



# Spojrzenie na wypowiedź pisemną na egzaminie gimnazjalnym przez pryzmat praktyczności testowania

[ Marcin Smolik ]

## **Dla osób niebędących nauczycielami ocenianie umiejętności tworzenia wypowiedzi pisemnej przez uczniów może wydawać się czynnością najprostszą pod słońcem: podaje się temat i uczniowie piszą. Wystarczy jednak krótka rozmowa z każdym nauczycielem języka polskiego, obcego lub historii, aby przekonać się, ile obaw towarzyszy procesowi oceniania wypowiedzi pisemnych i ile spraw związanych z tym zagadnieniem nadal pozostaje niejasnych.**

**S**koro tak liczne są wątpliwości na poziomie klasy, można sobie tylko wyobrazić, ile kłopotów musi nastroczać ocenianie umiejętności pisania w egzaminach ogólnokrajowych, do których przystępuje prawie pół miliona uczniów. W artykule omówię najważniejsze zagadnienia związane z ocenianiem umiejętności tworzenia wypowiedzi pisemnej w trzeciej części egzaminu gimnazjalnego na poziomie rozszerzonym przez pryzmat użyteczności testu w ujęciu Bachmana i Palmera (1996).

### **Testować, czy nie testować?**

Chociaż może się to wydawać dziwne, nie ma jednomyślności co do tego, czy umiejętność tworzenia wypowiedzi pisemnej lub szerzej – umiejętność pisania – w ogóle powinna być sprawdzana na egzaminach z języka obcego. Przyczyn takiego stanu rzeczy jest wiele, najważniejsze z nich wymieniam poniżej.

- Liczne próby stworzenia jednolitego psycholingwistycznego modelu tworzenia wypowiedzi pisemnej w języku drugim/obcym nie zakończyły się – jak dotąd – sukcesem. Wyniki badań wskazują na skomplikowane sieci połączeń wiążące wiedzę lingwistyczną (wiadomości i umiejętności) w zakresie języka rodzimego i drugiego/obcego oraz wiedzę związaną z tworzeniem spójnego tekstu tak pisanego, jak i mówionego. Rozsupłanie tych zależności, prowadzące do wyłuskania z tej pogmatwanej sieci powiązań wyłącznie tego, co można by

określić jako wiedzę związaną z tworzeniem wypowiedzi pisemnej w języku drugim/obcym, umożliwiłoby opracowanie zasad pożądaných działań dydaktycznych, a w późniejszym etapie – działań nakierowanych na mierzenie interesującej nas umiejętności. Jak dotąd jednak się to nie udało, chociaż oczywiście podejmowane są próby opracowania takiej definicji w – nazwijmy to – wersji roboczej, niekoniecznie nieskazitelnej pod względem empirycznym, ale umożliwiającej podjęcie praktycznych działań w klasie i na egzaminie. Wiele tych prób to działania naprawdę udane.

- Niektórzy kwestionują celowość sprawdzania umiejętności tworzenia dłuższej wypowiedzi pisemnej z szeroko pojętych względów kulturowych. Pomijając kontekst szkolny i akademicki, w życiu codziennym rzadko stajemy przed koniecznością stworzenia dłuższego tekstu samodzielnie. Nie oznacza to, że poza kontekstem edukacyjnym nie piszemy. Jest wręcz przeciwnie. Można by zaryzykować stwierdzenie, że piszemy więcej, niż miało to miejsce kiedyś, zmianie uległa jednak natura tworzonego tekstu. Wypowiedzi takie jak SMS czy e-mail mają charakter interaktywny i – tak naprawdę – pomijając kanał transmisji, noszą więcej cech charakterystycznych dla tekstów mówionych niż dla tekstów pisanych (Bachman i Palmer 1996:75).
- Nie bez znaczenia jest również wysuwany przez niektórych postulat dotyczący nieautentyczności tworzenia wypowiedzi

pisemnej pod presją ograniczonego czasu. Tworzenie tekstu – tutaj specjaliści wydają się być wyjątkowo jednomyślni – jest procesem twórczym, w którym cyklicznie powtarzane są ściśle ze sobą powiązane działania, takie jak sporządzanie planu, pisanie kolejnych wersji roboczych, analizowanie, strukturyzowanie wypowiedzi itp. (Hyland 2003). Próba odtworzenia tego procesu na egzaminie z języka obcego – nawet jeśli nie jest niemożliwa – byłaby ogromnym wyzwaniem<sup>1</sup>. W większości przypadków tekst stworzony na egzaminie traktowany jest jako produkt – ocenie podlega nie sam proces tworzenia, ale efekt tego procesu, tj. namacalny wytwór w postaci zapisu na kartce albo wydruku. To behawiorystyczne w swojej naturze podejście spotyka się z ostrą krytyką ze strony wielu specjalistów zajmujących się zawodowo pisaniem.

- Ściśle związana z autentycznością jest również kwestia celu tworzenia wypowiedzi pisemnej. W życiu codziennym każdy tekst piszemy po coś oraz dla kogoś, nawet jeśli tym kimś jesteśmy my sami, jak dzieje się np. w przypadku pamiętnika albo dziennika. Na egzaminie sytuacja ta jest w znacznym stopniu zaburzona. Jedynym realnym odbiorcą tworzonej wypowiedzi jest egzaminator, nawet jeżeli polecenie zawiera zapis, że tekst jest tworzony dla kolegi. Natomiast jedynym prawdziwym celem tworzenia wypowiedzi jest wykazanie się, w jakim stopniu zdający posiadał wiedzę związaną z tworzeniem wypowiedzi pisemnej w języku drugim/obcym (o problemach związanych z tym zagadnieniem pisałem już wyżej).
- Jako ostatnie – ale pod żadnym pozorem nie najmniej ważne – wymienię kwestie związane z samą sytuacją pomiarową. Nie ma zgody co do tego, które aspekty wypowiedzi oceniać i jak to robić. Czy powinniśmy skupić się na – najogólniej rzecz ujmując – treści, jakości języka, czy obydwu tych cechach? Czy powinniśmy stosować ocenianie analityczne, czy holistyczne (Weigle 2002)? Kto powinien dokonywać oceny? Ile wypowiedzi pozyskać od każdego zdającego? Te i wiele innych pytań rodzi mnóstwo wątpliwości związanych z trafnością i rzetelnością pomiaru (Weir 2005).

Powyższa lista z pewnością nie wyczerpuje zagadnienia. Skoro zatem sprawdzanie umiejętności tworzenia wypowiedzi pisemnej nastęrcza tylu wątpliwości oraz problemów, dlaczego nie sposób znaleźć egzaminu, który nie zawierałby komponentu o nazwie *Pisanie* lub podobnej? Po pierwsze, wykluczenie tej umiejętności

z egzaminów wiązałoby się zapewne z ograniczeniem jej obecności w kształceniu językowym, w którym pisanie funkcjonuje nie tylko jako cel sam w sobie, ale stanowi również środek (technikę nauczania) umożliwiający nauczanie i uczenie się gramatyki, słownictwa, rozumienia tekstów pisanych. Po drugie, pisanie zwiększa repertuar technik, które nauczyciel ma do dyspozycji, umożliwiających zaangażowanie uczniów w wykorzystanie nowo zdobytej wiedzy w sposób twórczy. Po trzecie wreszcie, samodzielne tworzenie wypowiedzi pisemnej – zwłaszcza dłuższej niż kilkudzaniowa – to ćwiczenie wymagające przekazania własnych myśli w sposób spójny i logiczny, wymagające zaangażowania nie tylko wiedzy językowej, ale również strategicznej, w tym metakognitywnej, związanej z planowaniem, monitorowaniem i ewaluacją osiągniętych rezultatów.

Pomimo licznych problemów związanych ze sprawdzaniem umiejętności tworzenia wypowiedzi pisemnej nie sposób zaprzeczyć, że jest to działanie użyteczne, a użyteczność jest – według Bachmana i Palmera (1996:17) – najważniejszą cechą każdego testu. W następnej części artykułu omówię zagadnienie użyteczności testu na przykładzie trzeciej części egzaminu gimnazjalnego z języka obcego nowożytnego na poziomie rozszerzonym.

### Użyteczność jako nadrzędna cecha dobrego testu

Bachman i Palmer (1996:17) definiują użyteczność jako wypadkową sześciu cech: trafności, autentyczności, interakcyjności (zaangażowania intelektualnego i emocjonalnego), oddziaływania na proces kształcenia, rzetelności i praktyczności. Choć wszystkie powyższe aspekty są równie ważne, należy od razu zaznaczyć, że opracowanie procedury testowania, którą charakteryzowałoby występowanie wszystkich tych cech w maksymalnym natężeniu, jest praktycznie niewykonalne. Najbliższe ideałowi było – niestety wyłącznie w swych założeniach – podejście komunikacyjne w testowaniu, jednakże *testowanie w pełni komunikacyjne i naturalne byłoby [...] nierealistyczne i niepraktyczne – wymagałoby organizowania wielogodzinnych, zindywidualizowanych egzaminów, na co nie pozwala czas, lokal i budżet* (Komorowska 2005:48). Te oraz inne ograniczenia sprawiają, że w praktycznych sytuacjach sprawdzania wiedzy uczniów konieczne jest osiągnięcie stanu, w którym cechy te osiągną natężenie właściwe dla danego kontekstu i tego, czego oczekuje się od wyników testowania. Przejdę teraz do omówienia tych cech po kolei. Ze względu na ograniczone miejsce zostaną one przedstawione jedynie pobieżnie; skupię się na wybranych aspektach każdego zagadnienia.

<sup>1</sup> Istnieją oczywiście alternatywne formy sprawdzania umiejętności tworzenia wypowiedzi pisemnej w ogromnym stopniu realizujące założenia pisania jako procesu, np. portfolio.

**Trafność to słuszność interpretacji wyniku uzyskanego w procesie testowania oraz wykorzystania tego wyniku. Test ma sprawdzać to, co zamierzaliśmy, a wyniki pozyskiwane podczas testowania powinny być właściwie interpretowane oraz wykorzystywane zgodnie z ich zakładanym przeznaczeniem.**

### Trafność

Trafność jest jednym z najbardziej problematycznych zagadnień nie tylko w sprawdzaniu umiejętności tworzenia wypowiedzi pisemnej, ale w ogóle w teorii pomiaru (Kane 2006; Komorowska 2002; Fulcher i Davidson 2007; Smolik 2008; Weir 2005); szczególnie jej opisanie zdecydowanie wykraczałoby poza zakres omawianego zagadnienia. Dla potrzeb niniejszego artykułu wystarczy, jeżeli zdefiniuję ją jako słuszność interpretacji wyniku uzyskanego w procesie testowania oraz wykorzystania tego wyniku (Messick 1989; Niemierko 1999). Dalej upraszczając, chodzi o to, aby test mierzył to, co zamierzaliśmy, że ma mierzyć, a wyniki pozyskiwane w wyniku procesu testowania były właściwie interpretowane oraz wykorzystywane zgodnie z ich zakładanym przeznaczeniem.

Trafność nie jest tworem jednowymiarowym; składają się nań różne sposoby gromadzenia danych o słuszności i poprawności interpretacji wyniku. Dwa podstawowe wymiary trafności to trafność konstruktu (ang. *construct validity*), zwana również teoretyczną, oraz trafność oparta na analizie treści (ang. *content validity*). W zadaniach/testach sprawdzających szeroko rozumianą umiejętność pisania trafność konstruktu ustala się, odnosząc zadanie do konkretnego modelu teoretycznego pisania bądź wybranych aspektów kilku takich modeli. Natomiast w przypadku trafności opartej na analizie treści – porównuje się wymagania w danym zadaniu np. z wymogami określonymi w programie nauczania bądź podręczniku.

W przypadku części egzaminu gimnazjalnego z języka obcego nowożytnego oba te wymiary trafności pokrywają się w znacznym stopniu ze względu na fakt, że podstawą tworzenia zadań egzaminacyjnych są wymagania ogólne i szczegółowe określone w podstawie programowej kształcenia ogólnego dla III etapu edukacyjnego (MEN 2008). Wymagania stanowią niejako definicję konstruktów, a jednocześnie określają treści, które powinny być nauczane oraz sprawdzane. Oczywiście nie

nie stoi na przeszkodzie, aby wymagania podstawy programowej odnieść do wybranego modelu teoretycznego pisania albo – nawet szerzej – kompetencji komunikacyjnej (np. Bachman 1990)<sup>2</sup>.

Jeżeli chodzi o tworzenie wypowiedzi, podstawa programowa III.1 (MEN 2008) określa, że uczeń III klasy gimnazjum, który kontynuuje naukę języka obcego ze szkoły podstawowej, czyli uczy się tego samego języka około sześciu lat, *tworzy krótkie, proste i zrozumiałe wypowiedzi pisemne* (np. *notatka, ogłoszenie, zaproszenie, pozdrowienia, życzenia, wiadomość, ankieta, pocztówka, e-mail, opis, krótki list prywatny*). To wymaganie szczególnie zostało doprecyzowane w *Informatorze o egzaminie gimnazjalnym* (CKE 2010), w którym określono, że zadanie sprawdzające umiejętność tworzenia wypowiedzi pisemnej *polega na napisaniu krótkiego (50-100 słów) tekstu użytkowego (listu prywatnego, wiadomości, e-maila) z elementami opisu, relacjonowania, zaproszenia, wyrażania i uzasadniania opinii i uczuć itp., zgodnie ze szczegółowymi wskazówkami podanymi w poleceniu. W każdym poleceniu podane są trzy elementy, które uczeń powinien rozwinąć w wypowiedzi* (str. 102). Innymi słowy – wymagania podstawy programowej wyznaczyły zakres tego, co może być sprawdzane na egzaminie.

Przyjrzyjmy się teraz, w jaki sposób powyższe zapisy zostały zrealizowane praktycznie w zadaniach, które były wykorzystane podczas sesji egzaminacyjnej w kwietniu br. W tym roku uczniowie mogli wybrać jeden z języków obcych, którego uczyli się w gimnazjum jako przedmiotu obowiązkowego. Mieli wybór spośród następujących: język angielski, język francuski, język hiszpański, język niemiecki, język rosyjski, język włoski. Ponieważ wszyscy uczniowie przystępowali do części trzeciej egzaminu tego samego dnia, możliwe było wykorzystanie

<sup>2</sup> Warto w tym miejscu podkreślić, że wymagania podstawy programowej dla języka obcego nowożytnego w dużej mierze bazują na skalach biegłości zawartych w *Europejskim Systemie Opisu Kształcenia Językowego* (CODN 2003). Skale zawarte w ESOKJ nie są jednak modelem kompetencji komunikacyjnej. Szerzej o problemach związanych ze skalami w ESOKJ piszą m.in. Weir (2005), Alderson (2007), Hulstijn (2007), Davidson i Fulcher (2007).

tego samego zadania w arkuszach z różnych języków. W Tabeli 1. przedstawiony jest opis obu zadań wraz z wymaganiami z podstawy programowej. Warto w tym miejscu podkreślić, że pomimo ewidentnych różnic, oba zadania równie dużo łączy, zwłaszcza jeżeli chodzi o wymagania szczegółowe, których spełnienie jest sprawdzane.

Tabela 1. Wypowiedź pisemna na egzaminie gimnazjalnym z języków obcych w kwietniu 2012 r.

	Zadanie 1.	Zadanie 2.
<b>Treść</b>	<p>Napisz wiadomość e-mail do kolegi/koleżanki z (Anglii/Hiszpanii/Włoch):</p> <ul style="list-style-type: none"> <li>■ poinformuj, w jakich okolicznościach straciłeś(-łaś) swój telefon komórkowy;</li> <li>■ opisz, jak na to zdarzenie zareagowali Twoi rodzice;</li> <li>■ wyjaśnij, jakie problemy masz z powodu braku telefonu.</li> </ul>	<p>W Twojej szkole odbędzie się spotkanie ze sławną osobą. W e-mailu do kolegi/koleżanki z (Francji/Niemiec/Rosji):</p> <ul style="list-style-type: none"> <li>■ napisz, dlaczego zaprosiliście tę osobę do szkoły;</li> <li>■ wyjaśnij, na czym polega Twój udział w przygotowaniach do tego spotkania;</li> <li>■ opisz, jak zamierzasz ubrać się w tym dniu.</li> </ul>
<b>Języki</b>	angielski, hiszpański, włoski	francuski, niemiecki, rosyjski
<b>Wymagania ogólne</b>	<p>I. Znajomość środków językowych            III. Tworzenie wypowiedzi            IV. Reagowanie na wypowiedzi</p>	<p>I. Znajomość środków językowych            III. Tworzenie wypowiedzi            IV. Reagowanie na wypowiedzi</p>
<b>Wymagania szczegółowe</b>	<p>1. Uczeń posługuje się podstawowym zasobem środków językowych (leksykalnych, gramatycznych, ortograficznych), umożliwiającą realizację pozostałych wymagań ogólnych w zakresie następujących tematów:            1) człowiek;            2) nauka i technika.</p> <p>5. Uczeń tworzy krótkie, proste i zrozumiałe wypowiedzi pisemne, np. e-mail:            4) relacjonuje wydarzenia z przeszłości;            6) przedstawia opinie innych osób;            1) opisuje ludzi (emocje);            8) opisuje swoje doświadczenia;            9) stosuje formalny lub nieformalny styl wypowiedzi w zależności od sytuacji.</p> <p>7. Uczeń reaguje w formie prostego tekstu pisanego, np. e-mail, w typowych sytuacjach:            2) uzyskuje i przekazuje informacje i wyjaśnienia.</p>	<p>1. Uczeń posługuje się podstawowym zasobem środków językowych (leksykalnych, gramatycznych, ortograficznych), umożliwiającą realizację pozostałych wymagań ogólnych w zakresie następujących tematów:            1) człowiek;            3) szkoła;            9) kultura.</p> <p>5. Uczeń tworzy krótkie, proste i zrozumiałe wypowiedzi pisemne, np. e-mail:            1) opisuje ludzi i czynności;            4) relacjonuje wydarzenia z przeszłości;            7) opisuje plany na przyszłość;            9) stosuje formalny lub nieformalny styl wypowiedzi w zależności od sytuacji.</p> <p>7. Uczeń reaguje w formie prostego tekstu pisanego, np. e-mail, w typowych sytuacjach:            2) uzyskuje i przekazuje informacje i wyjaśnienia.</p>

Aby sprawdzić, czy zakładane umiejętności zostały istotnie sprawdzone, przeanalizujemy kilka przykładowych wypowiedzi zaczerpniętych z prac gimnazjalistów, którzy przystąpili do egzaminu z języka angielskiego (Zadanie 1. w Tabeli 1.). W cytowanych poniżej fragmentach zachowano oryginalną pisownię.

<p><i>I was yesterday in the cinema and lost my mobile.</i></p>	<p>5.4) Uczeń relacjonuje wydarzenia z przeszłości. 7.2) Uczeń przekazuje informacje i wyjaśnienia.</p>
<p><i>My mum and dad weren't happy when I told them about this situation. They screamed at me whole weekend. I think they have dissapointed on me and I feel terrible with that.</i></p>	<p>5.4) Uczeń relacjonuje wydarzenia z przeszłości. 5.6) Uczeń przedstawia opinie innych osób. 5.8) Uczeń opisuje swoje doświadczenia.</p>
<p><i>Now I feel terrible, because I can't text with my friends and even ask them about homework. What's more I've lost all my photos and all numbers.</i></p>	<p>5.4) Uczeń relacjonuje wydarzenia z przeszłości. 5.8) Uczeń opisuje swoje doświadczenia. 1.12) Uczeń posługuje się podstawowym zasobem środków językowych (leksykalnych, gramatycznych, ortograficznych), umożliwiających realizację pozostałych wymagań ogólnych w zakresie tematu nauka i technika.</p>

Oczywiście wśród setek tysięcy prac egzaminacyjnych znajdują się również takie, w których nie znajdziemy dowodów na to, że uczeń potrafi przekazać wymagane informacje w formie pisemnej. Nie stanowi to jednak problemu z punktu widzenia pomiaru, ponieważ taka wypowiedź po prostu nie uzyska punktów za treść, co oznacza, że wynik tego ucznia będzie odzwierciedlał brak umiejętności, które podlegały sprawdzeniu.

W tych z konieczności krótkich rozważaniach nad trafnością nie można pominąć fundamentalnego pytania, tj. przejawem jakich umiejętności jest właściwie stworzona przez ucznia wypowiedź? Przecież w testowaniu nie może chodzić wyłącznie o to, aby dowiedzieć się, czy uczeń potrafi wykonać konkretne zadanie, w naszym przypadku – napisać e-mail o straconym telefonie komórkowym. Zadanie stanowi jedynie pretekst do tego, aby dowiedzieć się czegoś o szerszych umiejętnościach ucznia, niezbędne jest jednak wyznaczenie granic takiej generalizacji wyniku. Z pomocą może nam tu przyjść wspomniany wcześniej Europejski System Opisu Kształcenia Językowego (CODN 2003), do którego w komentarzu nawiązują autorzy podstawy programowej. Określają oni, że poziom III.1 podstawy ma wymagania zbliżone do poziomu A2+ w sześciostopniowej skali poziomów biegłości zdefiniowanej przez Radę Europy<sup>3</sup>. Jeżeli sięgniemy do skali nazwanej *Pisanie twórcze*, znajdziemy tam następujący deskryptor dla poziomu A2:

*Potrafi sporządzić bardzo krótkie, proste opisy obecnych i przeszłych wydarzeń i osobistych doświadczeń. Potrafi w formie połączonych zdań pisać o codziennych sprawach własnego otoczenia, np.: o ludziach, miejscach [...] (CODN 2003:76).*

Można wstępnie przyjąć, że wynik, jaki uzyska uczeń za wypowiedź pisemną na egzaminie gimnazjalnym, można interpretować jako wskaźnik takich umiejętności, bez nadmiernej obawy o przekraczanie granic generalizacji. Oczywiście wnioski takie jest wyłącznie efektem analiz teoretycznych, dla pełnego jego potwierdzenia konieczne byłyby pogłębione badania empiryczne. Cała powyższa dyskusja wskazuje jednak, że trafność zawsze stanowiła dla twórców egzaminu ważny element w rozważaniach nad użytecznością wypowiedzi pisemnej na egzaminie gimnazjalnym.

### Autentyczność

Bachman i Palmer (1996:23) definiują autentyczność jako stopień, w jakim zadanie egzaminacyjne odzwierciedla typy zadań, z jakimi zdający muszą sobie radzić w sytuacjach nieegzaminacyjnych. Z jednej strony sprawa jest prosta – komunikacja elektroniczna praktycznie zastąpiła tradycyjny list, zatem można by zakończyć rozważania stwierdzeniem o wysokiej autentyczności zadania i przejść do omówienia kolejnej cechy. Z drugiej strony jednak, po głębszym zastanowieniu wnioski nie wydają się już tak oczywiste.

<sup>3</sup> Więcej o powiązaniach pomiędzy częścią egzaminu gimnazjalnego z języka obcego nowożytnego i *ESOKJ* można znaleźć w: Smolik i Trzcińska (2011).

Po pierwsze, samo zagadnienie autentyczności w testowaniu jest niezwykle skomplikowane (Lewkowicz 2000); niektórzy w ogóle wątpią w możliwość uzyskania autentycznej sytuacji komunikacyjnej na egzaminie językowym<sup>4</sup>. Po drugie, wychodząc z założenia, że autentyczne jest to, co się wydaje autentyczne odbiorcy, można przyjąć, że uczniowie, którzy nie potrzebują umiejętności pisania do realizacji swoich potrzeb komunikacyjnych w języku obcym, z góry założą, że zadanie jest nieautentyczne [szerzej o uwzględnieniu potrzeb komunikacyjnych uczących się w kształceniu i egzaminowaniu pisze Komorowska (2005)]. Chociaż powyższe argumenty trudno byłoby uznać za niesłuszne, trzeba jednak zauważyć, że zadaniem szkoły jest przygotować uczniów nie tylko do funkcjonowania tu i teraz, ale również starać się przewidzieć ich potrzeby, które mogą pojawić się po ukończeniu nauki. Nie możemy na etapie gimnazjum wykluczyć, że za kilka lat uczeń, któremu w tej chwili wydaje się, że pisanie e-maila po angielsku jest zadaniem jałowym, stanie przed koniecznością porozumiewania się w języku obcym chociażby w kontekstach zawodowych z innymi użytkownikami języka angielskiego jako obcego. Stąd, między innymi, decyzja o wprowadzeniu elementu tworzenia wypowiedzi pisemnej na egzaminie.

Powracając do samego zadania, warto wskazać cechy nadające mu znamiona autentyczności:

- podobnie jak w e-mailach, które wysyłamy do znajomych w życiu codziennym, w wypowiedzi na egzaminie najważniejsze jest pełne przekazanie informacji w spójny i logiczny sposób; język stanowi jedynie środek do osiągnięcia tego celu. Znajduje to odzwierciedlenie w skali oceniania: uczeń może zdobyć maksymalnie 6 punktów za treść i spójność wypowiedzi, a 4 punkty za zakres i poprawność środków językowych;
- na stworzenie wiadomości uczeń ma ograniczony czas – około 15-20 minut. W autentycznej sytuacji komunikacyjnej nie poświęcilibyśmy pół dnia na napisanie trzydzielnego e-maila do kolegi, nie ma zatem podstaw, aby na egzaminie stwarzać taką sytuację. E-mail prywatny jest z założenia szybką formą komunikacji niewymagającą nadmiernej staranności stylistycznej, choć również nieczwalać na gwałcenie podstawowych norm językowych, co znajduje odzwierciedlenie w skali oceniania;
- polecenie określa sytuację, w której pisanie w języku obcym byłoby uzasadnione, tj. e-mail do kolegi z kraju, w którym język docelowy stanowi podstawowy sposób komunikacji;

<sup>4</sup> Perren (1967, za Stevenson 1985:47), pisząc o egzaminach ustnych, stwierdził: *nonsensem byłoby przypuszczać, że sytuacja egzaminator-kandydat odzwierciedla prawdziwe życie albo komunikację społeczną; obaj uczestnicy wiedzą doskonale, że uczestniczą w teście, a nie w podwieczorku przy herbacie, i obaj podlegają presjom psychologicznym oraz takim ograniczeniom stylu i rejestru, jakie obydwu stronom wydają się stosowne do okazji.*

- tekst tworzony przez ucznia dotyczy zagadnień bliskich doświadczeniom jego oraz/lub jego kolegów i koleżanek.

### Interakcyjność: zaangażowanie intelektualne i emocjonalne

Bachman i Palmer (1996:25) podkreślają, że zadania egzaminacyjne powinny odzwierciedlać taki poziom zaangażowania intelektualnego i emocjonalnego (ang. *interactiveness*) zdających, jaki towarzyszyłby analogicznemu działaniu językowemu w sytuacji nieegzaminacyjnej. Uzyskanie możliwie wysokiej interakcyjności zadania jest istotne, ponieważ brak zaangażowania ze strony zdającego może w istotny sposób wpłynąć na poziom wykonania zadania ze względu na zmniejszoną motywację do działania.

W zestawie zadań na poziomie rozszerzonym znajdują się zadania o stosunkowo niskim poziomie interakcyjności. Są to zadania sprawdzające znajomość środków językowych, w których zdający musi jedynie dokonać odpowiednich przekształceń syntaktycznych, morfologicznych i leksykalnych, aby stworzyć poprawne gramatycznie i leksykalnie zdania. W tym przypadku trudno oczekiwać szczególnego zaangażowania emocjonalnego; działanie jest w znacznym stopniu mechaniczne i w średnim stopniu angażuje zdających poznawczo.

W przypadku wypowiedzi pisemnej jednakże można przypuszczać, że zaangażowanie intelektualne i emocjonalne będzie większe. Konieczność stworzenia własnego tekstu angażuje nie tylko wiedzę językową, ale również strategiczną, wymaga odpowiedniego planowania, oszacowania dostępnych środków językowych i podjęcia kroków umożliwiających pełne przekazanie informacji, nawet pomimo braków leksykalnych i gramatycznych. Aby otrzymać maksymalny wynik za treść, zdający muszą nie tylko odnieść się do wszystkich podpunktów z polecenia, ale również rozwinąć swoją wypowiedź, co czasami może się wiązać z koniecznością porozumienia się z odbiorcą pomimo braków w znajomości środków językowych (kompetencja plurilingwalna; CODN 2003, Komorowska 2005).

Odpowiednio opracowana skala oceniania nie ogranicza odpowiedzi uczniów do inwentarza oczekiwanych fraz kluczowych, których egzaminator wyszukiwałby w wypowiedzi (słownego klucza, o którym – z mniejszym lub większym sensem – rozpisuje się prasa). Uczeń ma ogromną swobodę w realizacji tematu, co skutkuje niejednokrotnie niestandardowymi rozwiązaniami, w których bez trudu odnajdziemy ślady pełnego zaangażowania emocjonalnego w tworzenie wypowiedzi. Poniższy przykład (przycięty z całości) stanowi dobrą ilustrację omawianej kwestii. W wypowiedzi zachowano oryginalną pisownię.

*Hi Mark!*

*I know you were calling me, but 3 days ago I lost my mobile. I was going home when 3 bold guys went to me.*

*One ask me for a time, and another steal my phone.*

*My mum was very angry, but my dad understands me.*

*He even wanted to ride on that 3 guys, with a hatchet, but we didn't, because mum was crying.*

*Now I cant answer my phone, so no one has a contact with me. It makes my very angry. Life is brutal.*

*Take care*

*XYZ*

Podsumowując, nawet jeżeli nie można przyjąć a priori, że zadanie ma „wbudowaną” wysoką interakcyjność, można śmiało uznać, że nie wyklucza stworzenia wypowiedzi, która będzie zaangażowana emocjonalnie i – co istotne – wysoko oceniona.

### Oddziaływanie na proces kształcenia

Każdy egzamin wysokiej stawki (ang. *high-stakes exam*), a za taki z pewnością można uznać egzamin gimnazjalny, wywiera wpływ na wszystkich uczestników i wszystkie elementy procesu kształcenia, np. uczniów, nauczycieli, materiały nauczania [szerzej pisałem o tym w poprzednim numerze tego czasopisma (Smolik 2012), problem omawia również Komorowska (2005)]. Troska o to, aby to oddziaływanie było jak najbardziej korzystne, powinna stanowić integralną część procesu testowania od momentu rozpoczęcia prac nad egzaminem. Tak właśnie było w przypadku egzaminu gimnazjalnego z języka obcego – mieliśmy pełną świadomość, że wprowadzenie elementu z pisania będzie miało wpływ na działania pedagogiczne podejmowane przez nauczycieli w procesie kształcenia, jak również na zmianę przekonania nauczycieli i uczniów na temat tego, czego warto nauczać i się uczyć oraz jak kierować nauczaniem, aby wspierało proces uczenia się. Bez wątplenia kluczowym elementem tej układanki jest podstawa programowa kształcenia ogólnego, która reguluje i integruje dwa, skądinąd nierozzerwalne, procesy kształcenia językowego, tj. nauczanie i sprawdzanie wiedzy, dzięki temu, że wymagania ogólne i szczegółowe wyznaczają zarówno to, co ma być nauczane, jak również to, co ma być sprawdzane na egzaminie (Smolik 2012). Dzięki temu egzamin gimnazjalny z języka obcego przestaje być wyłącznie testem biegłości (ang. *proficiency test*), a staje się – w pewnym stopniu – testem osiągnięć (ang. *achievement test*). Powstaje swoistego rodzaju hybryda tych dwóch rodzajów testowania, nazywana często w literaturze *proachievement testing* (np. Gonzalez Pino 1989; Shrum i Glisan 2009).

Nie można oczywiście przyjąć prostego założenia, że wprowadzenie wypowiedzi pisemnej na egzaminie całkowicie zmieni wszystkich i wszystko. Efekt zwrotny testowania to niezmiernie skomplikowana kwestia, obejmująca nie tylko zmiany techniczne, prawne, regulaminowe, ale również zagadnienia związane z osobistymi poglądami i przekonaniem nauczycieli i uczniów. Nie można również zapomnieć o wymiarze stricte konkursowym egzaminu. Cokolwiek by o tym nie sądzić, wyniki egzaminacyjne są brane pod uwagę w procesie rekrutacji do szkół wyższego szczebla. Sytuacja taka nie ma jeszcze miejsca w przypadku części egzaminu z języka obcego na poziomie rozszerzonym<sup>5</sup>, tym niemniej dyrektorzy szkół mogą wykorzystać te wyniki, aby wstępnie zakwalifikować uczniów do grup o odpowiednim poziomie zaawansowania językowego. Kilkuletni okres karencji dla wyniku z egzaminu na poziomie rozszerzonym należy przyjąć jako przejaw troski o wyrównanie szans dla wszystkich uczniów. Nie każdy nauczyciel i nie każda szkoła przykładali równie dużo uwagi kształceniu umiejętności pisania, należy dać im czas na nadrobienie zaległości.

### Rzetelność

Dbałość o rzetelność procesu testowania to – w dużej mierze – dbałość o jego trafność, ponieważ obie te cechy łączy wręcz symbiotyczna zależność. Testowanie nie jest rzetelne, jeżeli nie umożliwia pozyskania wyników, które są w dużej mierze niezależne od partykularnych charakterystyk sytuacji testowej. Jeżeli testowanie nie jest rzetelne, najprawdopodobniej nie będzie również trafne.

Lista czynników, które rzutują na rzetelność testowania jest długa i nie sposób omówić ich wszystkich w tak krótkim tekście. Dla porządku wspomnę tylko o samym zadaniu, skali oceniania, sposobie przygotowania osób sprawdzających, warunkach pisania oraz oceniania wypowiedzi (por. Barkaoui 2007). Zmienność tych czynników wpływa na rzetelność, dlatego należy dołożyć starań, aby zmienność ta była jak najmniejsza. W przypadku wypowiedzi pisemnej na egzaminie gimnazjalnym z języka obcego:

- wszyscy uczniowie przystępujący do egzaminu z tego samego języka tworzą wypowiedź na ten sam temat (rodzi to inne problemy związane np. z brakiem wyboru tematu przez ucznia oraz wnioskowaniem o umiejętnościach na podstawie jednej próbki; por. Weigle, 2002, 2012);
- każde zadanie jest zbudowane w identyczny sposób, składa się z wstępu oraz trzech podpunktów; określone jest, jak długa powinna być wypowiedź (50-100 słów);

<sup>5</sup> Po raz pierwszy wyniki z poziomu rozszerzonego będą brane pod uwagę dopiero w roku szkolnym 2017/2018.



- wszyscy zdający przystępują do egzaminu tego samego dnia, o tej samej porze, mają mniej więcej tyle samo czasu na stworzenie tekstu;
- zadania wykorzystane w zestawach w różnych językach są do siebie podobne;
- prace oceniają wyłącznie przeszkoleni egzaminatorzy;
- prace oceniane są według tych samych kryteriów: treść, spójność i logika wypowiedzi, zakres środków językowych oraz poprawność środków językowych;
- ogólne kryteria oceniania zostają uszczegółowione, tzn. dostosowane do konkretnego zadania, zanim egzaminatorzy przystąpią do sprawdzania prac.

Skuteczność wszystkich tych zabiegów – powziętych przed egzaminem – będzie podlegała ocenie na podstawie danych empirycznych zebranych po sesji egzaminacyjnej, która jeszcze trwa. Wstępne oszacowania, możliwe dzięki danym zgromadzonym podczas szkoleń dla egzaminatorów, pozwalają przypuszczać, że rzetelność oceniania jest na zadowalającym poziomie.

## Praktyczność

Patrząc na testowanie przez pryzmat praktyczności, najlepszym testem byłby taki, który do przeprowadzenia wymagałby minimalnego zaangażowania środków finansowych, zasobów ludzkich i czasu. Traktując priorytetowo praktyczność – automatycznie zmniejszamy udział pozostałych cech w ogólnym obrazie użyteczności. Biorąc jednak pod uwagę fakt, że egzamin językowy musi sprawdzić poziom opanowania wielu elementów w ograniczonym czasie, często poświęca się trafność albo autentyczność na rzecz praktyczności. Weir (1990:34) twierdzi nawet, że rzetelny i trafny test na niewiele się zda, jeżeli nie będzie również praktyczny.

Przeprowadzenie egzaminu gimnazjalnego z języka obcego na poziomie rozszerzonym wiąże się ze znacznymi nakładami finansowymi wynikającymi z:

- opracowania materiałów szkoleniowych dla kandydatów na egzaminatorów oraz przeprowadzenia szkoleń, również przed samym egzaminem;
- transportu arkuszy z i do okręgowych komisji egzaminacyjnych oraz ośrodków oceniania;
- zatrudnienia tysięcy egzaminatorów do sprawdzenia prawie trzystu tysięcy prac egzaminacyjnych, z których część jest sprawdzana dwukrotnie.

Wszystkie te działania wymagają ogromnego zaangażowania zarówno egzaminatorów, którzy przez dwa lub więcej dni sprawdzają prace w warunkach skoszarowania, jak i pracowników Centralnej Komisji Egzaminacyjnej i – przede wszystkim – okręgowych komisji, którzy koordynują cały proces.

Biorąc powyższe pod uwagę, trudno byłoby w przypadku omawianego egzaminu mówić o primacie praktyczności nad trafnością lub autentycznością. I bardzo dobrze. Praktyczność nadal pozostaje przedmiotem troski, publiczne pieniądze i ludzki wysiłek na pewno nie idą marne, ale egzamin sprawdza szerszy zakres kluczowych umiejętności.

## Podsumowanie

Opracowanie ogólnokrajowego egzaminu przypomina czasem chodzenie po linie – zadanie jest ogromnym wyzwaniem, wiąże się jednak z ciągłą koniecznością balansowania, aby utrzymać równowagę pomiędzy trafnością, autentycznością, interakcyjnością, oddziaływaniem na kształcenie językowe, rzetelnością i praktycznością. Zachowanie tej równowagi wiąże się czasem z koniecznością zrezygnowania z pewnej cechy na rzecz innej, ale zawsze prowadzi do zoptymalizowania cechy nadrzędnej, tj. użyteczności testowania. W artykule starałem się przedstawić (z przyczyn oczywistych – wyłącznie w zarysie) swoją subiektywną opinię na temat użyteczności części egzaminu gimnazjalnego na poziomie rozszerzonym sprawdzającej umiejętność tworzenia wypowiedzi pisemnej. Na Wykresie 1. podejmuje próbę podsumowania dyskusji w formie graficznej.

Wykres 1. Natężenie sześciu cech użyteczności (Bachman i Palmer 1996) w przypadku zadania sprawdzającego umiejętność tworzenia wypowiedzi pisemnej w trzeciej części egzaminu gimnazjalnego na poziomie rozszerzonym (ujęcie bardzo subiektywne).

Trafność	Au- tentyczność	Interakcyjność	Oddziaływanie na kształcenie	Rzetelność	Praktyczność

Kluczowe w przypadku opisywanego zadania wydają mi się jego trafność i potencjalny efekt, jaki wprowadzenie tego zadania do egzaminu powinno wywrzeć na działania podejmowane przez nauczycieli i uczniów, być może również na ich przekonania i postawy (Smolik 2012). Cechą, którą trzeba było – trochę z natury rzeczy – poświęcić, jest interakcyjność zadania. Wydaje mi się jednak, że jest to nie tyle wada egzaminu gimnazjalnego jako takiego, ale ogólny problem wszystkich egzaminów sprawdzających tzw. umiejętności językowe (słuchanie, czytanie, mówienie, pisanie).

## Bibliografia

- Alderson, J. C. (2007) The CEFR and the need for more research. *The Modern Language Journal*, t. 91, 659-663.
- Bachman, L. (1990) *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bachman, L., Palmer, A. (1996) *Language Testing in Practice*. Oxford: Oxford University Press.
- Barkaoui, K. (2007) Participants, Texts, and Processes in ESL/EFL Essay Tests: A Narrative Review of the Literature. *The Canadian Modern Language Journal*, t. 64, nr 1, 99-134.
- CKE. (2010) *Informator o egzaminie gimnazjalnym od roku szkolnego 2011/2012* [dostęp 12.02.2012] <[http://www.cke.edu.pl/images/stories/0012\\_Gimnazjum/100826\\_Informator%20gimnazjalny.pdf](http://www.cke.edu.pl/images/stories/0012_Gimnazjum/100826_Informator%20gimnazjalny.pdf)> .
- CODN. (2003) *Europejski System Opisu Kształcenia Językowego: uczenie się, nauczanie, ocenianie*. Warszawa: Wydawnictwo CODN.
- Davidson, F., Fulcher, G. (2007) The Common European Framework of Reference (CEFR) and the design of language tests: A matter of effect. *Language Teaching*, t. 40, 231-241.
- Fulcher, G., Davidson, F. (2007) *Language Testing and Assessment: An advanced resource book*. London & New York: Routledge.
- Gonzalez Pino, B. (1989) Prochievement Testing of Speaking. *Foreign Language Annals*, t. 22, nr 5, 487-496.
- Hulstijn, J. H. (2007) The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency. *The Modern Language Journal*, t. 91, 663-667.
- Hyland, K. (2003) *Second language writing*. Cambridge: Cambridge University Press.
- Kane, M. (2006) Validation. W: R. L. Brennan (red.) *Educational Measurement: Fourth Edition*. Westport, CT: Praeger Publishers.
- Komorowska, H. (2002) *Sprawdzanie Umiejętności w Nauce Języka Obcego. Kontrola – Ocena – Testowanie*. Warszawa: Fraszka Edukacyjna.
- Komorowska, H. (2005) Testowanie językowych osiągnięć uczniów – trudności i zagrożenia. *Języki Obce w Szkole*, t. XLIX, nr 6, 41-51.
- Lewkowicz, J. (2000) Authenticity in language testing: some outstanding questions. *Language Testing*, t. 17, nr 1, 43-64.
- MEN. (2008) *Rozporządzenie Ministra Edukacji Narodowej z dnia 23 grudnia 2008 r. w sprawie podstawy programowej wychowania przedszkolnego oraz kształcenia ogólnego w poszczególnych typach szkół*. [dostęp 07.02.2012] <[http://bip.men.gov.pl/men\\_bip/akty\\_prawne/rozporzadzenie\\_20081223\\_v2.pdf](http://bip.men.gov.pl/men_bip/akty_prawne/rozporzadzenie_20081223_v2.pdf)> .
- Messick, S. (1989) Validity. W: R. L. Linn (red.) *Educational Measurement: Third Edition*. New York: American Council on Education & Macmillan.
- Niemierko, B. (1999) *Pomiar wyników kształcenia*. Warszawa: WSiP.
- Shrum, J. L., Glisan, J. W. (2009) *Teacher's Handbook. Contextualised language instruction*. Boston, MA: Heinle.
- Smolik, M. (2008) *Investigating scoring validity. A study of the 'nowa matura' speaking exam in English at the basic level*. Niepublikowana rozprawa doktorska (UMCS Lublin).
- Smolik, M. (2012). Języki obce na egzaminach zewnętrznych: innowacje w latach 2012-2015. *Języki Obce w Szkole*, t. 1, nr 1, 22-33.
- Smolik, M., Trzcńska, B. (2011) *Zmiany w systemie egzaminów zewnętrznych a Europejski System Opisu Kształcenia Językowego*. Prezentacja podczas seminarium: *Odnoszenie egzaminów językowych do europejskiego systemu opisu kształcenia językowego (ESOKJ)*, Warszawa, 28 kwietnia 2011. [dostęp 06.05.2012] <<http://beta.eduentuzjasci.pl/images/stories/newsy/01.pdf>> .
- Stevenson, D. K. (1985) Authenticity, validity and a tea party. *Language Testing*, t. 2, nr 1, 41-47.
- Weigle, S. C. (2002) *Assessing writing*. Cambridge: Cambridge University Press.
- Weigle, S. C. (2012) *Assessing Writing*. W: Ch. Coombe, P. Davidson, B. O'Sullivan, S. Stoyhoff (red.) *The Cambridge Guide to Second Language Assessment*. Cambridge: Cambridge University Press.
- Weir, C. J. (1990) *Communicative Language Testing*. London: Prentice-Hall Regent.
- Weir, C. J. (2005) *Language Testing and Validation. An Evidence-Based Approach*. Basingstoke: Palgrave Macmillan.

## [ dr Marcin Smolik ]

Adiunkt w Zakładzie Metodyki Nauczania Języka Angielskiego Uniwersytetu Marii Curie-Skłodowskiej w Lublinie. Kieruje pracami trzech zespołów ds. modernizacji systemu egzaminów zewnętrznych z języków obcych nowożytnych w Centralnej Komisji Egzaminacyjnej, jest również koordynatorem zespołu języków obcych nowożytnych w projekcie VII CKE-EFS *Budowa banków zadań w OKE Łódź*.

e-mail: marcin.smolik@cke-efs.pl