wpływ części mowy na poprawność odpowiedzi. Wraz z wiekiem rośnie procent poprawnych odpowiedzi. Zależność ta nie ma charakteru liniowego:

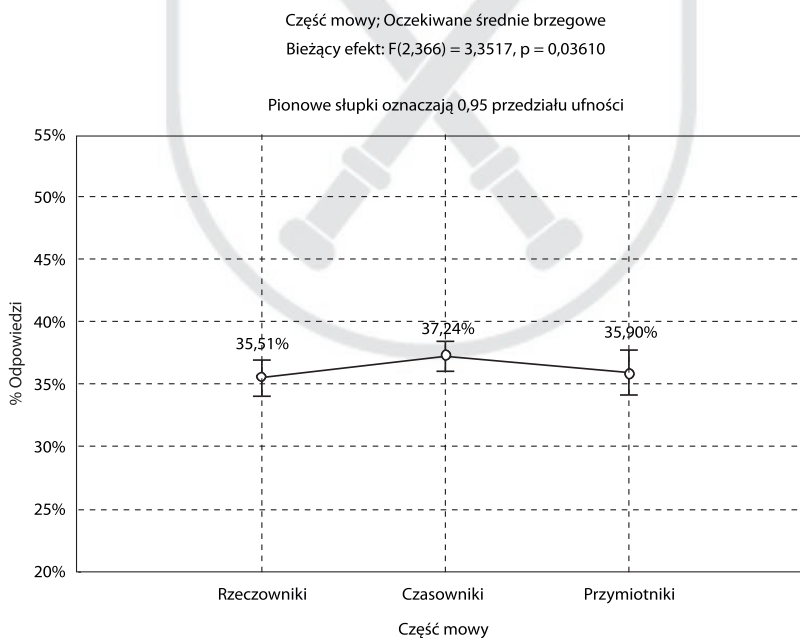
u dzieci 2-letnich różnice pomiędzy każdymi kolejnymi dwoma przedziałami nie są istotne (test Duncana, $p > 0,05$), a u dzieci 3-letnich istotne różnice wystąpiły między przedziałami 3–1 i 3–2 (test Duncana, $p < 0,03$). Dwie pary kolejnych przedziałów osiągnęły różnice bliskie istotnym: 2–1 i 2–2 (test Duncana, $p > 0,14$) oraz 3–3 i 3–4 (test Duncana, $p > 0,11$). Czasowniki okazały się istotnie łatwiejsze od rzeczowników (test NIR Fishera, $p < 0,008$) i marginalnie łatwiejsze od przymiotników (test NIR Fishera, $p < 0,058$). Różnice między rzeczownikami i przymiotnikami nie są istotne.

Tabela 2. Średnie sumy (i odchylenia standardowe) poprawnych odpowiedzi w kwartalnych przedziałach wiekowych dla każdej z płci – dzieci młodsze (C – chłopcy; D – dziewczynki; N – liczba osób w grupie)

Wiek	Płeć	Rzecz. (max = 160) średnia	Rzecz. odch. st.	Czas. (max = 75) średnia	Czas. odch. st.	Przym. (max = 29) średnia	Przym. odch. st	N
2–1	C	43,43	6,68	20,43	3,05	7,57	3,15	7
2–1	D	47,67	4,37	21,67	7,87	8,83	3,43	6
2–2	C	50,91	12,28	23,45	6,11	9,64	3,35	11
2–2	D	49,90	10,55	25,30	6,34	8,70	3,71	10
2–3	C	52,58	7,20	25,58	4,48	8,75	3,31	12
2–3	D	56,50	9,28	26,78	6,42	10,89	2,30	18
2–4	C	57,69	15,36	25,08	5,59	10,31	4,53	13
2–4	D	54,48	11,46	26,67	4,36	10,48	2,75	21
3–1	C	49,22	6,76	25,44	4,13	9,11	2,32	9
3–1	D	51,00	7,22	28,69	3,64	9,92	1,98	13
3–2	C	59,42	16,92	29,33	7,62	10,75	2,83	12
3–2	D	60,90	12,95	33,20	5,47	12,50	3,78	10
3–3	C	64,13	22,83	31,00	5,95	11,50	2,88	8
3–3	D	67,13	23,06	32,87	9,40	11,60	5,53	15
3–4	C	75,43	24,39	38,00	6,79	13,29	3,81	14
3–4	D	68,55	20,15	33,40	7,63	12,75	3,88	20
	Łącznie	58,12	16,82	28,57	7,47	10,68	3,68	199



Wykres 1. Wpływ wieku na poprawność odpowiedzi: analiza wariancji w schemacie wiek \times płeć \times część mowy dla dzieci młodszych



Wykres 2. Wpływ kategorii gramatycznej (części mowy) na poprawność odpowiedzi: analiza wariancji w schemacie wiek \times płeć \times część mowy dla dzieci młodszych

Analiza typów błędów dla dzieci młodszych (przedziały kwartalne) – 264 pozycje

Analiza wariancji w schemacie $8 \times 2 \times 3 \times 3$ (wiek \times płeć \times część mowy \times typ błędu) z powtarzającymi pomiarami (dla czynników ‘część mowy’ i ‘typ błędu’) dla danych procentowych wykazała istotne efekty główne wieku, części mowy i typu błędu. W przypadku zmiennych ‘wiek’ i ‘część mowy’ wyniki były analogiczne w stosunku do analizy poprawnych odpowiedzi: wraz z wiekiem maleje proporcja błędnych odpowiedzi, czasowniki mają mniejszą proporcję błędów niż rzeczowniki i przymiotniki. W przypadku zmiennej ‘typ błędu’ różnica pomiędzy poszczególnymi typami błędów polegała na większej proporcji błędów semantycznych zarówno w stosunku do błędów fonetycznych, jak i tematycznych (jednowymiarowe testy istotności dla porównań zaplanowanych prezentowane są w tabeli 4), co obrazuje wykres 3. Ponownie nie odnotowano żadnego wpływu płci na wyniki – brak efektu głównego ($F(1;183) = 0,44$, $p > 0,5$) i jakichkolwiek interakcji innych zmiennych z płcią. Tabela 3 zawiera szczegółowe wyniki analizy wariancji (bez uwzględnienia zmiennej ‘płeć’). Wystąpiły interakcje wieku i typu błędu, części mowy i typu błędu oraz interakcja tych trzech zmiennych. O wyższej proporcji błędów semantycznych zdecydowała ich przewaga w rzeczownikach (dla interakcji część mowy \times typ błędu jednowymiarowe testy

istotności dla porównań zaplanowanych prezentowane są w tabeli 5 oraz na wykresie 4) w porównaniu z czasownikami i przymiotnikami. W przypadku błędów fonetycznych największy ich procent odnotowano dla przymiotników, a najmniejszy dla czasowników (wszystkie kontrasty istotne – tabela 5). Błędy tematyczne pojawiały się jednakowo często w przypadku wszystkich trzech części mowy (żaden kontrast nie był istotny – tabela 5). Zmiany rozwojowe w proporcji poszczególnych typów błędów są przedstawione na wykresie 5. Błędy semantyczne utrzymują się przez cały badany okres na stałym poziomie (wszystkie kontrasty nieistotne z wyjątkiem jednego – dla największej i najmniejszej wartości – pomiędzy grupami wiekowymi 2–3 i 3–4, $F(1;183) = 10,24$; $p < 0,002$), podczas gdy proporcja błędów tematycznych systematycznie maleje, a proporcja błędów fonetycznych zmienia się nieliniowo, ale również z tendencją malejącą. Gdy rozpatrywać zmiany rozwojowe w proporcji poszczególnych błędów w zależności od części mowy (co jest uwidocznione na wykresie 6), wyraźnie widać, że systematyczny spadek proporcji błędów tematycznych wraz z wiekiem dotyczy wszystkich części mowy. W przypadku błędów fonetycznych obniżenie wraz z wiekiem ich częstości dotyczy przede wszystkim czasowników (w ostatnich trzech grupach wiekowych). Natomiast błędy semantyczne utrzymują się na względnie stałym poziomie przez cały badany okres dla wszystkich części mowy.

Tabela 3. Wyniki analizy wariancji dla dzieci 2–3-letnich (pula 264 pozycji) w schemacie $8 \times 2 \times 3 \times 3$ (wiek \times płeć \times część mowy \times typ błędu). Zmienne: 1 – wiek (kwartalne przedziały), 2 – część mowy (reczowniki, czasowniki, przymiotniki), 3 – typ błędu (fonetyczny, semantyczny, tematyczny)

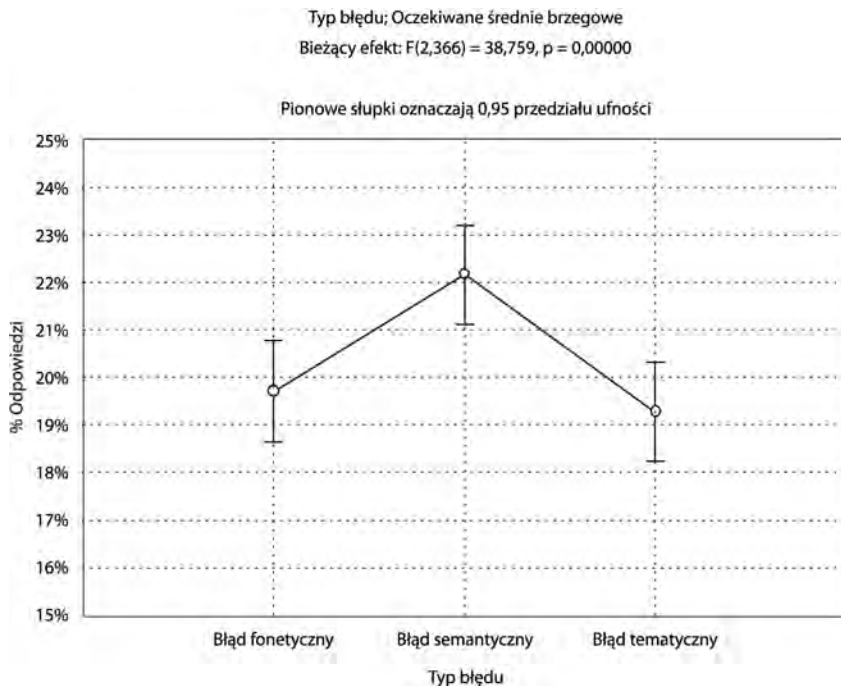
Zmienna	F	Stopnie swobody	p
1	7,72	7; 183	0,001
2	34,67	2; 366	0,001
3	38,76	2; 366	0,001
1, 2	0,74	14; 366	0,736
1, 3	3,49	14; 366	0,001
2, 3	7,54	4; 732	0,001
1, 2, 3	1,82	28; 732	0,006

Tabela 4. Wyniki jednowymiarowych testów istotności dla porównań zaplanowanych dla zmiennej ‘typ błędu’ u dzieci młodszych. Poziomy zmiennej: 1 – błąd fonetyczny; 2 – błąd semantyczny; 3 – błąd tematyczny

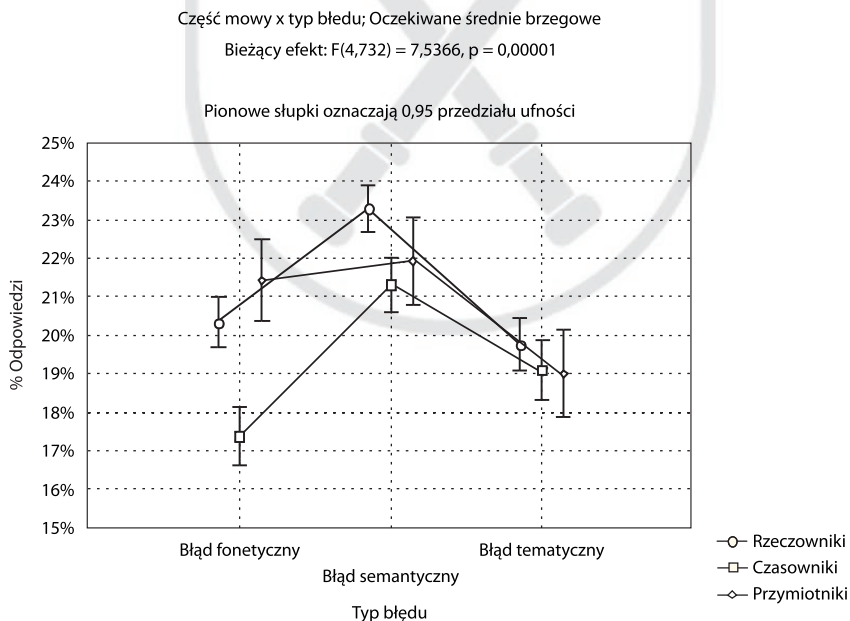
Porównanie	F	Stopnie swobody	p
1 vs. 2	47,24	1; 183	0,001
2 vs. 3	61,5	1; 183	0,001
1 vs. 3	1,58	1; 183	0,21

Tabela 5. Wyniki jednowymiarowych testów istotności dla porównań zaplanowanych dla interakcji zmiennych ‘część mowy’ i ‘typ błędu’. Poziomy zmiennej ‘część mowy’: 1 – rzeczowniki; 2 – czasowniki; 3 – przymiotniki

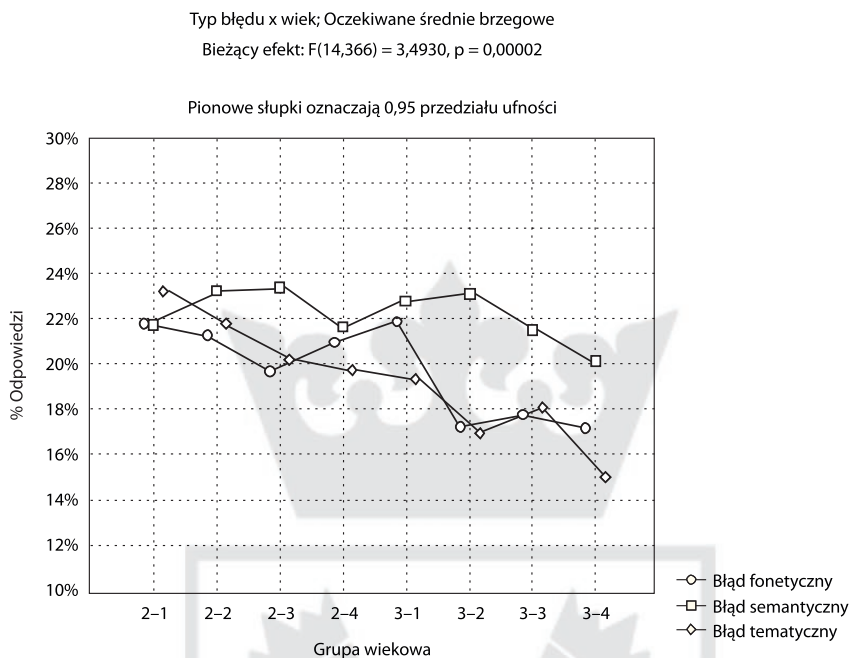
Porównanie	F	Stopnie swobody	p
dla błędów fonetycznych			
1 vs. 2	50,88	1; 183	0,001
1 vs. 3	4,84	1; 183	0,029
2 vs. 3	49,78	1; 183	0,001
dla błędów semantycznych			
1 vs. 2	27,26	1; 183	0,001
1 vs. 3	5,86	1; 183	0,017
2 vs. 3	1,01	1; 183	0,3
dla błędów tematycznych			
1 vs. 2	2,39	1; 183	0,12
1 vs. 3	1,67	1; 183	0,19
2 vs. 3	0,02	1; 183	0,89



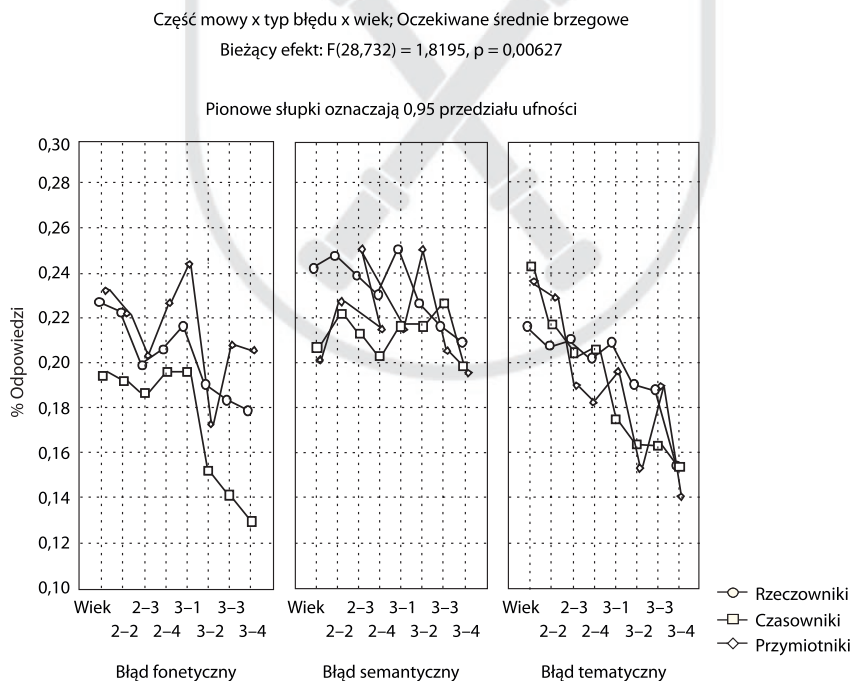
Wykres 3. Proporcje poszczególnych typów błędów: analiza wariancji w schemacie wiek \times płeć \times część mowy \times typ błędu dla dzieci młodszych



Wykres 4. Proporcje poszczególnych typów błędów w zależności od kategorii gramatycznej (części mowy): analiza wariancji w schemacie wiek \times płeć \times część mowy \times typ błędu dla dzieci młodszych



Wykres 5. Zmiany rozwojowe w proporcji poszczególnych typów błędów: analiza wariancji w schemacie wiek \times płeć \times część mowy \times typ błędu dla dzieci młodszych



Wykres 6. Zmiany rozwojowe w proporcji poszczególnych typów błędów w zależności od części mowy: analiza wariancji w schemacie wiek \times płeć \times część mowy \times typ błędu dla dzieci młodszych

Analizy poprawności odpowiedzi oraz typów popełnianych błędów dla dzieci starszych (4–6-latków)

Analiza poprawności odpowiedzi dla dzieci starszych – 497 pozycji

Dzieci starsze (badane jako pierwsze) odpowiadały w badaniu na 497 pytań testowych. Wyniki dzieci 4-letnich były analizowane w podziale na półroczne przedziały wiekowe, a dzieci 5–6-letnich na przedziały roczne. Wstępne obliczenia przeprowadzone w trakcie badań wskazywały na spore zróżnicowanie w obrębie grupy 4-latków, dlatego też zdecydowano o jej podziale. Tabela 6 zawiera średnie dla liczby poprawnych odpowiedzi w poszczególnych przedziałach wiekowych z podziałem na płeć badanych w każdej z trzech części mowy.

Analiza wariancji w schemacie $4 \times 2 \times 3$ (wiek \times płeć \times część mowy) z powtarzaniem pomiarem na danych procentowych wykazała istotny wpływ wieku i części mowy na poprawność odpowiedzi oraz interakcję tych dwóch zmiennych. Wpływ zmiennej ‘płeć’ okazał się nieistotny. Tabela 7 przedstawia szczegółowe

wyniki analizy. Z wykresu 7 wynika, że wraz z wiekiem istotnie rośnie procent poprawnych odpowiedzi, a zależność ta ma charakter liniowy. Testy Duncana wykazały, że kontrasty są istotne między każdymi dwoma kolejnymi przedziałami wieku ($p < 0,005$). Wykres 8 ukazuje wpływ części mowy na poprawność odpowiedzi: czasowniki (podobnie jak w grupie młodszej) okazały się najprostsze, rzeczowniki nieco trudniejsze, najmniejszy procent poprawnych odpowiedzi uzyskały przymiotniki (wszystkie kontrasty mierzone testami NIR Fishera istotne, $p < 0,001$). Wykres 9 przedstawia interakcję wieku i części mowy. Wynika z niego, że przymiotniki są nie tylko najmniej poprawnie rozpoznawaną częścią mowy, ale także proporcja poprawnych odpowiedzi w tej kategorii rośnie z wiekiem stosunkowo najwolniej. Kontrasty pomiędzy poszczególnymi częściami mowy obliczane dla każdego przedziału wiekowego wskazują na istotne różnice pomiędzy przymiotnikami a rzeczownikami i czasownikami bez względu na grupę wiekową (testy NIR Fishera, $p < 0,05$). Wyniki dla rzeczowników i czasowników dla każdej kolejnej pary grup wiekowych są istotnie różne (testy

Tabela 6. Średnie sumy (i odchylenia standardowe) poprawnych odpowiedzi w półrocznych lub rocznych przedziałach wiekowych* dla każdej z płci – dzieci starsze (C – chłopcy; D – dziewczynki; N – liczba osób w grupie)

Wiek	Płeć	Rzeczowniki (max = 289) średnia	Rzeczowniki odch. st.	Czasowniki (max = 131) średnia	Czasowniki odch. st.	Przymiotniki (max = 77) średnia	Przymiotniki odch. st.	N
4–(1–2)	C	152,33	25,54	70,80	12,95	35,60	5,77	15
4–(1–2)	D	159,00	13,66	73,18	7,83	38,91	5,22	11
4–(3–4)	C	169,33	23,64	81,50	13,87	39,67	6,57	12
4–(3–4)	D	174,44	20,26	81,31	12,26	38,63	5,34	16
5–(1–4)	C	185,22	28,78	86,88	16,92	43,19	6,70	32
5–(1–4)	D	192,96	22,39	90,88	11,89	43,65	6,32	26
6–(1–4)	C	205,42	24,37	97,00	9,60	49,74	7,72	19
6–(1–4)	D	214,24	15,60	100,00	7,48	50,62	4,44	21
	Łącznie	185,55	29,65	87,05	15,28	43,30	7,73	152

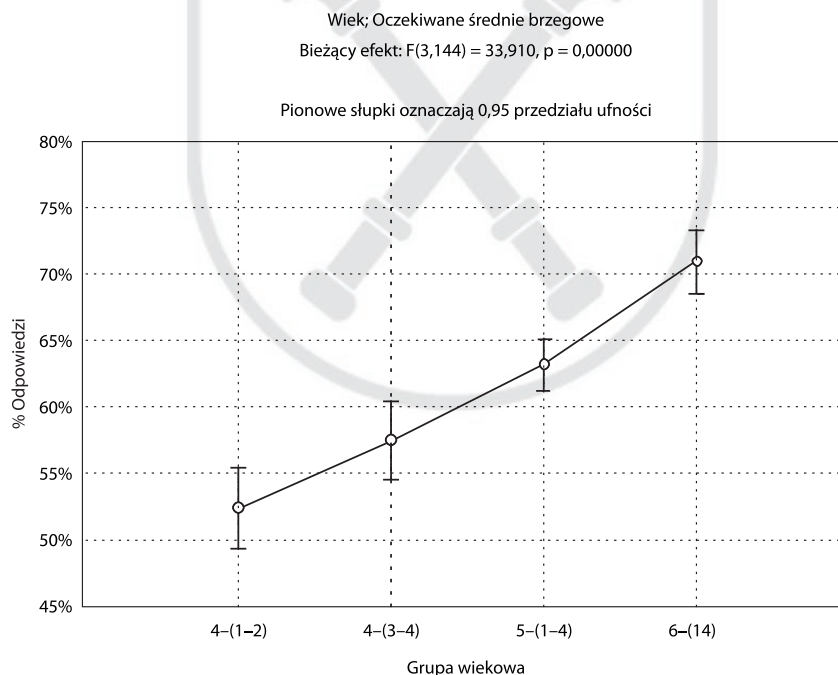
* liczby w nawiasie w kolumnie ‘wiek’ oznaczają kwartały danego roku

NIR Fishera, $p < 0,05$), a więc w całym badanym okresie procent poprawnych odpowiedzi dla obu tych części mowy rośnie, natomiast dla przymiotników kontrast pomiędzy dwiema

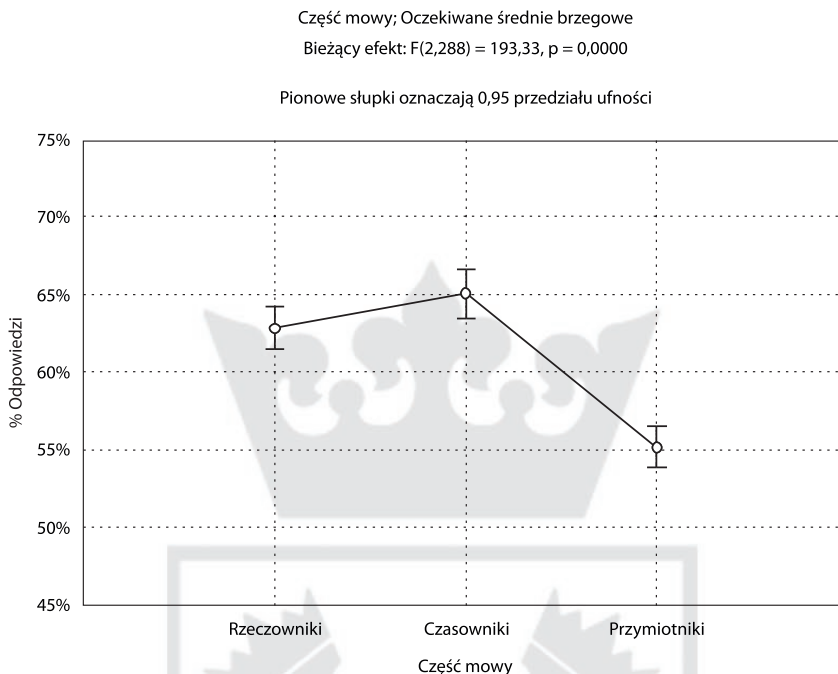
pierwszymi grupami wiekowymi jest nieistotny (test NIR Fishera, $p > 0,24$), a istotny wzrost zaczyna się dopiero na przełomie drugiego półrocza czwartego i piątego roku życia.

Tabela 7. Wyniki analizy wariancji dla dzieci 4–6-letnich (pula 497 pozycji) w schemacie $4 \times 2 \times 3$ (wiek \times płeć \times część mowy) dla poprawnych odpowiedzi. Zmienne: 1 – wiek, 2 – płeć, 3 – część mowy (rzeczowniki, czasowniki, przymiotniki)

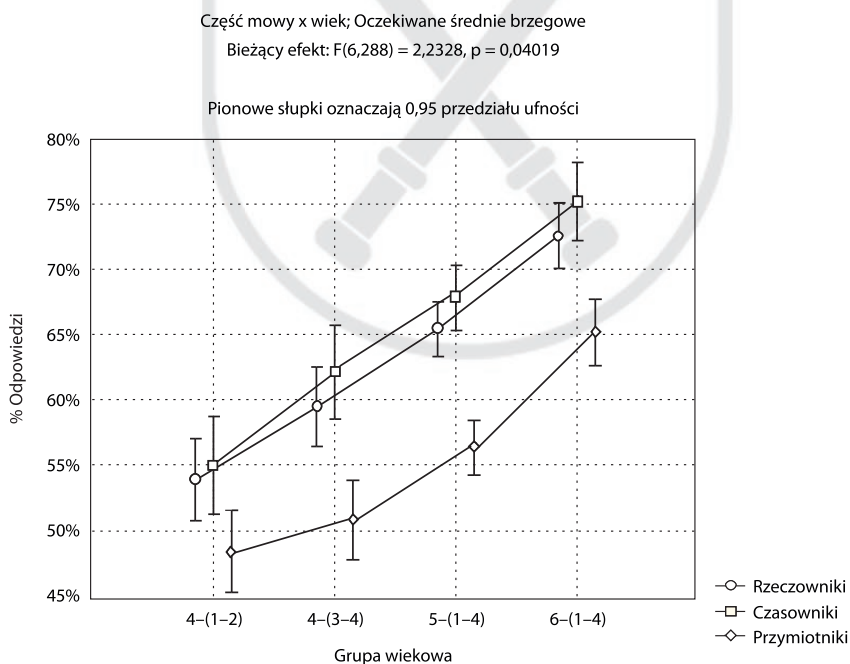
Zmienna	F	Stopnie swobody	p
1	33,91	3; 144	0,001
2	1,79	1; 144	0,18
3	193,33	2; 288	0,001
1, 2	0,16	3; 144	0,923
1, 3	2,23	6; 288	0,041
2, 3	0,74	2; 288	0,47
1, 2, 3	0,7	6; 288	0,66



Wykres 7. Wpływ wieku na poprawność odpowiedzi: analiza wariancji w schemacie wiek \times płeć \times część mowy dla dzieci starszych



Wykres 8. Wpływ kategorii gramatycznej (części mowy) na poprawność odpowiedzi: analiza wariancji w schemacie wiek \times płeć \times część mowy dla dzieci starszych



Wykres 9. Interakcja kategorii gramatycznej (części mowy) i wieku: analiza wariancji w schemacie wiek \times płeć \times część mowy dla dzieci starszych

Analiza typów błędów dla dzieci starszych (4–6-latków) – 497 pozycji

Analiza wariancji w schemacie: wiek (4) × płeć (2) × część mowy (3) × typ błędu (3) z powtarzanymi pomiarami (dla czynników ‘część mowy’ i ‘typ błędu’) dla danych procentowych wykazała istotne efekty główne wieku, części mowy i typu błędu. Szczegółowe wyniki analizy prezentowane są w tabeli 8. W przypadku zmiennych ‘wiek’ i ‘część mowy’ wyniki były analogiczne w stosunku do analizy poprawnych odpowiedzi: wraz z wiekiem maleje proporcja błędnych odpowiedzi, czasowniki mają najmniejszą proporcję błędów, a przymiotniki największą, rzeczowniki lokują się pomiędzy tymi dwiema kategoriami (wszystkie kontrasty są istotne, testy NIR Fishera, $p < 0,001$). W przypadku zmiennej ‘typ błędu’ różnica pomiędzy poszczególnymi typami

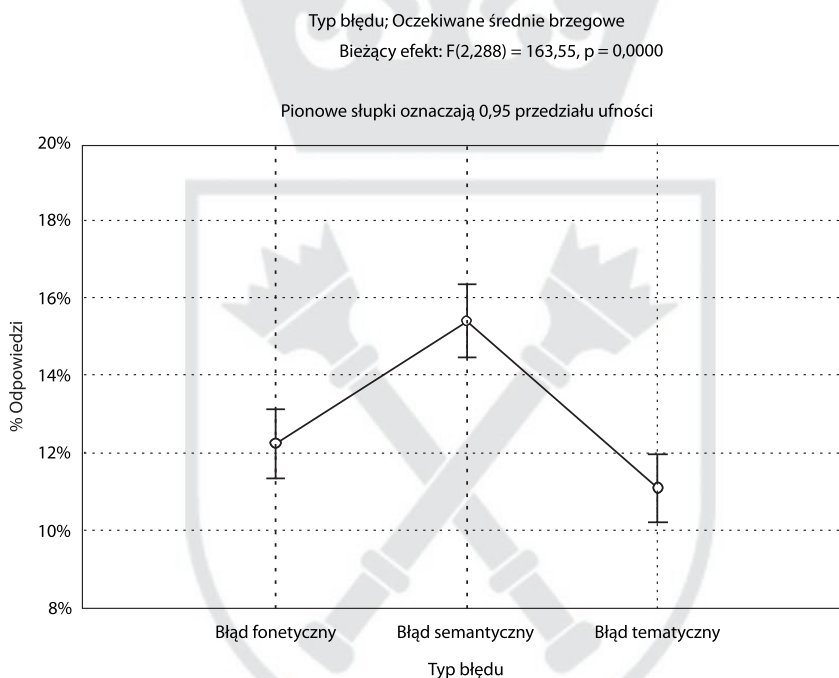
błędów polegała na istotnie większej proporcji błędów semantycznych zarówno w stosunku do błędów fonetycznych, jak i tematycznych, błędów tematycznych było jednak istotnie mniej niż fonetycznych (jednowymiarowe testy istotności dla porównań zaplanowanych prezentowane są w tabeli 9), co obrazuje wykres 10. Okazało się także, że u dzieci starszych wystąpiła interakcja kategorii gramatycznej (części mowy) i typu błędu. Proporcja błędów fonetycznych w poszczególnych częściach mowy była istotnie różna (najwięcej błędów tego typu wystąpiło w przymiotnikach, a najmniej w czasownikach) – co widoczne jest na wykresie 11. Wystąpiła także interakcja typu błędu z płcią – wykres 12 przedstawia tę zależność: chłopcy popełniali istotnie więcej błędów semantycznych niż dziewczynki, dla pozostałych kategorii błędów nie było różnic pomiędzy płciami.

Tabela 8. Wyniki analizy wariancji dla dzieci 4–6-letnich (pula 497 pozycji) w schemacie wiek (4) × płeć (2) × część mowy (3) × typ błędu (3). Zmienne: 1 – wiek (kwartalne przedziały), 2 – płeć, 3 – część mowy (rzeczowniki, czasowniki, przymiotniki), 4 – typ błędu (fonetyczny, semantyczny, tematyczny)

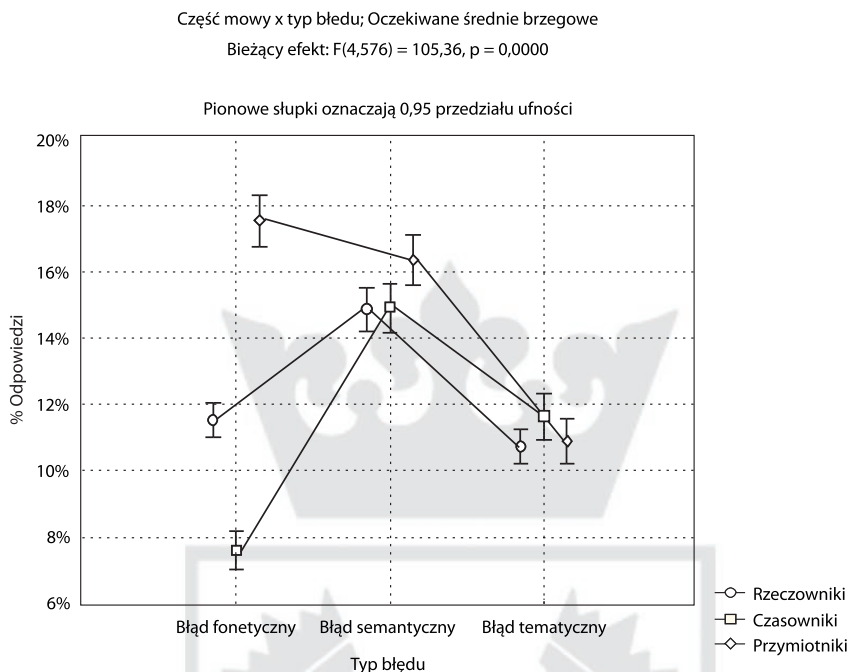
Zmienna	F	Stopnie swobody	p
1	33,97	3; 144	0,001
2	1,774	1; 144	0,185
3	217,5	2; 288	0,001
4	163,5	2; 288	0,001
1, 2	0,162	3; 144	0,922
1, 3	2,247	6; 288	0,039
1, 4	0,332	6; 288	0,920
2, 3	0,740	6; 288	0,478
2, 4	6,790	2; 288	0,002
3, 4	105,4	4; 576	0,001
1, 2, 3	0,696	6; 288	0,653
1, 2, 4	0,572	6; 288	0,753
1, 3, 4	0,979	12; 576	0,468
2, 3, 4	2,849	4; 576	0,023
1, 2, 3, 4	1,868	12; 576	0,036

Tabela 9. Wyniki jednowymiarowych testów istotności dla porównań zaplanowanych dla zmiennej 'typ błędu' u dzieci starszych. Poziomy zmiennej: 1 – błąd fonetyczny; 2 – błąd semantyczny; 3 – błąd tematyczny

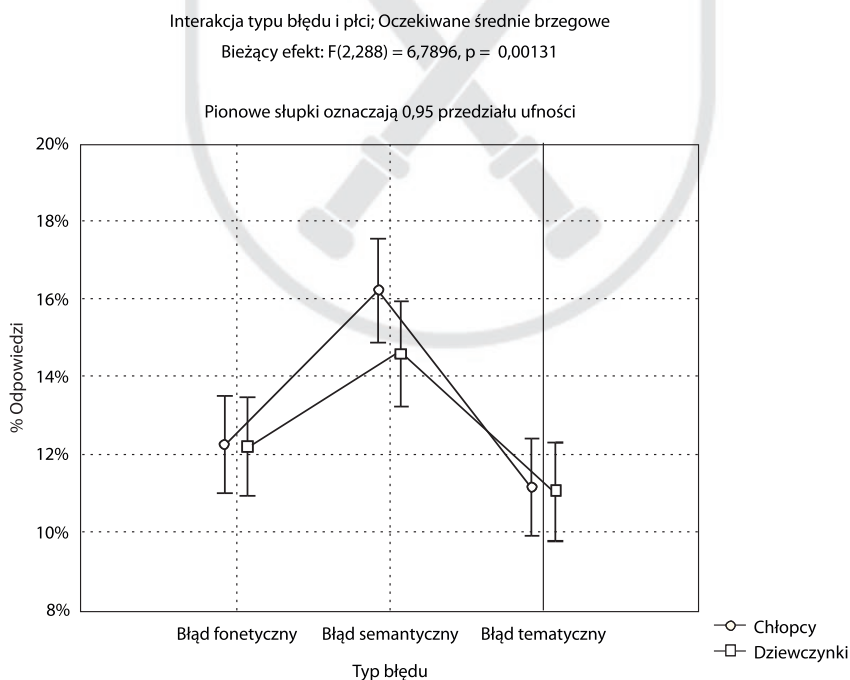
Porównanie	F	Stopnie swobody	p
1 vs. 2	141,4	1; 144	0,001
2 vs. 3	357,9	1; 144	0,001
1 vs. 3	21,63	1; 144	0,001



Wykres 10. Proporcje poszczególnych typów błędów: analiza wariancji w schemacie wiek × płęć × część mowy × typ błędu dla dzieci starszych



Wykres 11. Proporcje poszczególnych typów błędów w zależności od części mowy: analiza wariancji w schemacie wiek \times płeć \times część mowy \times typ błędu dla dzieci starszych



Wykres 12. Proporcje poszczególnych typów błędów w zależności od płci: analiza wariancji w schemacie wiek \times płeć \times część mowy \times typ błędu dla dzieci starszych

DYSKUSJA WYNIKÓW

Przedstawione wyniki można interpretować z jednej strony jako świadectwo zmian rozwojowych i charakterystyki rozwoju słownиковego dzieci w wieku 2–6 lat. Z drugiej strony prezentowane badania miały także służyć analizie właściwości użytego materiału testowego – psychometrycznej charakterystyce poszczególnych pozycji testowych i ogólnych właściwości narzędzia testowego. Ze względu na pierwsze zastosowanie opisywanej tu metody (które jest przedmiotem tego artykułu), a także brak wcześniejszych danych na temat charakterystyki narzędzia testowego wnioski dotyczące rozwoju leksykalnego mogą być obciążone błędem związanym właśnie z nieznaną specyfiką narzędzia. Wnioski te formułujemy więc, mając na uwadze powyższe zastrzeżenie. Dzięki analizie uzyskanych wyników mogliśmy jednak przygotować nową, zrównoważoną pod względem wielu czynników wersję narzędzia. Sam proces tworzenia nowej wersji nie jest w tym artykule szczegółowo opisywany pod względem technicznym, ale w ramach dyskusji przedstawiamy zasady, którymi kierowaliśmy się przy tworzeniu kolejnej wersji Obrazkowego Testu Słownikowego.

Na podstawie uzyskanych wyników, w aspekcie interpretacji dotyczącej rozwoju leksykalnego, możemy stwierdzić, że wraz z wiekiem rośnie zasób słownictwa. U dzieci młodszych (2–3 lata) rośnie od 28% do 46% (ok. 74 do 121 pozycji testowych), ale wzrost ten nie ma charakteru liniowego. Oznacza to, że istotny skok rozwojowy nie następuje pomiędzy każdymi kolejnymi dwoma przedziałami wiekowymi. Mogłoby to sugerować, że dynamika rozwoju w tym wieku nie jest jeszcze zbyt duża, co przeczyłoby dotychczasowym wynikom badań, zgodnie z którymi jest to okres bardzo intensywnych zmian w rozwoju słownictwa (np. Fenson i in., 2007). Możliwe zatem, że o braku istotnych różnic w tym przypadku zdecydowała albo charakterystyka narzędzia, które ewidentnie zawierało dużo pozycji trudnych dla dzieci w tym wieku, skoro nawet dzieci w czwartym kwartale trzeciego roku życia średnio odpowiadały poprawnie na mniej niż

50% pozycji testowych, albo przyjęty podział na kwartalne przedziały wiekowe. W przypadku dzieci starszych (4–6 lat), gdzie przedziały wiekowe ustalono w odstępach półrocznych (dla 4-latków) lub rocznych (dla 5- i 6-latków), różnice pomiędzy każdą parą kolejnych przedziałów były istotne, a więc dynamikę zmian można by interpretować jako wyraźną. Dzieci te średnio odpowiadały poprawnie w przypadku ponad 50% pozycji (wzrost od 52% do 71%). Warto jednak zauważyć, że w grupie dzieci młodszych na przestrzeni całego analizowanego okresu (2 lata) wzrost wyniósł 18%, a w grupie dzieci starszych (okres 3 lat) jedynie 19%. A więc biorąc pod uwagę długość analizowanych okresów, u dzieci starszych wzrost następował wolniej. Tu jednak należy uwzględnić fakt, że większa pula testowa (497 pozycji) dla dzieci starszych zawierała większą liczbę pozycji trudnych (które u dzieci młodszych nie zostały użyte), i to mogło zadecydować o pozornie wolniejszym wzroście zasobu słów.

Wyniki wskazują także na różnice w rozumieniu poszczególnych części mowy: zarówno u dzieci młodszych, jak i starszych największy odsetek poprawnych odpowiedzi uzyskiwały czasowniki (odpowiednio 37% i 67%). W przypadku dzieci młodszych rzeczowniki i przymiotniki były na zbliżonym poziomie (35–36%), a u dzieci starszych rzeczowniki przewyższały przymiotniki (65% vs. 57%). Jednak wyciąganie wniosku o większej łatwości w posługiwaniu się czasownikami mogłoby być pochopne. Możliwe, że wynik ten przede wszystkim świadczy o tym, że mimo staranności w doborze słów kluczowych zbiór czasowników zawierał więcej słów prostych dla dzieci. Ta ostatnia możliwość jest tym bardziej prawdopodobna, że aktualne teorie rozwoju słownikowego przewidują pierwszeństwo rzeczowników nad czasownikami (np. Gentner i Boroditsky, 2001). Jest to zatem wskazówka, żeby przy doborze pozycji do kolejnej wersji testu zadbać szczególnie o zrównoważenie dobieranych pozycji pod względem trudności w poszczególnych częściach mowy.

Ani w grupie młodszej, ani w starszej w analizach poprawności nie wystąpił efekt główny płci ani interakcja płci z innymi zmiennymi.

Różnic takich – na korzyść dziewczynek – można by oczekiwać zwłaszcza u dzieci młodszych, gdyż badania są zgodne co do tego, że dziewczynki wcześniej zaczynają mówić, a zatem początkowo mają większy zasób słownictwa (Fenson i in., 2007). Nasz test w obecnej postaci okazał się niewrażliwy na tego typu różnice. Jedyna istotna różnica dotycząca płci wystąpiła w przypadku analizy błędów u dzieci starszych – u chłopców zaobserwowano istotnie więcej błędów semantycznych niż u dziewczynek. Może to świadczyć o silniejszej u chłopców tendencji do tworzenia związków paradygmatycznych (opartych na kategoryzacji) w przypadku niepełnej lub niedokładnej znajomości znaczenia słowa.

W przypadku analizy typów błędów w obu badanych grupach przeważały błędy semantyczne (22% w grupie młodszej, 15% w grupie starszej). Można to interpretować jako przewagę związków paradygmatycznych (semantycznych) w słowniku umysłowym nad związkami syntagmatycznymi (opartymi na relacjach tematycznych) i fonetycznymi (opartymi na podobieństwie brzmieniowym). Potwierdzają to także zmiany rozwojowe – proporcja błędów tematycznych maleje systematycznie w obu badanych grupach wraz z wiekiem, a dla grupy starszej jest istotnie mniejsza od pozostałych dwóch typów błędów w całym badanym okresie. Inaczej natomiast wygląda w obu grupach interakcja typów błędów i części mowy. O ile u dzieci młodszych widoczna jest znacząco niższa proporcja błędów fonetycznych w czasownikach (w stosunku do obu pozostałych kategorii), a wyższa proporcja błędów semantycznych w rzeczownikach (wśród błędów tematycznych nie ma różnic pomiędzy kategoriami części mowy), o tyle u dzieci starszych różnice występują wyłącznie w obrębie błędów fonetycznych: przymiotniki mają ich najwięcej, a czasowniki najmniej. Zatem przymiotniki, jako generalnie najtrudniej rozpoznawana część mowy, dla dzieci starszych stanowiły

trudność przede wszystkim pod względem fonetycznym. Systematyczne różnice pomiędzy trzema typami błędów występujące zarówno u dzieci młodszych, jak i starszych świadczą także o tym, że dzieci w jakimś stopniu postrzegają relacje łączące słowo kluczowe i słowa dystraktory – ich błędy nie wynikają jedynie z całkowitej nieznaności słowa kluczowego, ale częściowo mogą być interpretowane jako jego ‘niepełna’ interpretacja – mylenie słowa kluczowego z innym bliskim semantycznie (a więc należącym do tej samej kategorii) może stanowić pewien ślad niedokładnej znajomości znaczenia.

Przedstawiona analiza wyników pierwszej fazy konstruowania Obrazkowego Testu Słownikowego pozwoliła na wyciągnięcie wstępnych wniosków dotyczących właściwości rozwoju słownika umysłowego dzieci w wieku od 2 do 6 lat, których prezentacja była celem tego artykułu. Uzyskane dane stały się także podstawą do tworzenia nowszej wersji testu – o znacząco mniejszej liczbie pozycji. Przygotowanie w pierwszej fazie badań tak dużej liczby pozycji testowych (497) umożliwiło też wyłonienie z nich dwóch równoległych wariantów testu o zbliżonych parametrach. Przy doborze słów do nowszej wersji testu opieraliśmy się na założeniu, że wersja ta powinna zawierać pozycje uporządkowane według trudności (od najłatwiejszych do najtrudniejszych), tak żeby w wersji ostatecznej możliwe było przerywanie testu, gdy dziecko zacznie udzielać w przeważającej mierze odpowiedzi niepoprawnych. Do nowej wersji testu wybrane zostały pozycje, które uzyskały nieujemną moc dyskryminacyjną, miały niemalejący trend rozwojowy (dla procentu poprawnych odpowiedzi) oraz różniły się pomiędzy sobą ogólnym poziomem trudności (pozycje o podobnym poziomie trudności przypisywane były do różnych wariantów testu). Tak przygotowane dwa warianty testu będą podstawą kolejnych badań, które pozwolą na dalsze skrócenie testu i bardziej precyzyjne ustalenie właściwości psychometrycznych testu.

PRZYPISY

¹ Prezentowane badania były finansowane ze środków publicznych w ramach grantu MNiSz oraz w ramach projektu BST finansowanego przez Wydział Psychologii UW. Autorzy składają gorące podziękowania dr hab. Magdalenie Smoczyńskiej (UJ) za wsparcie i pomoc merytoryczną przy opracowaniu prezentowanego narzędzia.

BIBLIOGRAFIA

- Aitchison J. (2003), *Words in the Mind: An Introduction to the Mental Lexicon*. 3rd edition. Oxford: Basil Blackwell.
- Berko Gleason J., Bernstein Ratner N. (red.) (1998/2005), *Psycholinguistics*. Orlando: Harcourt. Wydanie polskie: *Psycholingwistyka*. Gdańsk: GWP.
- Bloom P. (1996/2007), Controversies in Language Acquisition: Word Learning and the Part of Speech [w:] R. Gelman, T.K. Au (red.). *Perceptual and Cognitive Development*, 151–184. Wydanie polskie: Kontrowersje wokół przyswajania języka: uczenie się wyrazów i części mowy [w:] B. Bokus, G.W. Shugar (red.). *Psychologia języka dziecka. Osiągnięcia, nowe perspektywy*, 175–210. Gdańsk: GWP (tłum. Maciej Haman).
- Bloom P. (2000), *How Children Learn the Meaning of Words*. Cambridge, MA: MIT Press.
- Bornstein M.H., Cote L.R., Maital S., Painter K., Park S.Y., Pascual L., Pecheux M.G., Ruel J., Venuti P., Vyt A. (2004), Cross-linguistic Analysis of Vocabulary in Young Children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development*, 75 (4), 1115–1139.
- Caselli M.C., Bates E., Casadio P., Fenson L., Fenson J., Sanderl L., Weir J. (1995), A Cross-linguistic Study of Early Lexical Development. *Cognitive Development*, 10, 159–199.
- Caselli M.C., Casadio P., Bates E. (1999), A Comparison of the Transition from First Words to Grammar in English and Italian. *Journal of Child Language*, 26, 69–111.
- Chomsky N. (1957), *Syntactic Structures*. Haga: Mouton.
- Ciechanowicz A. (1992), *Skala Dojrzałości Umysłowej Columbia. Podręcznik. Polska standaryzacja 1990*. Warszawa: PTP.
- Clark E.V. (1993), *The Lexicon in Acquisition*. Cambridge: Cambridge University Press.
- Clark E.V. (1995/2007), Language Acquisition: The Lexicon and Syntax [w:] J.L. Miller, P.D. Eimas (red.). *Speech, Language and Communication*, 303–337. San Diego: Academic Press. Wydanie polskie: Przyswajanie języka: słownik i składnia [w:] B. Bokus, G.W. Shugar (red.). *Psychologia języka dziecka. Osiągnięcia, nowe perspektywy*, 175–210. Gdańsk: GWP (tłum. Ewa Haman).
- Clark E.V. (2003), *First Language Acquisition*. Cambridge: Cambridge University Press.
- Dale P.S., Dionne G., Eley T.C., Plomin R. (2000), Lexical and Grammatical Development: a Behavioural Genetic Perspective. *Journal of Child Language*, 27, 619–642.
- Dionne G., Dale P.S., Boivin M., Plomin R. (2003), Genetic Evidence for Bidirectional Effects of Early Lexical and Grammatical Development. *Child Development*, 74 (2), 394–412.
- Dixon A.J., Marchman V.A. (2007), Grammar and the Lexicon: Developmental Ordering in Language Acquisition. *Child Development*, 78 (1), 190–212.
- Dunn L., Dunn L. (1997), *Peabody Picture Vocabulary Test – Revised*. 3rd edition. Circle Pines, MN: American Guidance Service.
- Fenson L., Marchman V., Thal D.J., Dale P.S., Reznick J.S., Bates E. (2007), *MacArthur-Bates Communicative Development Inventories. User's Guide and Technical Manual*. 2nd edition. Baltimore: Paul Brooks Publishing Co.
- Gentner D., Boroditsky L. (2001), Individuation, Relativity and Early Word Learning [w:] M. Bowerman, S. Levinson (red.). *Language Acquisition and Conceptual Development*, 215–256. Cambridge, UK: Cambridge University Press.
- Gleitman L., Landau B. (1994), *The Acquisition of the Lexicon: Remarks on Lexical Knowledge*. Cambridge, MA: MIT Press.

- Golinkoff R.M., Hirsh-Pasek K., Bloom L., Smith L., Woodward A., Akhtar N. (2000), *Becoming a Word Learner: A Debate on Lexical Acquisition*. New York, NY: Oxford University Press.
- Gopnik A., Choi S., Baumberger T. (1996), Cross-linguistic Differences in Early Semantic and Cognitive Development. *Cognitive Development*, 11 (2), 197–225.
- Hall D.G., Waxman S.R. (2004), *Weaving a Lexicon*. Cambridge, MA: MIT Press.
- Jaworowska A., Matczak A., Szustrowa T. (1996), *Międzynarodowa Wykonaniowa Skala Leitera. Polska standaryzacja: 1993*. Warszawa: PTP.
- Jurkowski A. (1997), *Test Językowy Leksykon*. Warszawa: PTP.
- Kurcz I. (2005), *Psychologia języka i komunikacji*. Warszawa: Scholar.
- Leonard L.B. (1998/2006), *Children with Specific Language Impairment*. Cambridge, MA: the MIT Press.
Wydanie polskie: *SLI – Specyficzne zaburzenie rozwoju językowego*. Gdańsk: GWP (tłum. Mikołaj Hernik).
- Lista Frekwencyjna Korpusu Języka Polskiego PWN (2004).
- Lyytinen H., Erskine J., Tolvanen A., Torppa M., Poikkeus A.M., Lyytinen P. (2006), Trajectories of Reading Development: A Follow-up From Birth to School age of Children With and Without Risk for Dyslexia. *Merrill-Palmer Quarterly*, 52 (3), 514–546.
- Marchman V.A., Bates E. (1994), Continuity in Lexical and Morphological Development: A Test of the Critical Mass Hypothesis. *Journal of Child Language*, 21 (2), 339–366.
- Marchman V.A., Fernald A. (2008), Speed of Word Recognition and Vocabulary Knowledge in Infancy Predict Cognitive and Language Outcomes in Later Childhood. *Developmental Science*, 11 (3), F9–F16.
- Markman E.M. (1989), *Categorization and Naming in Children*. Cambridge, MA: MIT Press.
- Matczak A. (2001), *Diagnoza Możliwości Intelktualnych – 2*. Warszawa: PTP.
- Matczak A., Piotrowska A., Ciarkowska W. (1998), *Skala Inteligencji Wechslera dla dzieci wersja zmodyfikowana WISC-R*. Warszawa: PTP.
- Matczak A., Jaworowska A., Ciecchanowicz A., Fecenec D., Stańczak J., Zalewska E. (2007), *Dziecięca Skala Rozwojowa*. Warszawa: PTP.
- Michnick-Golinkoff R.M., Hirsh-Pasek K., Bloom L., Smith L., Woodward A., Akhtar N. (2000), *Becoming a Word Learner: A Debate on Lexical Acquisition*. New York, NY: Oxford University Press.
- Naigles L. (1990), Children Use Syntax to Learn Verb Meanings. *Journal of Child Language*, 17 (2), 357–374.
- Naigles L. (1996), The Use of Multiple Frames in Verb Learning Via Syntactic Bootstrapping. *Cognition*, 58, 221–251.
- Nelson K. (1996), *Language in Cognitive Development: Emergence of the Mediated Mind*. Cambridge: Cambridge University Press.
- Reeves L.M., Hirsch-Pasek K., Golinkoff R. (1998/2005), Words and Meaning: from Primitives to Complex Organization [w:] J. Berko Gleason, N. Bernstein Ratner (red.). *Psycholinguistics*, 157–226. Orlando: Harcourt. Wydanie polskie: Słowa i znaczenia: od pojęć pierwotnych do złożonych struktur [w:] J. Berko Gleason, N. Bernstein Ratner (red.). *Psycholingwistyka*, 173–248. Gdańsk: GWP (tłum. Jerzy Bobryk).
- Smoczyńska M. (1999), *Inwentarz Rozwoju Mowy i Komunikacji*. Kraków: Uniwersytet Jagielloński.
- Snow C.E., Griffin P.M., Burns S. (2005), *Knowledge to Support the Teaching of Reading*. San Francisco: Yossey-Bass Publishers.
- Tomasello M. (2003), *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.
- Wagner R.K., Muse A.E., Tannenbaum K.R. (2006), *Vocabulary Acquisition. Implications for Reading Comprehension*. New York: Guilford Press.
- Watson A.C., Painter K., Bornstein M.H. (2001), Longitudinal Relations between 2-year-olds' Language and 4-year-olds' Theory of Mind. *Journal of Cognition and Development*, 2 (4), 449–457.