

MAŁGORZATA BANACH\* | UNIwersYTET Jagielloński, KRAKÓW

# Rozumienie tekstu pisanego przez zdających egzamin certyfikacyjny z języka polskiego jako obcego

Słowa kluczowe: język polski jako obcy, egzaminy certyfikacyjne, rozumienie tekstów pisanych, glottodydaktyka porównawcza.

doi: <http://dx.doi.org/10.31286/JP.98.2.4>

## 1. Wprowadzenie

Przez wiele lat badania znajomości języka polskiego poza Polską koncentrowały się głównie na opanowaniu przez przedstawicieli Polonii systemu naszego języka (Miodunka 2013b: 66). W glottodydaktyce polonistycznej nowe perspektywy badawcze otworzyło uruchomienie w 2004 roku, na mocy nowelizacji z 30 kwietnia 2003 roku ustawy o języku polskim (Dz. U. z 2003 r. Nr 73, poz. 661), systemu państwowych egzaminów certyfikacyjnych z języka polskiego jako obcego, uznawanych za najistotniejszą na progu XXI wieku nowość w nauczaniu cudzoziemców naszego języka, jak również zjawisko bezprecedensowe w całych dziejach tego nauczania (Miodunka 2013b: 53–54, 66). Na tle rozmaitych form oceny znajomości języków obcych certyfikacje wyróżniają się między innymi wielką wartością symboliczną (Barni 2007: 333). Instytucjonalne potwierdzenie stopnia biegłości językowej posiada istotny wymiar społeczny (Vedovelli 2007: 12–14) i nierzadko wpływa na losy jednostek, stąd testy certyfikacyjne zaliczane są do tzw. egzaminów doniosłych (Niemierko 1999: 248). Do egzaminów certyfikacyjnych z języka polskiego, przeprowadzanych w latach 2004–2015 na trzech poziomach zaawansowania: B1 (podstawowym), B2 (średnim ogólnym) i C2 (zaawansowanym), tylko w pierwszym jedenastolecu istnienia certyfikacji przystąpiło ponad dziesięć tysięcy kandydatów<sup>1</sup>.

Opanowanie fonetyki, gramatyki, leksyki i ortografii nie jest w nauce języka celem samym w sobie, o czym pamiętali twórcy certyfikacji, definiujący znajomość języka polskiego jako „umiejętność rozumienia współczesnej polszczyzny pisanej i mówionej oraz umiejętność używania języka polskiego w mowie i piśmie” (*Państwowe Egzaminy Certyfikacyjne...* 2003: 11). Kształt egzaminu – w jego pięciu częściach sprawdzano kolejno: rozumienie ze słuchu, poprawność gramatyczną, rozumienie tekstów pisanych, pisanie i mówienie – zwiększył znaczenie sprawności językowych w nauczaniu obcokrajowców, a jednocześnie pozwalał na podjęcie

\* malgorzata.banach@uj.edu.pl

1 Zob. <http://certyfikatpolski.pl/> (dostęp: 14 grudnia 2017). Od 2016 r. egzaminy certyfikacyjne przeprowadzane są w nieco innym, zreformowanym systemie, regulowanym nową ustawą o języku polskim oraz nowymi rozporządzeniami ministerialnymi.

badań nad stopniem opanowania tych sprawności przez osoby ubiegające się o certyfikat (Miodunka 2013a: 22; 2013b: 66)<sup>2</sup>.

Ponieważ ekspozycja na język jest warunkiem rozwoju kompetencji uczących się, wśród sprawności (czy też w nowym ujęciu – działań) językowych bardzo ważną rolę w procesie nauki odgrywa recepcja pisemna, czyli rozumienie tekstów pisanych (zob. Janowska 2010: 568; Seretny, Lipińska 2005: 188; Chodkiewicz 2016: 71). Praktyka glottodydaktyczna pokazuje zarazem, że czytanie ze zrozumieniem niejednokrotnie jest dla uczących się zadaniem trudnym (Janowska 2010: 561; Czetwertyńska 2004: 48).

Niniejszy artykuł przedstawia w syntetyczny sposób wyniki badań przeprowadzonych przez autorkę w ramach pracy doktorskiej napisanej pod kierunkiem prof. dr. hab. Władysława Miodunki i obronionej w 2015 roku (zob. Banach 2015). Analizy te zmierzały do odpowiedzi na pytanie o umiejętności zdających egzamin certyfikacyjny na poziomie B2<sup>3</sup> w latach 2005–2008 w zakresie czytania ze zrozumieniem w języku polskim, a także o sposób testowania rozumienia tekstów pisanych stosowany na tym egzaminie w badanym okresie<sup>4</sup>. Zachętą do podjęcia tej problematyki było zarówno przekonanie o znaczeniu sprawności czytania dla rozwoju znajomości języka, jak i konstatacja, że wśród zdających egzamin rozumienie tekstów pisanych, testowane w części C egzaminu, bywało w pewnych okresach sprawnością deficytową (Miodunka 2013a: 20; 2013b: 59). Badanie zostało jednak podjęte także z myślą o przyszłości, poszukiwanie nowych rozwiązań i ciągle ulepszanie metod testowania stanowi bowiem integralną część ewolucji systemów certyfikacyjnych.

## 2. Rozumienie tekstów pisanych jako sprawność, proces i produkt. Poziomy rozumienia – zarys problematyki

Mówiąc o rozumieniu tekstów pisanych, warto pamiętać, że termin ten stosowany jest zarówno jako określenie pewnej ogólnej kompetencji, sprawności, którą użytkownicy języka opanowali na danym poziomie, jak i w odniesieniu do konkretnych jej przejawów podczas czynności czytania (Chodkiewicz 2016: 80). W analizach tej czynności można skupić się na procesie rozumienia zachodzącym w jej trakcie lub na jej produkcie, czyli stopniu zrozumienia przetwarzanego tekstu przez osobę czytającą (zob. Alderson 2005: 3–7; Chodkiewicz 2003; Koda 2010: 228–229; Pawłowska 1992: 14–41).

W definicjach procesu czytania akcentowana jest kompleksowa natura tego zjawiska o charakterze językowym i poznawczym (Krasowicz-Kupis 1999: 19; Dakowska 2001: 92–97; Piasecka 2000: 256; Alderson 2005: 1–3), wymagająca interdyscyplinarnego podejścia do jego badania i opisu (zob. Chodkiewicz 2016: 79–80; Pawłowska 1992: 18). Operacje zachodzące

2 Wyniki egzaminów certyfikacyjnych uwzględniła w swoich badaniach nad sprawnością rozumienia ze słuchu w języku polskim jako obcym Adriana Prizel-Kania (2011). Materiał zgromadzony w procesie certyfikacji był też analizowany np. w badaniach porównawczych kompetencji polskich maturzystów i cudzoziemców zdających egzamin certyfikacyjny na poziomie zaawansowanym (zob. Miodunka, Przechodzka 2006, 2007).

3 Do 2011 r. był to najczęściej wybierany poziom na egzaminach certyfikacyjnych z języka polskiego (zob. Miodunka 2013a: 18).

4 Badania nad sprawnością czytania w języku polskim jako obcym poza kontekstem certyfikacji prowadziły np. Tamara Czerkies (2012) i Dominika Bucko (2014).

w trakcie czytania dzieli się często na tzw. procesy wyższego i niższego rzędu. Pierwsze z nich polegają na identyfikacji znaków graficznych i rozpoznawaniu słów, przetwarzaniu informacji składniowych i budowaniu znaczenia na poziomie fraz i zdań (czasem wymienia się wśród tych procesów także aktywizację pamięci operacyjnej, umożliwiającej przetwarzanie danych); procesy wyższego rzędu z kolei związane są z konstruowaniem znaczenia na poziomie dyskursu, wyciąganiem wniosków, interpretacją przekazu (Grabe, Stoller 2002: 20; Grabe 2010; por. Chodkiewicz 2016: 80; Seretny 2010: 551–552; Piasecka 2011: 260–261). W powstających w różnych okresach modelach psycholingwistycznych procesu czytania odmiennie rozkładano akcenty, jeśli chodzi o kolejność i znaczenie poszczególnych poziomów przetwarzania oraz możliwość współdziałania między nimi. Obecnie sporą popularnością cieszą się modele zakładające ciągłą interakcję przetwarzania typu *dół-góra* (związanego z identyfikacją danych płynących z tekstu) i typu *góra-dół* (polegającego na wykorzystywaniu przez czytelnika posiadanej wiedzy, także tej uzyskanej już podczas lektury) (zob. Dakowska 2001: 100–105), choć jednocześnie coraz częściej powraca się do podkreślania roli sprawnego dekodowania w procesie rozumienia (zob. Grabe 2010: 102–103; Wolf i in. 2005: 456–457; Alderson 2005: 18, 20; Koda 2010: 39; Chodkiewicz 2016: 81).

W glottodydaktyce czytanie bywa opisywane (niejednokrotnie w nawiązaniu do teorii psycholingwistycznych) w kategoriach składających się na nie umiejętności cząstkowych (*sub skills, reading skills*)<sup>5</sup>, co ma praktyczne znaczenie dla rozwijania i testowania tej sprawności. Testując recepcję pisemną, dużą wagę przykładana się do umiejętności określanych niekiedy odrębnym mianem poziomów rozumienia; warto również podkreślić, że zdecydowanie dominuje tu perspektywa rozumienia jako produktu, nie da się bowiem zaobserwować i ocenić procesów, które zachodzą w umyśle osób piszących test (zob. Alderson 2005: 7–9, 20; Chodkiewicz 2003)<sup>6</sup>.

Przykłady poziomów rozumienia odnajdziemy w *Europejskim systemie opisu kształcenia językowego* (ESOKJ), którego autorzy wyliczają, że osoba czytająca może sięgać po tekst po to, by:

- ogólnie zrozumieć jego treść,
- uchwycić konkretne informacje,
- zrozumieć szczegółowo całą informację,
- wyciągnąć wnioski (ESOKJ 2003: 70).

Pierwsze trzy pozycje z powyższej listy funkcjonują również w literaturze przedmiotu pod nazwą rozumienia globalnego, selektywnego i szczegółowego (zob. Seretny, Lipińska 2005: 194, 201–202)<sup>7</sup>. Hanna Komorowska wśród umiejętności, jakie powinni opanować uczący się języka, wymienia dodatkowo „rozumienie wewnętrznej organizacji tekstu” oraz „umiejętność wydobycia i spożytkowania zdobytych informacji” (Komorowska 2002a: 144).

5 Analizując listy tych umiejętności tworzone przez poszczególnych autorów, można jednak dostrzec, że istnieje między nimi wiele różnic zarówno ilościowych, jak i jakościowych (zob. Banach 2015).

6 Zainteresowanie samym procesem czytania w glottodydaktyce wiąże się np. z badaniem i rozwijaniem strategii rozumienia tekstów czy form oceny alternatywnej (Chodkiewicz 2003; zob. Banach 2016).

7 Niekiedy rezygnuje się z dokonywania rozróżnień pomiędzy rozumieniem selektywnym i szczegółowym, mówi się wtedy o rozumieniu globalnym i szczegółowym (zob. Komorowska 2002a: 144).

Do tych koncepcji nawiązywali twórcy testów certyfikatowych z języka polskiego. Opis kompetencji oczekiwanej w zakresie rozumienia tekstów od zdających egzamin na poziomie B2 w analizowanym okresie odnaleźć można w *Standardach wymagań egzaminacyjnych (Państwowe Egzaminacje Certyfikatowe... 2003: 35)*, natomiast rozwinięcie opisu ze *Standardów...* i praktyczne wskazówki dla kandydatów przynosił dodatkowo zbiór zadań z czytania opublikowany na stronie internetowej certyfikacji<sup>8</sup>. W publikacji tej zaznaczono, że w części C egzaminu na poziomie średnim ogólnym testowaniu podlega rozumienie zasadniczej myśli tekstu (rozumienie globalne), wyszukiwanie i rozumienie konkretnych informacji (rozumienie selektywne), rozumienie pojedynczych słów bądź wyrażzeń (rozumienie szczegółowe), a także „umiejętność śledzenia strukturalnej i narracyjnej spójności tekstu”<sup>9</sup>.

### 3. Opis badań

#### 3.1. Grupa badawcza

W latach 2005–2008 do egzaminu certyfikatowego z języka polskiego jako obcego na poziomie B2 przystąpiło 516 osób pochodzących z 35 krajów, w tym najwięcej z Polski (o certyfikat mogą ubiegać się obywatele polscy zamieszkali na stałe za granicą), Niemiec, Ukrainy i Stanów Zjednoczonych. Deklaracje kandydatów odnośnie do kraju pochodzenia wypadła jednak traktować z pewną ostrożnością<sup>10</sup>. Analiza danych na temat kraju i języka zdających pokazała dodatkowo, że 96 osób (blisko 19 procent badanych) jednocześnie identyfikowało się z Polską jako jedynym lub jednym z dwóch krajów oraz podało język polski jako jedyny język. Największa grupa zdających – 349 kandydatów (czyli niemal 68 procent badanych) – wskazała wyłącznie kraj inny niż Polska i język inny niż polski. Najmniejszą grupę (71 osób) tworzyli kandydaci, których deklaracje nie pozwalały zaliczyć ich do żadnej z powyższych grup<sup>11</sup>.

Wśród zdających przeważały osoby młode i bardzo młode, gdyż niemal 82 procent kandydatów (422 osoby) to osoby w wieku poniżej 30 lat. Pozostali kandydaci to nieco częściej osoby w wieku 30–39 lat (ok. 11 procent) niż powyżej 40 lat (ok. 7 procent rozwiązujących test). Badana populacja składała się w zdecydowanej większości z kobiet (333 kobiety, czyli ponad 64 procent zdających).

#### 3.2. Analizowany materiał

W badaniu analizie poddano teksty i oparte na nich zadania, które pojawiły się w 19 testach użytych w latach 2005–2008 w części C egzaminu certyfikatowego z języka polskiego jako obcego na poziomie B2, jak również testy jako całości oraz odpowiedzi udzielone w części C

8 Zbiór ten dostępny jest także pod adresem: <http://certyfikatpolski.pl/dla-zdajacych/zbiory-zadan/> (dostęp: 10 marca 2018).

9 Zob. [http://certyfikatpolski.pl/wp-content/uploads/2017/03/6\\_B2\\_RT.pdf](http://certyfikatpolski.pl/wp-content/uploads/2017/03/6_B2_RT.pdf), s. 3 (dostęp: 10 marca 2018).

10 Jak wskazuje W. Miodunka (2013a: 19), kandydaci podający podczas egzaminu jako kraj pochodzenia Polskę to zarówno Polacy mieszkający za granicą, związani bardziej z naszym krajem niż z miejscem osiedlenia, jak i obcokrajowcy zamieszkali w Polsce, identyfikujący się z nowym miejscem zamieszkania.

11 Byli to zdający o polskiej identyfikacji (lub osoby, które podały Polskę jako jedno z dwóch państw) wskazujący język inny niż polski albo język polski jako jeden z dwóch języków, a także osoby wskazujące kraj inny niż Polska, ale deklarujące posługiwanie się językiem polskim.

przez zdających, utrwalone na 516 arkuszach egzaminacyjnych (zob. tab. 1). Dodatkowy materiał stanowiło 50 arkuszy testowych wypełnionych przez grupę kontrolną, złożoną z rodzimych użytkowników języka polskiego zamieszkałych i kształcących się w Polsce.

Jednym z celów przyświecających autorom ESOKJ było umożliwienie Europejczykom „porównywania i wzajemnego uznawania kwalifikacji językowych” (ESOKJ 2003: 17). Twórcy polskiej certyfikacji odwoływali się do opisów biegłości językowej zawartych w ESOKJ oraz do standardów grupy ALTE, stowarzyszenia skupiającego europejskie instytucje zajmujące się testowaniem znajomości języków narodowych.

Ten szerszy kontekst miał wpływ na przyjęcie w badaniu podejścia porównawczego<sup>12</sup>. W analizie tłem dla polskiego materiału stały się analogiczne testy rozumienia tekstów pisanych z języka angielskiego i włoskiego, pochodzące z egzaminów certyfikatowych: *First Certificate in English* (FCE) oraz *Certificato di Conoscenza della Lingua Italiana Livello 3* (CELI 3). Oba te egzaminy, podobnie jak państwowy egzamin z języka polskiego jako obcego na poziomie średnim ogólnym, oznaczane są jako reprezentujące poziom 3 skali ALTE, który odpowiada poziomowi B2 w ESOKJ.

Do analizy porównawczej wybrano cztery testy angielskie z części *Reading* egzaminu FCE i cztery testy włoskie z części *Comprensione della lettura* egzaminu CELI 3. Siedem testów miało taki format, jaki obowiązywał na tych egzaminach m.in. w latach 2005–2008, jeden test natomiast stanowił przykład formuły stosowanej na egzaminie Cambridge w latach 2008–2014.

Badanie w zasadniczej części dotyczyło sytuacji zastanej, a podstawowy materiał, który poddano analizie, został pierwotnie wytworzony w procesie certyfikacji znajomości języka, a więc w innym celu niż na potrzeby tego badania. Czynniki te miały wpływ na dobór procedur badawczych i sposób interpretacji wyników.

Tabela 1. Analizowany materiał

EGZAMIN	PL-B2	FCE	CELI 3
Liczba testów	19 testów z części C egzaminu	4 testy z części <i>Reading</i>	4 testy z części <i>Comprensione della lettura</i>
Liczba tekstów stanowiących podstawę zadań	– 26 dłuższych tekstów – 76 tekstów o długości akapitu – dwa zestawy pojedynczych napisów	15	16
Liczba zadań	– 36 zadań – 94 samodzielne jednostki testu typu „wielokrotny wybór”	15	16
Liczba wypełnionych arkuszy	– 516 arkuszy egzaminacyjnych – 50 arkuszy testowych wypełnionych przez grupę kontrolną	—	—

12 Podobny charakter miały analizy A. Prizel-Kani (2011).

### 3.3. Procedury badawcze

Badanie miało charakter wieloetapowy, stworzenie aparatu badawczego – wykorzystującego zarówno metody jakościowe, jak i ilościowe – wymagało odwołania się do polskich i zagranicznych publikacji z zakresu glottodydaktyki (ze szczególnym uwzględnieniem opracowań na temat testowania znajomości języka) oraz językoznawstwa<sup>13</sup>.

Pierwszym etapem badania była analiza porównawcza tekstów polskich i zagranicznych pojawiających się w analizowanych testach, mająca na celu udzielenie odpowiedzi na pytanie o: ich rodzaj i obecność w teście informacji na ich temat (takich jak: tytuł, źródło, autor, data publikacji), ich długość, tematykę, stopień zanurzenia w lokalnym kontekście kulturowym, dominujące funkcje, stopień złożoności składniowej oraz warstwę leksykalną. Za wskaźnik złożoności syntaktycznej analizowanych tekstów przyjęto średnią liczbę wyrazów w zdaniu<sup>14</sup>. Obliczono także odsetek zdań złożonych w poszczególnych grupach tekstów. Przedmiotem analizy leksykalnej było przede wszystkim słownictwo tekstów polskich. W badaniu za punkt odniesienia przyjęto listy uporządkowane metodą gron wyrazowych w trzech przedziałach częstości przez Annę Seretny (2011). W ramach analizy sprawdzano, jaki odsetek słów w tekstach pochodził spoza tych list. Aby dokładniej je scharakteryzować pod względem leksykalnym, wskazano również procentowy udział w tekstach nazw własnych i ich derywatów. Odsetek nazw własnych obliczono też dla tekstów zagranicznych.

Kolejny etap stanowiła prowadzona w ujęciu porównawczym analiza jakościowa zadań, w której trakcie zwrócono uwagę m.in. na zastosowane w nich typy jednostek, liczbę jednostek w zadaniach, specyficzne cechy konstrukcji zadań i testowane poziomy rozumienia. Analiza jakościowa polskich i zagranicznych testów jako całości miała zaś na celu opis i porównanie: ich budowy, łącznej długości tekstów, czasu przeznaczanego na rozwiązanie komponentu, sposobu formułowania instrukcji, przyjętej punktacji oraz stosowanego typu oceniania.

Do ustalenia kompetencji badanej populacji w zakresie rozumienia tekstów pisanych posłużyła analiza statystyczna odpowiedzi udzielonych przez badaną grupę zdających w części C egzaminu certyfikatowego z języka polskiego jako obcego na poziomie B2. Analizę przeprowadzono na poziomie komponentu, zadań i ich jednostek. Średnie wyniki kandydatów w części C zestawiono ponadto z ich średnimi rezultatami uzyskanymi w pozostałych komponentach egzaminu. W tej fazie badania zasadniczą rolę odgrywała analiza ilościowa polskich zadań. Za jedną z ważniejszych miar przyjęto wskaźnik łatwości<sup>15</sup>. Poszukiwano odpowiedzi na pytanie o czynniki wpływające na trudność zadań dla zdających. W identyfikacji tych czynników pomocna była również analiza na poziomie jednostek tworzących zadania. Dotyczyła

13 Były to m.in.: *Standardy wymagań egzaminacyjnych (Państwowe Egzaminacje Certyfikatowe...* 2003); *Listy zagadnień kontrolnych do analizy testów biegłości językowej* 2004; ESOKJ 2003; *Przewodnik dla autorów zadań do testów językowych* 2004; Bachman 1990; Bachman, Palmer 2000; Alderson 2005; Weir 2005; Komorowska 2002b; Seretny 2004, 2011; Żydek-Bednarczuk 2005; Prizel-Kania 2011.

14 Bardziej złożone wskaźniki czytelności tekstu nie mogły być jednocześnie zastosowane do analizy wszystkich trzech grup tekstów.

15 Wyraża on stosunek liczby punktów zdobytych w zadaniu/jednostce przez osoby rozwiązujące test do maksymalnej liczby punktów, którą grupa ta mogła uzyskać w danym zadaniu/jednostce, i przyjmuje wartości od 0 do 1 (zob. Niemierko 1999: 152–153).



ona zadań rozwiązywanych przez grupy liczące przynajmniej dziesięć osób. W badaniu dla jednostek takich zadań obliczone zostały wskaźnik łatwości oraz licznik wszystkich błędnych odpowiedzi i braku odpowiedzi; w opisie skoncentrowano się na jednostkach, które okazały się dla rozwiązujących je kandydatów najtrudniejsze, oraz na błędnych odpowiedziach o dużej frekwencji.

Analizując arkusze egzaminacyjne, poszukiwano także odpowiedzi na pytanie o to, które jednostki różnicowały lepiej i gorzej czytających. Ze względu na specyfikę badanego materiału badanie mocy różnicującej jednostek przeprowadzono dla dwóch testów. Wskaźnik mocy różnicującej jednostek został obliczony za pomocą metody E1-3.

Ostatnim etapem analizy była weryfikacja niektórych danych przez badanie w grupie kontrolnej, złożonej z rodzimych użytkowników języka polskiego zamieszkałych w Polsce. W badaniu tym wzięło udział 50 uczniów z dwóch klas maturalnych I Liceum Ogólnokształcącego w Bochni. W badaniu kontrolnym użyte zostały dwa testy złożone z 40 jednostek; każdy z testów wypełniało 25 osób. Na ich rozwiązanie maturzyści mieli – podobnie jak zdający egzamin – 45 minut. Testy zawierały w sumie osiem zadań (były to całe zadania egzaminacyjne lub ich fragmenty), które z różnych względów okazały się trudne dla ubiegających się o certyfikat; w obu testach znalazły się pewne zadania wspólne, pełniące funkcję kotwiczącą. Analiza wyników na poziomie zadań i ich jednostek miała dostarczyć dodatkowych wskazówek, pozwalających wnioskować pośrednio o źródłach trudności, jakie podczas pracy z testem mogli napotkać zdający egzamin. Maturzystów poproszono ponadto o subiektywną ocenę trudności rozwiązywanych zadań.

## 4. Wyniki przeprowadzonych badań

### 4.1. Wyniki analizy tekstów

Najważniejszym źródłem tekstów polskich, które w latach 2005–2008 pojawiły się w części C egzaminu certyfikatowego na poziomie B2, była prasa (lub Internet), co zbliża polskie egzaminy do testów włoskich i angielskich. Wyłącznie na polskim egzaminie testowano rozumienie pojedynczych napisów i tekstów o długości akapitu. W odróżnieniu od autorów zagranicznych autorzy zadań polskich – zwłaszcza na początku objętego badaniem okresu – sięgali między innymi do krótkich tekstów użytkowych (np. reklam, horoskopów, informacji na etykietach). W analizowanym materiale znalazły się również fragmenty książek, w tym fragmenty literackie, w korpusie tekstów polskich odgrywały one jednak niewielką rolę.

Inaczej niż w wypadku testów zagranicznych, gdzie standardem było zamieszczanie tekstów wraz z tytułem oraz podanie ogólnej informacji na temat rodzaju tekstu (w testach FCE) albo jego pełnego adresu bibliograficznego (w testach CELI 3), teksty pojawiające się w analizowanych testach polskich opisywane były dość niekonsekwentnie (od braku jakichkolwiek informacji przez podanie samego tytułu aż do umieszczenia pod tekstem notki bibliograficznej).

Autorzy polskiego systemu certyfikatowego przyjęli rozwiązanie polegające na sprawdzaniu w jednym teście rozumienia większej liczby tekstów różnej długości (zob. tab. 2). Na egzaminach angielskim i włoskim testowano natomiast rozumienie niewielkiej liczby stosunkowo

długich tekstów. Porównanie długości tekstów polskich i zagranicznych wykazało, że nawet dłuższe teksty polskie (średnio liczyły one nieco ponad 272 wyrazy) były zdecydowanie krótsze od tekstów angielskich oraz na ogół krótsze od tekstów włoskich (ich średnia długość to odpowiednio: ok. 600 i 421 wyrazów).

Tabela 2. Długość analizowanych tekstów – ujęcie zbiorcze

TEKSTY Z EGZAMINU		ŚREDNIA LICZBA WYRAZÓW W TEKŚCIE	MINIMALNA LICZBA WYRAZÓW	MAKSYMALNA LICZBA WYRAZÓW
PL-B2	Napisy	4,5	2	12
	Akapity	42,9	14	72
	Dłuższe teksty	272,4	94	407
FCE		599,7	498	695
CELI 3		421,4	231	733

W tekstach polskich pojawiały się także zdania statystycznie krótsze od zdań w tekstach zagranicznych (różnica ta dla tekstów dłuższych wynosiła ok. 6 wyrazów) (zob. tab. 3). W porównaniu z tekstami angielskimi i włoskimi materiał polski charakteryzował się również zdecydowanie mniejszym procentowym udziałem zdań złożonych (było ich w dłuższych tekstach polskich 48,5 procent, natomiast w materiale angielskim i włoskim odpowiednio: 69 i 66 procent). Może to wskazywać, że teksty polskie były w większym stopniu modyfikowane i upraszczane składniowo na potrzeby zadań.

Analiza tekstów polskich ujawniła, że poruszają one problematykę należącą do wielu kręgów tematycznych (pod tym względem podobne były do nich teksty włoskie, podczas gdy teksty angielskie koncentrowały się wyraźniej wokół kilku zagadnień). Teksty polskie okazały się jednak bardziej abstrakcyjne i mniej osobiste od tekstów angielskich i włoskich, duże znaczenie miała też w wypadku wielu z nich funkcja informacyjna, podczas gdy w swobodniejszych pod względem struktury tekstach zagranicznych częściej prezentowano osobiste przekonania bohaterów i autora, co sprawiało, że częściej dochodziła w nich do głosu funkcja ekspresywno-emocjonalna. W każdej grupie tekstów pojawiły się wypowiedzi o dominującej funkcji impresyjnej. Wśród tekstów angielskich trafiały się też dłuższe fragmenty literackie o dominującej funkcji estetycznej.

Wśród tekstów polskich były zarówno fragmenty neutralne kulturowo, jak i fragmenty mocno osadzone w kulturze i realiach polskich. Realia odnaleźć można było przede wszystkim w nagłówkach prasowych i tekstach o długości akapitu, podczas gdy dłuższe teksty podejmowały często tematy należące do kultury wysokiej, a ich bohaterowie to zwykle Polacy reprezentujący świat literatury, nauki, sztuki. Pod względem obecności elementu kulturowego teksty polskie bliższe były tekstom włoskim, te ostatnie jednak koncentrowały się na realiach życia



włoskiego, a ich bohaterami byli głównie zwykli mieszkańcy Półwyspu Apenińskiego. Duża różnica występowała natomiast między tekstami polskimi a angielskimi, gdyż w tych ostatnich przedstawiano przeważnie zwykłych Brytyjczyków lub znane postacie należące do świata popkultury; teksty angielskie cechowały się przy tym na ogół największą neutralnością kulturową w całym analizowanym materiale. Obserwacje na temat warstwy kulturowej tekstów potwierdziła analiza leksykalna – w tekstach angielskich procentowy udział nazw własnych był niższy niż w tekstach polskich czy włoskich (zob. tab. 4).

Tabela 3. Złożoność składniowa analizowanych tekstów – ujęcie zbiorcze

TEKSTY Z EGZAMINU		ŚREDNIA LICZBA WYRAZÓW W ZDANIU	ODSETEK ZDAŃ ZŁOŻONYCH W POSZCZEGÓLNYCH GRUPACH TEKSTÓW
PL-B2	Napisy	4,5	9%
	Akapity	12,5	53%
	Teksty dłuższe	10,5	48,5%
FCE		16,5	69%
CELI 3		16,7	66%

Tabela 4. Udział nazw własnych w analizowanych tekstach – ujęcie zbiorcze

TEKSTY Z EGZAMINU		NAZWY WŁASNE (%)
PL-B2	Napisy	5
	Akapity	3,6
	Teksty dłuższe	3,7
FCE		2,4
CELI 3		3,7

Jeśli chodzi o obecność wyrazów spoza list leksykalnych zaproponowanych przez A. Seretny, analiza trzech grup polskich tekstów (pojedynczych napisów, tekstów o długości akapitu i dłuższych tekstów) wykazała, że w każdej z tych grup po wyłączeniu z obliczeń nazw własnych i ich derywatów słowa o niskiej frekwencji stanowiły kilkanaście procent słownictwa (zob. tab. 5)<sup>16</sup>.

16 W większości dłuższych tekstów traktowanych indywidualnie wyrazy spoza trzech przedziałów częstości pokrywały ponad 10 procent tekstu (w kilku tekstach było to nawet ponad 20 procent), choć po odjęciu nazw własnych i ich derywatów proporcja ta ulegała czasem zmianie.

Za element przyczyniający się do zwiększenia trudności badanych tekstów należy uznać obecne w nich wyrażenia idiomatyczne, natomiast internacjonalizmy czy wyrazy spokrewnione słowotwórczo ze słowami pojawiającymi się na listach leksykalnych mogły być postrzegane przez część kandydatów jako łatwiejsze, niż sugerowałyby to wyłącznie dane liczbowe.

Tabela 5. Procentowy udział słów o niskiej frekwencji w tekstach polskich – ujęcie zbiorcze

GRUPA TEKSTÓW	SŁOWNICTWO SPOZA TRZECH PRZEDZIAŁÓW CZĘSTOŚCI (%)	SŁOWNICTWO SPOZA TRZECH PRZEDZIAŁÓW CZĘSTOŚCI BEZ NAZW WŁASNYCH (%)
Napisy	21	16
Akapyty	17,2	13,6
Dłuższe teksty	14,7	11

#### 4.2. Wyniki analizy jakościowej zadań i testów

W każdym z analizowanych w badaniu systemów certyfikatowych przyjęto nieco odmienne rozwiązania w zakresie testowania sprawności rozumienia tekstów pisanych. Największą różnorodnością pod względem wykorzystywanych technik testujących charakteryzował się analizowany materiał polski (pojawiało się w nim sześć typów zadań), a najmniejszą – włoski (przy konstruowaniu testów stosowano jedynie trzy techniki) (zob. tab. 6). Podobnie jak zadania angielskie, wszystkie zadania polskie opierały się na pytaniach zamkniętych, wymagających od kandydatów nie samodzielnego formułowania odpowiedzi, lecz jej wyboru (pytania otwarte znalazły się natomiast w testach włoskich).

Jeśli chodzi o materiał polski, często te same zadania pojawiały się w co najmniej dwóch testach (wprowadzano do nich również poprawki). Modyfikacje dokonywane przez autorów polskich testów w kolejnych latach analizowanego okresu polegały także na rezygnacji z pewnego typu jednostek czy sięgnięciu po inne techniki testujące.

Tabela 6. Techniki testujące rozumienie tekstów pisanych w testach PL-B2, FCE i CELI 3

EGZAMIN	PL-B2	FCE	CELI 3
Typ jednostek testu	<ul style="list-style-type: none"> <li>- Wielokrotny wybór informacyjny</li> <li>- Prawda/fałsz/brak informacji (P/F/B1)</li> <li>- Wielokrotny wybór leksykalny</li> <li>- Luki zdaniowe</li> <li>- Ustalanie kolejności</li> <li>- Dobieranie</li> </ul>	<ul style="list-style-type: none"> <li>- Wielokrotny wybór informacyjny</li> <li>- Luki zdaniowe</li> <li>- Dobieranie</li> <li>- Dobieranie wielokrotne</li> </ul>	<ul style="list-style-type: none"> <li>- Wielokrotny wybór informacyjny</li> <li>- Dobieranie wielokrotne</li> <li>- Krótka odpowiedź</li> </ul>

Porównanie zadań polskich z zadaniami angielskimi i włoskimi tego samego typu wykazało, że często tej samej techniki testującej w różnych systemach używano w nieco odmienny sposób, sprawdzając za jej pomocą na przykład różne poziomy rozumienia. Tylko w testach polskich pojawił się ponadto typ zadania ukierunkowany bezpośrednio na testowanie znajomości leksyki (zadania z jednostkami typu „wielokrotny wybór leksykalny”). Zadania polskie różniły się nie tylko od swoich zagranicznych odpowiedników, niejednorodne wewnętrznie (np. pod względem liczby jednostek w zadaniu, liczby opcji do wyboru w jednostce, konkretnych rozwiązań technicznych czy graficznych) okazały się również poszczególne grupy analizowanych zadań polskich. W porównaniu z polskim materiałem niektóre typy zadań angielskich i włoskich wykazywały większą jednorodność konstrukcyjną.

W materiale polskim znajdowały się zadania testujące rozumienie globalne akapitów, rozumienie konkretnych informacji zawartych w tekstach, rozumienie na poziomie pojedynczych wyrażen i struktur, a także umiejętność śledzenia struktury tekstu. Inaczej niż na egzaminie FCE testowaniu nie podlegało rozumienie dłuższych tekstów jako całości, określanie ich tonu czy intencji autora. Zbliżało to zadania polskie do zadań włoskich, które również nie zawierały pytań tego typu.

Niewielka długość tekstów stanowiących podstawę polskich zadań oraz stosunkowo duża liczba jednostek w zadaniach sprzyjały dokładnej lekturze tekstów. Analizowane zadania zagraniczne opierały się zwykle na tekstach dłuższych od tekstów polskich i składały się na ogół z mniejszej liczby jednostek niż zadania polskie. Mogło to zachęcać kandydatów do bardziej strategicznej lektury – np. szybkiego przeszukiwania tekstu w celu znalezienia potrzebnej informacji.

Tabela 7. Wyniki analizy komponentów – ujęcie zbiorcze

EGZAMIN	PL-B2	FCE	CELI 3
Liczba zadań	4 lub 5	4 (1996–2008) 3 (2008–2014)	4 (w 3 sekcjach)
Liczba jednostek	40	35 (1996–2008) 30 (2008–2014)	23
Czas przeznaczony na rozwiązanie testu	45 minut	75 minut (1996–2008) 60 minut (2008–2014)	Czytanie i pisanie: 135 minut
Średnia długość testu	1178 wyrazów	2249 wyrazów	1686 wyrazów

W odróżnieniu od testów czytania z egzaminów FCE i CELI 3 polski komponent sprawdzający rozumienie tekstów pisanych nie miał stałego schematu. Analizowane testy polskie i zagraniczne składały się z podobnej liczby części i zadań (zob. tab. 7). Komponent polski od zagranicznych odróżniały: krótszy czas przeznaczony na jego rozwiązanie, większa liczba

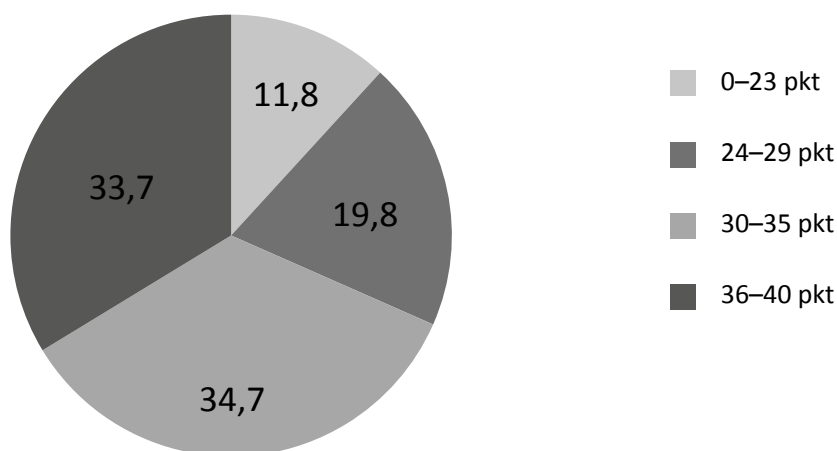
jednostek, mniejsza łączna długość tekstów w teście, identyczna punktacja wszystkich jednostek, brak ważenia wyników oraz większa transparentność oceniania (tylko w polskich testach podawano zawsze informacje na temat punktacji poszczególnych zadań). Instrukcje do zadań i przykłady obecne w testach polskich były bardziej związane z tekstem niż instrukcje i przykłady w testach włoskich, ale nie wprowadzały w kontekst w równym stopniu jak instrukcje do zadań z egzaminu FCE.

#### 4.3. Wyniki analizy odpowiedzi zdających

Analiza odpowiedzi kandydatów na poziomie komponentu wykazała, że większość badanych dobrze lub bardzo dobrze opanowała sprawność czytania. W analizowanym materiale dominowały prace dobre i bardzo dobre (zob. wykres 1). Ponad 68 procent badanych uzyskało wynik powyżej 30 punktów, z czego w przedziale 36–40 punktów zmieściło się niemal 34 procent prac. Prace w przedziale punktowym 24–29 to z kolei około 20 procent arkuszy.

Mimo że rozumienie tekstów pisanych było w badanej grupie sprawnością opanowaną na ogół dobrze i bardzo dobrze, to znalazła się w tej populacji pewna liczba kandydatów, których kompetencja w tym zakresie okazała się niska i nie pozwoliła na zdanie egzaminu<sup>17</sup>. Czytanie było drugą po poprawności gramatycznej częścią egzaminu z największą liczbą prac niedostatecznych. Prace poniżej progu zaliczenia, wynoszącego 24 punkty, stanowiły blisko 12 procent wszystkich analizowanych arkuszy.

Wykres 1. Procentowy rozkład wyników badanej grupy w poszczególnych przedziałach punktowych



<sup>17</sup> Warunkiem zaliczenia egzaminu było uzyskanie co najmniej 60 procent możliwych do zdobycia punktów w każdym z jego pięciu komponentów.

Tabela 8. Wyniki badanej grupy na poziomie komponentu – zestawienie podstawowych danych statystycznych

ROK	ŚREDNIA LICZBA PUNKTÓW	ODCHYLENIE STANDAR-DOWE	MEDIANA	NAJNIŻSZY WYNIK	NAJWYŻSZY WYNIK	NAJCZĘSTSZY WYNIK
2005	32,2	7,1	35	11	40	37
2006	33,4	5,6	35	11	40	38
2007	29,5	6,2	30	14	39	34
2008	31,5	6,6	34	2	39	36
2005–2008	31,6	6,4	33,5	2	40	36

Średnia arytmetyczna wyników we wszystkich analizowanych testach łącznie wyniosła 31,6 pkt, mediana – 33,5 pkt, a najczęściej uzyskiwany przez kandydatów rezultat to 36 pkt (zob. tab. 8). W skali całego analizowanego okresu w certyfikатовym teście rozumienia tekstów pisanych statystycznie najlepsi okazali się kandydaci przystępujący do egzaminu w roku 2005 i 2006 (ich średnia to odpowiednio: 32,2 oraz 33,4 pkt). Nieco gorszy wynik (średnio 31,5 pkt) uzyskała grupa zdająca egzamin w 2008 roku, najgorsze rezultaty osiągnęli natomiast zdający w 2007 roku (średnio tylko 29,5 pkt), w którym czytanie było na poziomie B2 jedną z dwóch najsłabiej opanowanych sprawności. Tendencje typowe dla całości analizowanego czterolecia najlepiej odzwierciedlały wyniki ubiegających się o certyfikat w 2008 roku.

Biorąc pod uwagę średnie wyników uzyskanych przez badaną grupę w poszczególnych częściach egzaminu (zob. tab. 9), rozumienie tekstów pisanych okazało się w łącznym zestawieniu za lata 2005–2008 sprawnością statystycznie trudniejszą niż mówienie i rozumienie ze słuchu, ale łatwiejszą od poprawności gramatycznej i redakcji tekstów (pozycja ta jednak nie była identyczna w poszczególnych latach objętych badaniem).

Tabela 9. Średnia wyników w poszczególnych częściach egzaminu certyfikатовego na poziomie B2 w latach 2005–2008<sup>18</sup>

CZĘŚĆ EGZAMINU	RS	PG	RT	P	M
Średnia liczba punktów	32	30,5	31,6	30,1	33,3

Zdający o polskiej identyfikacji (wskazujący Polskę jako swój kraj i język polski jako jedyny język) mieli nad kandydatami o identyfikacji cudzoziemskiej pewną przewagę (zob. tab. 10), ale w skali całego badanego okresu różnice wyników w obu tych grupach nie były znaczne

<sup>18</sup> Przedstawione w tabeli 9 wyniki z pozostałych części egzaminu zostały obliczone na podstawie danych udostępnionych przez sekretariat ówczesnej Państwowej Komisji Poświadczania Znajomości Języka Polskiego jako Obcego. W tabeli przyjęto następujące oznaczenia kolejnych komponentów egzaminu: RS – Rozumienie ze słuchu, PG – Poprawność gramatyczna, RT – Rozumienie tekstów pisanych, P – Pisanie, M – Mówienie.

(przykładowo różnica średnich wyniosła 0,6 pkt, wyniki cudzoziemców charakteryzowały się jednak większym rozproszeniem). Rezultaty te można interpretować w rozmaity sposób. Przede wszystkim należy pamiętać, że nie ma możliwości ustalenia, w jakim wieku wyjechały z kraju osoby wpisujące w odpowiednich rubrykach ankiety Polskę i język polski czy w jakim języku odebrały one wykształcenie. Część kandydatów o cudzoziemskiej identyfikacji narodowej i językowej to z kolei osoby polskiego pochodzenia – sam ten fakt nie przesądza o dobrej znajomości języka polskiego, jakkolwiek w sprzyjających okolicznościach mógł stanowić znaczne ułatwienie. Nie można jednak również pomniejszać sukcesu kandydatów, którzy uzyskali wysokie wyniki na egzaminie, nie będąc Polakami ani nie posiadając polskich korzeni.

Na tym tle statystycznie gorzej wypadali kandydaci z grupy trzeciej – najbardziej niejednorodnej i o najmniej jednoznacznej identyfikacji. Wydaje się, że znaczna ich część to osoby urodzone lub wychowane za granicą, mające wprawdzie polskie pochodzenie, ale znające język polski nieco słabiej.

Tabela 10. Wyniki uzyskane w teście przez osoby o polskiej identyfikacji wskazujące język polski jako jedyny język na tle wyników pozostałych grup

DANE	GRUPA			RAZEM
	POLSKA IDENTYFIKACJA I JĘZYK POLSKI JAKO JEDYNY JĘZYK	NIEPOLSKA IDENTYFIKACJA I JĘZYK/JĘZYKI INNY/INNE NIŻ POLSKI	INNI	
Liczba osób	96	349	71	516
Średnia liczba punktów	32,2	31,6	30,8	31,6
Odchylenie standardowe	5,9	6,5	6,9	6,4

Bardziej szczegółowy obraz kompetencji badanych przyniosła analiza odpowiedzi udzielonych przez nich w poszczególnych zadaniach i ich jednostkach. Wskaźnik łatwości 20 spośród 36 analizowanych zadań (zob. tab. 11) wynosił mniej niż 0,8, przy czym najliczniejszą grupę stanowiły zadania o wskaźniku łatwości w granicach 0,7–0,79 (było ich 15), natomiast dla pięciu najtrudniejszych zadań wskaźnik ten mieścił się w przedziale 0,5–0,69. Wskaźnik łatwości w granicach 0,8–0,89 miało aż 14 zadań (była to więc druga co do wielkości grupa), a tylko dwa – wyższy. Interpretacja wyników była w pewnym stopniu utrudniona, gdyż analizowane zadania w opisywanym okresie rozwiązywały różne grupy zdających, zależnie od testu, w którym zostały wykorzystane<sup>19</sup>.

Rozważając kwestię zależności między stosowaną metodą testowania a wynikami zdających, zaobserwowano jednak pewne tendencje, chociaż nie stanowiły one bezwyjątkowych

19 Aby trafnie ocenić umiejętności kandydatów (a pośrednio czynniki wpływające na trudność poszczególnych zadań), ustalano m.in., jakie wyniki uzyskiwała konkretna grupa na tle innych grup w zadaniach wspólnych dla różnych testów (tzw. zadaniach kotwiczących).



reguł. Większość zadań z jednostkami na dobieranie i typu „prawda/fałsz/brak informacji” miała wskaźnik łatwości niższy niż 0,8 (po dwa zadania z obu tych grup znalazły się jednocześnie wśród pięciu zadań o najniższej rozwiązalności). W grupie zadań z jednostkami wielokrotnego wyboru leksykalnego przewaga liczbowa zadań o wskaźniku łatwości niższym niż 0,8 była nieznacząca<sup>20</sup>, natomiast wśród zadań z lukami zdaniowymi i zadań z jednostkami na ustalanie kolejności (wymagających uporządkowania rozsypanych fragmentów tekstu) dominowały takie, których rozwiązywalność wynosiła przynajmniej 80 procent.

Tabela 11. Wskaźnik łatwości analizowanych zadań – ujęcie zbiorcze

PRZEDZIAŁ WSKAŹNIKA ŁATWOŚCI (FI)	0–0,49	0,5–0,69	0,7–0,79	0,8–0,89	0,9–1
Liczba zadań	0	5	15	14	2

W większości typów zadań zdecydowanie dominowały jednostki o wskaźniku łatwości wynoszącym co najmniej 0,8 (zob. tab. 12). Jedynie w zadaniach typu „prawda/fałsz/brak informacji” i w zadaniach na dobieranie liczba jednostek w przedziale łatwości 0,8–1 nie przekroczyła 50 procent. W najniższym przedziale wskaźnika łatwości 0–0,49 znalazł się pewien odsetek jednostek typu „prawda/fałsz/brak informacji”, jednostek na dobieranie i pytań typu „wielokrotny wybór informacyjny”.

Tabela 12. Rozkład jednostek poszczególnych typów w kolejnych przedziałach wskaźnika łatwości

TYP ZADANIA	LICZBA ANALIZOWANYCH JEDNOSTEK	ODSETEK JEDNOSTEK W PRZEDZIAŁACH ŁATWOŚCI			
		0–0,49	0,5–0,69	0,7–0,79	0,8–1
Wielokrotny wybór informacyjny	94	6,4%	9,6%	13,8%	70,2%
P/F/BI	53	13,2%	18,9%	20,7%	47,2%
Wielokrotny wybór leksykalny	45	0%	8,9%	28,9%	62,2%
Luki zdaniowe	30	0%	6,7%	23,3%	70%
Ustalanie kolejności	40	0%	15%	12,5%	72,5%
Dobieranie	90	7,8%	25,5%	18,9%	47,8%

W analizowanym okresie kandydaci na ogół gorzej radzili sobie zatem z zadaniami typu „prawda/fałsz/brak informacji” i na dobieranie. Przyczyn tego zjawiska można szukać zarówno po stronie kandydatów (mogły to być m.in.: nieznanomość kluczowych słów, brak obycia

20 W wypadku tych zadań w analizie na poziomie jednostek obraz był już nieco inny ze względu na to, że część jednostek wykorzystywano kilkakrotnie w różnych zadaniach opierających się na tym samym tekście.

testowego, nieuważna lektura tekstu lub treści zadania oraz błędy w rozumowaniu), jak i po stronie zadań. Konstrukcja pewnych jednostek typu „prawda/fałsz/brak informacji” sprawiała, że ich rozwiązanie było zadaniem złożonym pod względem kognitywnym. W zadaniach z jednostkami na dobieranie dużą rolę odegrały pewne cechy tekstów uwypuklone przez konstrukcję zadania, a więc podobieństwo tematyczne i leksykalne fragmentów znajdujących się w różnych jednostkach, nie do końca spójna struktura tekstu, ściśle powiązanie następujących po sobie jednostek pod względem treści i struktury, nieobecność informacji ważnych dla zrozumienia przekazu, wreszcie gry słów, których zrozumienie warunkowało wskazanie poprawnego połączenia. Sam fakt wyboru wskazanych technik testujących nie przesądza o stworzeniu zadania, które będzie miało niską rozwiązywalność, jednak może zwiększać takie prawdopodobieństwo.

Wydaje się, że takie właściwości analizowanych tekstów, jak średnia długość zdań czy stopień pokrycia wyrazami o niższej frekwencji nie miały – poza nielicznymi wyjątkami – bezpośredniego i zasadniczego wpływu na rozwiązywalność towarzyszących im zadań. Obecność trudnych elementów leksykalnych w tekście mogła jednak mieć znaczenie wtedy, gdy ich rozumienie było kluczowe dla udzielenia odpowiedzi w danej jednostce.

Na uwagę zasługuje to, że badani stosunkowo często mieli trudności z rozumieniem struktury tekstu. Śledzenie narracyjnej i strukturalnej spójności tekstu sprawiało kandydatom problem przede wszystkim wtedy, gdy rekonstrukcja wymagała operacji na poziomie wyższym niż zdanie (a więc gdy w jednostce brakowało wyraźnych wskazówek składniowych). Z przeprowadzonej na materiale dwóch testów analizy mocy różnicującej jednostek wynikało tymczasem, że umiejętność uchwycenia struktury tekstu była skutecznym prognostykiem dobrego wyniku w całym teście rozumienia. Kandydaci o wysokiej kompetencji w zakresie rozumienia tekstów pisanych zdecydowanie lepiej od grupy słabszej radzili sobie także z dopasowywaniem fragmentów w podobnych tematycznie jednostkach zadań na dobieranie; mieli oni również wyraźną przewagę w odszyfrowywaniu sensu związanych nagłówków prasowych.

#### 4.4. Wyniki badania w grupie kontrolnej

Zadania rozwiązywane przez maturzystów z grupy kontrolnej okazały się dla nich łatwiejsze niż dla zdających egzamin certyfikacyjny; wyniki polskich licealistów były też mniej rozproszone (zob. tab. 13)<sup>21</sup>, co nie stanowiło zaskoczenia. Różnice między rozwiązywalnością poszczególnych zadań w badaniu kontrolnym i podczas egzaminu wahały się jednak od 0,03 do 0,41; w niektórych wypadkach były one zatem nieznaczące.

Pięć z ośmiu zadań miało w grupie kontrolnej wskaźnik łatwości niższy niż 0,8, co świadczy o tym, że nie były one dla Polaków bardzo łatwe. Udzielenie poprawnej odpowiedzi w poszczególnych zadaniach lub jednostkach, które sprawiły trudności zdającym egzamin certyfikacyjny, stanowiło nierzadko wyzwanie także dla uczniów biorących udział w badaniu kontrolnym. Stosunkowo trudne okazały się dla licealistów na przykład zadania z jednostkami typu „prawda/fałsz/brak informacji”, których rozwiązanie wymagało uważnej lektury tekstów i pytań

21 Dwa zadania zastosowane w badaniu kontrolnym (zad. 1 i 3) składały się z połowy pytań wchodzących w skład zadań egzaminacyjnych, dlatego w tych wypadkach dodatkowo ustalono wynik zdających egzamin tylko w jednostkach wybranych do badania kontrolnego i to on był brany pod uwagę w porównaniu.

oraz logicznego myślenia (zad. 1 i 3). W zadaniach z jednostkami na dobieranie wymagających przede wszystkim dobrej znajomości słownictwa i odczytania gier słownych (np. zad. 7 i 8) licealiści na ogół radzili sobie bardzo dobrze. Tam, gdzie połączenie elementów wymagało jednak rekonstrukcji tekstu o luźnej strukturze lub wyboru spośród bardzo zbliżonych opcji (zad. 4 i 5), licealiści nierzadko popełniali błędy. Warto również podkreślić, że opinie poszczególnych maturzystów na temat trudności rozwiązywanych zadań bardzo często nie pokrywały się z uzyskanym przez nich rezultatem punktowym.

Tabela 13. Wyniki zdających egzamin oraz grupy kontrolnej na poziomie zadań

ZADANIE	TYP ZADANIA	ZDAJĄCY EGZAMIN CERTYFIKATOWY			POLSCY MATURZYŚCI		
		WSKAŹ- NIK ŁATWO- ŚCI	ODCHY- LENIE STANDAR- DOWE	MEDIA- NA	WSKAŹ- NIK ŁATWO- ŚCI	ODCHY- LENIE STANDAR- DOWE	MEDIA- NA
1	P/F/BI	0,49	1	3	0,58	0,9	3
2	Wielokrotny wybór informacyjny	22	–	–	0,72	1,1	4
3	P/F/BI	0,7	1	4	0,73	0,6	4
4	Dobieranie	0,71	2,6	8	0,77	2,1	8
5	Dobieranie	0,5	3,4	6	0,79	1,8	8
6	Ustalanie kolejności	0,69	2,5	7	0,82	2,9	10
7	Dobieranie	0,7	3,3	7,5	0,92	1,5	10
8	Dobieranie	0,75	2,7	8	0,97	0,8	10

W związku z tym, że w pewnych zadaniach rodzimi użytkownicy języka okazali się tylko nieznacznie lepsi od zdających egzamin certyfikатовy, można przypuszczać, że problemów kandydatów z rozwiązaniem tych zadań nie da się wyjaśnić tylko ich niższą kompetencją leksykalną czy gramatyczną w języku polskim. Ważnymi czynnikami mogły być również: brak wprawy w rozwiązywaniu danego typu jednostek, nieuważna lektura tekstów lub treści zadań, niewłaściwy tok rozumowania, ale także sama forma zadań lub ich części. Światło na tę ostatnią kwestię rzuciły też obserwacje maturzystów wypowiadających się na temat trudności napotkanych w trakcie rozwiązywania testów. W komentarzach wskazywali oni na przykład na bardzo luźną strukturę tekstów, które musieli rekonstruować, czy duże podobieństwo różnych opcji w zadaniach na dobieranie.

22 Brak wskaźników dla zadania 2 wynika stąd, że składało się z kilku niezależnych jednostek, które podczas egzaminu nigdy nie pojawiły się w takim zestawie i były rozwiązywane przez różne grupy zdających.

## 5. Podsumowanie

Oprócz odpowiedzi na pytanie o kompetencję opisywanej grupy w zakresie sprawności rozumienia tekstów pisanych oraz o sposób testowania tej sprawności na egzaminie certyfikacyjnym badanie przyniosło wnioski na temat możliwości udoskonalania testów certyfikacyjnych oraz wskazówki dla dydaktyki sprawności czytania.

Część postulatów dotyczących ulepszenia testów rozumienia tekstów pisanych na poziomie B2 jest wciąż aktualna. W ramach dalszych prac nad testami czytania warto m.in.: zmniejszyć liczbę jednostek w teście i równocześnie wydłużyć łączną długość tekstów w komponencie (wykorzystując na egzaminie także dłuższe teksty, liczące nawet około 450–500 wyrazów); zrezygnować z testowania pojedynczych, niepowiązanych ze sobą akapitów; stosować jednostki testujące rozumienie globalne dłuższych fragmentów, rozumienie intencji autora czy funkcji tekstu; w doborze tekstów uwzględniać fragmenty książek (np. współczesnych powieści, tytułów dla młodzieży, literatury podróżniczej); umiejętność śledzenia spójności tekstu sprawdzać w tekstach o przejrzystej, logicznej strukturze; rozbudować fazę pilotażu i analizy statystycznej zadań certyfikacyjnych.

W dydaktyce czytania warto poświęcić więcej uwagi rozwijaniu umiejętności rozumienia zasadniczego przesłania, a także struktury tekstu, nie skupiając się wyłącznie na kształceniu zdolności rozumienia szczegółowych informacji. Uczący się powinni zacząć postrzegać teksty jako całości. Praca na lekcji przy rozwiązywaniu zadań wymagających odtworzenia lub śledzenia struktury tekstu winna obejmować także praktyczną – dostosowaną do poziomu wiedzy metajęzykowej uczniów – analizę mechanizmów spójnościowych obecnych w rekonstruowanych tekstach. W nauczaniu recepcji pisemnej duże znaczenie ma rozwijanie umiejętności syntezy informacji, logicznego myślenia i wnioskowania, ale również regulowania sposobu lektury i jej tempa stosownie do wyznaczonego celu. Warto zatem konfrontować uczniów z tekstami długimi i – wprowadzając rozmaite strategie pracy z tekstem – ograniczać czas przeznaczony na wykonanie niektórych zadań w taki sposób, by uniemożliwić czytającym zbyt wolną i szczegółową lekturę. W przygotowaniu do egzaminu natomiast przydatna może się okazać także wiedza na temat strategii egzaminacyjnych czy konstrukcji różnego typu dystraktorów<sup>23</sup>.

Analizy przeprowadzone w ramach opisywanego badania wniosły wkład w rozwój badań nad sprawnością czytania w języku polskim jako obcym i nad testowaniem tej sprawności, nie wyczerpując wszakże omawianej problematyki. Z każdym rokiem działania certyfikacji powiększa się zbiór danych na temat kompetencji osób pragnących uzyskać potwierdzenie znajomości języka polskiego, także na nowych poziomach zaawansowania. Wiedzę zdobytą w wyniku analiz tego materiału warto również pogłębiać w dalszych badaniach o charakterze jakościowym, pozwalających uzyskać odpowiedzi na pytania, które mogą rodzić się w trakcie analiz ilościowych.

23 Są to nieprawidłowe odpowiedzi, podane do wyboru obok odpowiedzi poprawnej (tzw. werstraktora).

## Bibliografia

- Alderson J.C. 2005: *Assessing reading*, Cambridge University Press, Cambridge.
- Bachman L.F. 1990: *Fundamental considerations in language testing*, Oxford University Press, Oxford.
- Bachman L.F., Palmer A.S. 2000: *Language testing in practice: Designing and developing useful language tests*, Oxford University Press, Oxford.
- Banach M. 2015: *Rozumienie tekstu pisanego przez zdających egzamin certyfikacyjny z języka polskiego jako obcego na poziomie B2 w latach 2005–2008*, praca doktorska, Uniwersytet Jagielloński, Kraków.
- Banach M. 2016: *Alternatywne sposoby sprawdzania i oceny recepcji pisemnej w glottodydaktyce*, [w:] I. Janowska, P. Gębal (red.), *O lepsze jutro studiów polonistycznych w świecie. Glottodydaktyka polonistyczna dziś*, Księgarnia Akademicka, Kraków, s. 340–352.
- Barni M. 2007: *Etica e valutazione delle competenze in L2*, [w:] M. Vedovelli (red.), *Manuale della certificazione dell'italiano L2*, Carocci, Roma, s. 329–341.
- Bucko D. 2014: *Strategie czytania hipertekstu podejmowane przez uczących się języka polskiego jako obcego podczas realizacji zadań*, praca doktorska, Uniwersytet Jagielloński, Kraków.
- Chodkiewicz H. 2003: *Czytanie jako proces i produkt: implikacje dla nauczania języków obcych*, [w:] J. Arabski (red.), *Wybrane zagadnienia glottodydaktyki*, Wyższa Szkoła Zarządzania Marketingowego i Języków Obcych w Katowicach, Katowice, s. 11–24.
- Chodkiewicz H. 2016: *Szczególne miejsce czytania w nauce języka obcego*, „Języki Obce w Szkole”, nr 4, s. 79–85.
- Czerkies T. 2012: *Tekst literacki w nauczaniu języka polskiego jako obcego (z elementami pedagogiki dyskursywnej)*, Księgarnia Akademicka, Kraków.
- Czetwertyńska G. 2004: *Ile wart jest ten, kto uczy czytać?*, „Języki Obce w Szkole”, nr 6, s. 48–55.
- Dakowska M. 2001: *Psycholingwistyczne podstawy dydaktyki języków obcych*, Wydawnictwo Naukowe PWN, Warszawa.
- Egzamin certyfikacyjny z języka polskiego jako obcego. Rozumienie tekstów pisanych. B2. Poziom średni ogólny*, Państwowa Komisja ds. Poświadczania Znajomości Języka Polskiego jako Obcego (online: [http://certyfikatpolski.pl/wp-content/uploads/2017/03/6\\_B2\\_RT.pdf](http://certyfikatpolski.pl/wp-content/uploads/2017/03/6_B2_RT.pdf), dostęp: 9 marca 2018).
- ESOKJ: *Europejski system opisu kształcenia językowego. Uczenie się, nauczanie, ocenianie*, Rada Europy, Centralny Ośrodek Doskonalenia Nauczycieli, Warszawa 2003.
- Grabe W. 2010: *Reading in a second language. Moving from theory to practice*, Cambridge University Press, New York.
- Grabe W., Stoller F. 2002: *Teaching and researching reading*, Longman, Harlow.
- Janowska I. 2010: *Tekst i jego rozumienie w podejściu zadaniowym*, [w:] G. Zarzycka, G. Rudziński (red.), *Teksty i podteksty w nauczaniu języka polskiego jako obcego – 2*, „Acta Universitatis Lodzianensis. Kształcenie Polonistyczne Cudzoziemców”, nr 17, Wydawnictwo Uniwersytetu Łódzkiego, Łódź, s. 561–569.
- Koda K. 2010: *Insights into second language reading. A cross-linguistic approach*, Cambridge University Press, New York.
- Komorowska H. 2002a: *Metodyka nauczania języków obcych*, Fraszka Edukacyjna, Warszawa.
- Komorowska H. 2002b: *Sprawdzanie umiejętności w nauce języka obcego. Kontrola – Ocena – Testowanie*, Fraszka Edukacyjna, Warszawa.
- Krasowicz-Kupis G. 1999: *Rozwój metajęzykowy a osiągnięcia w czytaniu dzieci 6–9-letnich*, Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, Lublin.
- Listy zagadnień kontrolnych do analizy testów biegłości językowej*, przeł. M. Gaszyńska, A. Seretny, Towarzystwo Autorów i Wydawców Prac Naukowych „Universitas”, Kraków 2004.
- Miodunka W. 2013a: *10-lecie certyfikacji języka polskiego i jej wpływ na nauczanie polszczyzny cudzoziemców*, „Języki Obce w Szkole”, nr 3, s. 16–22.
- Miodunka W.T. 2013b: *Jakość polszczyzny używanej przez cudzoziemców. Metody analizy jakości na materiale egzaminów certyfikacyjnych z języka polskiego jako obcego w roku 2011*, „Poradnik Językowy”, nr 1, s. 53–68.
- Miodunka W., Przechodzka G. 2006: *Kompetencja w języku polskim polskich maturzystów i zaawansowanych cudzoziemców. Badania porównawcze*, „Polonistyka”, nr 7, s. 6–15.
- Miodunka W., Przechodzka G. 2007: *Mother tongue versus foreign language performance – a Polish case*, [w:] W. Martyniuk (red.), *Towards a Common European Framework of Reference for languages of school education? Proceedings of a conference*, Towarzystwo Autorów i Wydawców Prac Naukowych „Universitas”, Kraków, s. 301–308.
- Niemierko B. 1999: *Pomiar wyników kształcenia*, Wydawnictwa Szkolne i Pedagogiczne, Warszawa.

- Państwowe Egzaminy Certyfikatywne z Języka Polskiego jako Obcego. Standardy wymagań egzaminacyjnych, Państwowa Komisja Poświadczania Znajomości Języka Polskiego jako Obcego, Ministerstwo Edukacji Narodowej i Sportu, Warszawa 2003.
- Pawłowska R. 1992: *Lingwistyczna teoria nauki czytania*, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk.
- Piasecka L. 2000: *Implications of theoretical approaches to EFL reading*, [w:] J. Arabski (red.), *Studies in foreign language acquisition and teaching*, Wydawnictwo Uniwersytetu Śląskiego, Katowice, s. 256–264.
- Piasecka L. 2011: *Trudności w rozwijaniu sprawności receptywnych – jak wspierać ucznia w nauce czytania ze zrozumieniem*, [w:] H. Komorowska (red.), *Nauka języka obcego w perspektywie ucznia*, Oficyna Wydawnicza Łośgraf, Warszawa, s. 259–274.
- Prizel-Kania A. 2011: *Rozwijanie i testowanie sprawności rozumienia ze słuchu w języku polskim jako obcym na poziomie B2 z uwzględnieniem rezultatów egzaminów certyfikatywnych*, praca doktorska, Uniwersytet Jagielloński, Kraków.
- Przewodnik dla autorów zadań do testów językowych, przeł. M. Gaszyńska-Magiera, A. Seretny, Towarzystwo Autorów i Wydawców Prac Naukowych „Universitas”, Kraków 2004.
- Seretny A. 2004: *Testy językowe i ich rodzaje. Kryteria poprawności testu językowego*, [w:] A. Seretny, W. Martyniuk, E. Lipińska (red.), *Opisywanie, rozwijanie i testowanie znajomości języka polskiego jako obcego*, Towarzystwo Autorów i Wydawców Prac Naukowych „Universitas”, Kraków, s. 63–78.
- Seretny A. 2010: *Rozwijanie kompetencji leksykalnej uczących się przez obcowanie z tekstem*, [w:] G. Zarzycka, G. Rudziński (red.), *Teksty i podteksty w nauczaniu języka polskiego jako obcego – 2*, „Acta Universitatis Lodzianis. Kształcenie Polonistyczne Cudzoziemców”, nr 17, Wydawnictwo Uniwersytetu Łódzkiego, Łódź, s. 547–559.
- Seretny A. 2011: *Kompetencja leksykalna uczących się języka polskiego jako obcego w świetle badań ilościowych*, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków.
- Seretny A., Lipińska E. 2005: *ABC metodyki nauczania języka polskiego jako obcego*, Towarzystwo Autorów i Wydawców Prac Naukowych „Universitas”, Kraków.
- Vedovelli M. 2007: *Introduzione*, [w:] M. Vedovelli (red.), *Manuale della certificazione dell'italiano L2*, Carocci, Roma, s. 11–28.
- Weir C. 2005: *Language testing and validation. An evidence-based approach*, Palgrave Macmillan, Basingstoke.
- Wolf M., Vellutino F., Gleason J.B. 2005: *Psycholingwistyczna analiza czynności czytania*, [w:] J.B. Gleason, N. Bernstein Ratner (red.), *Psycholingwistyka*, przeł. J. Bobryk, Gdańskie Wydawnictwo Psychologiczne, Gdańsk, s. 439–476.
- Żydek-Bednarczuk U. 2005: *Wprowadzenie do lingwistycznej analizy tekstu*, Towarzystwo Autorów i Wydawców Prac Naukowych „Universitas”, Kraków.
- <http://certyfikatpolski.pl/> (ostatni dostęp: 9 marca 2018).

## Summary

---

### Reading comprehension among candidates taking state certification exams in Polish as a foreign language

Keywords: Polish as a foreign language, certificate exams, reading comprehension, comparative glottodidactics.

The article presents the results of research which investigated the reading ability of candidates taking the certification exam in Polish as a foreign language at the B2 level in the years 2005–2008. The first part of the analysis was concerned with texts, testing techniques and whole tests used in the exam at the B2 level. The Polish corpus was compared with analogous reading comprehension tests in English and Italian. The comparative study of the reading papers showed both similarities between the tests and solutions specific to each exam. A crucial part of the research was a quantitative analysis of answers given by the candidates in part C of the Polish exam; selected tasks were also tested in a control group comprised of Polish secondary school students. The research showed that two types of test items: *true/false/not stated* and *matching information* had appeared relatively difficult for the test takers. The candidates had also difficulty understanding text structure of more complex texts.