

Ocenianie znajomości języka angielskiego dla celów specjalistycznych

Kontekst akademicki

DOI: 10.47050/jows.2025.1.73-82

Innowacyjne podejście do oceniania umiejętności językowych w kontekstach specjalistycznych to jednocześnie wyzwanie i szansa. Testy oparte na autentycznych sytuacjach wymagają nie tylko biegłości językowej, lecz także głębokiej wiedzy merytorycznej, co znacząco podnosi trafność i rzetelność oceny. Autorka definiuje ocenianie w kontekście ESP, omawia jego przesłanki i wyzwania oraz prezentuje przykłady zadań i kryteriów oceny, które zapewnią możliwie najlepszy sposób mierzenia kompetencji specjalistycznych.

BARBARA SAWICKA
Politechnika Poznańska

ESP ASSESSMENT IN AN ACADEMIC CONTEXT

Testing languages for specific purposes (LSP) involves the development of assessments grounded in real-life contexts. LSP testing is characterized by the authenticity of tasks, which reflect the essential features of target language use situations, and the integration of domain-specific content knowledge. This interplay between linguistic and content knowledge enhances the validity and reliability of assessments. Consequently, test developers should formulate criteria specifically tailored to LSP tasks. This paper aims to define LSP/ESP testing, describe its theoretical underpinnings, and provide examples of ESP assessment tasks and criteria within an academic context.

KEY WORDS: LSP, ESP, EAP, ASSESSMENT, VALIDITY, AUTHENTICITY, CRITERIA

Testowanie języków dla potrzeb specjalistycznych jest częścią obszaru testowania językowego, w którym cel, treść i metody wywodzą się z analizy sytuacji użycia języka docelowego (*Target Language Use, TLU*), np. test z języka hiszpańskiego dla biznesu, japońskiego dla przewodników turystycznych, włoskiego dla nauczycieli języków obcych lub angielskiego dla kontroli ruchu lotniczego (Douglas 2002: 1). Początkowo testy z języka angielskiego specjalistycznego (*English for Specific Purposes, ESP*) były postrzegane jedynie jako testy biegłości zawierające słownictwo o określonym przeznaczeniu. Istnieją jednak dwa atrybuty testów ESP, które odróżniają je od ogólnych testów językowych: autentyczność zadania i „interakcja między znajomością języka a posiadaniem wiedzy specjalistycznej” (Douglas 2002: 2 – tłum. własne). Autentyczność zadania ESP oznacza, że powinno ono posiadać kluczowe cechy działań wykonywanych w sytuacji TLU, które są istotne dla zdającego. Powodem tego powiązania między zadaniami testowymi a przyszłymi kontekstami użycia języka w TLU jest zapewnienie transferu do rzeczywistych zadań, tj. zwiększenie prawdopodobieństwa, że zdający wykona zadanie testowe w taki sam sposób, w jaki zostałoby ono wykonane w sytuacji TLU. Związek między znajomością języka a wiedzą ogólną, który w testach ogólnego przeznaczenia jest zwykle postrzegany jako zmienna zakłócająca, przyczyniająca się do błędu pomiaru i która powinna być w jak największym stopniu zredukowana, jest istotną i pożądaną cechą testów ESP. Wiedza o konkretnym celu jest w nich „niezbędną, integralną częścią koncepcji umiejętności językowych dla celów specjalistycznych” (Douglas 2002: 2 – tłum. własne). Co więcej, można argumentować, że ponieważ testy wypowiedzi pisemnej w kontekście pracy lub studiów są pisane dla określonego odbiorcy, na określony temat związany z dyscypliną i w określonym celu, nie ma konfliktu między oceną języka a treścią, co z kolei zwiększa trafność oceny (Dudley-Evans i St John 1998: 225).

Początki egzaminów ESP sięgają 1913 roku, kiedy to University of Cambridge Local Examinations Syndicate (UCLES) przeprowadził egzamin Certificate of Proficiency in English, który w założeniu miał być testem z języka angielskiego specjalistycznego dla celów zawodowych (*English for Occupational Purposes, EOP*) dla kandydatów zagranicznych ubiegających się o stanowiska nauczycieli języka angielskiego w szkołach zagranicznych i służył do oceny ich biegłości językowej. Egzamin ten miał jednakże charakter akademicki i nie badał sprawności językowych potrzebnych do wykonywania zawodu nauczyciela języka angielskiego jako obcego. Początkowo był on wzorowany na tradycyjnym, opartym na esejach programie nauczania języka ojczystego. Egzamin pisemny obejmował pracę z literatury angielskiej, esej, pracę z fonetyki, część gramatyczną oraz tłumaczenie z francuskiego i niemieckiego na angielski, a także z angielskiego na francuski i niemiecki. Istniał również komponent ustny z dyktandem, czytaniem na głos i konwersacją (Weir 2003: 2).

Innym testem ESP, który można uznać za pierwszy test z języka angielskiego dla potrzeb akademickich (*English for Academic Purposes, EAP*), był egzamin Komisji Egzaminów Wstępnych na studia (College Entrance Examination Board) – English Competence, wprowadzony w USA w 1930 r. jako test dla międzynarodowych kandydatów do amerykańskich szkół wyższych i uniwersytetów. Test miał na celu ustalenie, czy potencjalny student był w stanie używać języka angielskiego w kontekście akademickim. Składał się z czterech komponentów: czytania czterech krótkich tekstów, którym towarzyszyły pytania typu prawda/fałsz, oraz dłuższego testu sprawdzającego rozumienie głównych myśli tekstu; dyktanda; pisanie eseju; rozmowy na jeden z wcześniej określonych tematów (Fulcher [2003] 2014: 2).

Oba opisane testy mają dobrze zdefiniowane cele związane odpowiednio z zawodowym i akademickim językiem angielskim, a zatem można je uznać za przykłady pierwszych testów ESP. Jednak, jak wspomniano powyżej, rzeczywiste testowanie ESP należy, po pierwsze, opierać przede wszystkim na analizie sytuacji użycia języka docelowego TLU, z której powinny wynikać istotne cechy treści testu i metodologii oceny. Po drugie, testowanie języka specjalistycznego powinno obejmować interakcję między znajomością języka a wiedzą

merytoryczną z danej dziedziny. Żaden z tych testów nie został opracowany na podstawie analizy języka używanego w nauczaniu lub w kontekście nauki akademickiej, a zadania w testach nie przypominały zadań wymaganych w pierwszym przypadku od nauczycieli i w drugim – od studentów (Douglas 2002: 3).

Faktyczne testowanie ESP rozpoczęło się od egzaminu Temporary Registration Assessment Board (TRAB), który został wprowadzony w 1975 roku przez brytyjską Naczelną Izbę Lekarską (General Medical Council) (Douglas 2002: 3–4). Jego celem była ocena umiejętności zawodowych i językowych lekarzy wyszkolonych poza Wielką Brytanią, którzy ubiegali się o tymczasowe stanowiska lekarskie w Wielkiej Brytanii. Egzamin miał sprawdzić zarówno kompetencje zawodowe, jak i umiejętności komunikowania się w języku angielskim. Komponent językowy składał się z zadań ze słuchu, pisemnego eseju i ustnej rozmowy, w której oceniano zarówno wiedzę zawodową, jak i umiejętności językowe. Komponent językowy TRAB został skonstruowany w wyniku analizy języka używanego przez lekarzy, pielęgniarki i pacjentów w brytyjskich szpitalach. Jak wspomniano powyżej, to analityczne podejście jest cechą wyróżniającą opracowywanie testów ESP. Co więcej, twórcy testów, którzy byli odpowiedzialni za komponent językowy TRAB, współpracowali z ekspertami medycznymi przy konstruowaniu zadań. Jest to ważny aspekt opracowywania testów specjalistycznych, ponieważ dziedziny specjalistyczne wykraczają poza obszary, w których lingwista ma kompetencje do wydawania niezależnych osądów (Rea-Dickins 1987: 196). Warto zauważyć, że twórcy TRAB starali się umożliwić interakcję między umiejętnościami językowymi zdających a wiedzą ogólną w zadaniach testowych poprzez zapewnienie odpowiednich cech kontekstowych w materiale testowym. TRAB nie jest już używany, ale stanowi dobry model do opracowywania testów ESP (Douglas 2002: 4–5).

Najprawdopodobniej pierwszym egzaminem z języka angielskiego dla celów akademickich (EAP), który został opracowany w wyniku szczegółowej analizy potrzeb językowych w zakresie studiów uniwersyteckich w języku angielskim, był Test in English for Educational Purposes (TEEP) wprowadzony przez The Associated Examining Board w 1983 roku. Badania leżące u podstaw tego egzaminu rozpoczęły się w 1970 roku i były oparte na procesorze potrzeb komunikacyjnych Johna Munby'ego (1978). Zebrano dane dotyczące trudności, jakie napotykają zagraniczni studenci w pisaniu tekstów akademickich. Problemy te zostały zbadane zarówno z punktu widzenia studentów, jak i wykładowców. W badaniach skoncentrowano się na następujących dziedzinach: sztuka, nauki społeczne, administracja i biznes oraz nauki ścisłe i inżynieria. Egzamin TEEP składał się z trzech części: czytania ze zrozumieniem, pisania i rozumienia mówionego języka angielskiego. W wyniku szczegółowej analizy potrzeb zadania egzaminacyjne zostały zintegrowane i powiązane tematycznie. Na przykład zadanie pisemne zostało tak skonstruowane, aby zdający wykorzystał notatki, które miał zrobić podczas słuchania i czytania, co odzwierciedlało rzeczywiste używanie języka przez studentów podczas studiów akademickich, a tym samym miało na celu zwiększenie trafności egzaminu (Dudley-Evans i St John 1998: 216; Jordan 2001: 176). Obecnie TEEP jest dostępny na Uniwersytecie w Reading i składa się z sześciu części: znajomość języka, zadanie tematyczne (etap przygotowawczy do zadania pisemnego), czytanie, słuchanie, pisanie i mówienie.

Rodzaje testów ESP

Trzy rodzaje narzędzi oceniania mają szczególne znaczenie w testowaniu specjalistycznym: testy poziomujące, testy postępów/osiągnięć i testy biegłości (Alderson 1981: 123; Hutchinson i Waters 1987: 145–152).

Test poziomujący ESP ma na celu określenie stanu wiedzy osoby uczącej się przed rozpoczęciem kursu ESP, tj. czy osoba ta musi w ogóle uczestniczyć w kursie, a jeśli zajdzie taka potrzeba, test powinien wskazać, jaki rodzaj kursu powinna zrealizować. W pewnym sensie można go więc uznać za przykład testu biegłości, ponieważ jeśli umiejętności językowe

przystępującej do niego osoby zostaną uznane za wystarczająco zaawansowane, osoba ta nie będzie miała potrzeby uczestniczenia w kursie ESP (Hutchinson i Waters 1987: 146).

Test osiągnięć i postępów w ESP jest ściśle związany z kursem angielskiego specjalistycznego. Główny cel stanowi w tym przypadku sprawdzenie, czy i w jakim stopniu założenia kursu zostały osiągnięte, a zatem odzwierciedlają one jego program nauczania (Alderson 1981: 123). Odpowiedni sposób oceny osiągnięć uczącego się powinien obejmować treści przypominające te, które były nauczane na kursie, oraz ocenę umiejętności produktywnych i receptywnych, a także interaktywnych i mediacyjnych, zgodnie ze sposobem, w jaki materiał był wprowadzany. Weryfikacja polega na sprawdzeniu, jak dobrze uczący się wykonują zadania językowe, używając języka angielskiego w kontekście specjalistycznym, tzn. w kontekście akademickim (EAP) bądź zawodowym (EOP). Dlatego konieczne jest dostosowanie technik oceny do treści kursu specjalistycznego. W wielu sytuacjach formalne testy mogą być nieodpowiednie, szczególnie w przypadku krótkich intensywnych kursów, gdy cały dostępny czas jest wykorzystywany do wprowadzenia treści i ćwiczenia wiedzy oraz umiejętności, a faktyczny postęp może być zaobserwowany po zakończeniu nauki w rzeczywistych sytuacjach użycia języka docelowego (TLU). Ocena w klasie jest kolejnym przypadkiem, w którym formalne testy nie powinny być nadmiernie wykorzystywane. Powinna się ona raczej opierać na ciągłej obserwacji i ocenie kształtującej, która koncentruje się na pracy uczących się wykonywanej w klasie oraz w domu przez pewien okres, zwykle semestr lub czas trwania kursu językowego. Postępy podczas pracy w klasie w kontekście ESP, a zwłaszcza w języku angielskim dla celów zawodowych (EOP), mogą być oceniane zarówno przez nauczyciela, jak i uczących się (samoocena i ocena rówieśnicza). Ten rodzaj oceny jest elastyczny, kształtujący, przyczynia się do rozwoju technik uczenia się kursantów i pomaga im w uzyskaniu większej samodzielności w nauce języka (Dudley-Evans i St John 1998: 210–213).

Testy biegłości opierają się na ocenie umiejętności językowych zdającego pod kątem stopnia, w jakim spełniają one kryteria uważane za kluczowe dla biegłości wymaganej do wykonania określonego zadania lub skutecznego funkcjonowania w sytuacji TLU. Testy biegłości językowej w dziedzinie ESP mają wysoką trafność fasadową, ponieważ raport z wykonanego eksperymentu wydaje się znacznie trafniejszym wskaźnikiem umiejętności językowych naukowca niż zadanie wymagające od niego napisania krótkiego opowiadania (Alderson 1981: 125; Hutchinson i Waters 1987: 150). Jedną z głównych zalet tych testów jest to, że tworzą profil umiejętności zdającego, który można dopasować do wymagań komunikacyjnych danego kierunku studiów. Może to następnie zachęcić instytucje edukacyjne do wyraźnego określenia oczekiwań lub wymagań w zakresie języka, których spełnienie pomoże słuchaczom odnieść sukces w pracy lub w nauce (Alderson 1981: 125). Chociaż testy biegłości ESP wydają się uzasadnionymi instrumentami oceny umiejętności językowych, ich stosowanie rodzi też pytania:

1. Jakie przyjąć kryteria? Czy powinny one być skorelowane z różnymi obszarami tematycznymi (np. angielski dla architektów lub angielski dla inżynierów chemików)? Czy powinny odnosić się też do bardziej powszechnych sytuacji związanych z językiem, np. umiejętności opisywania diagramów lub wyjaśniania procesów? Czy test z angielskiego dla inżynierów elektryków może być ważnym wskaźnikiem umiejętności językowych w innych gałęziach inżynierii? (Alderson 1981: 125; Hutchinson i Waters 1987: 150–151).
2. Jakie umiejętności i wiedza są niezbędne do wykonania określonego zadania w prawdziwym życiu? Jak konkretne powinny być te umiejętności? Jaki jest poziom specjalistyczności/ogólności, na którym wiedza przedmiotowa jest oceniana zamiast wiedzy językowej lub umiejętności komunikacyjnych? (Alderson 1981: 126; Hutchinson i Waters 1987: 150)

Nie ma jednoznacznych odpowiedzi na te pytania, ponieważ zarówno kryteria oceny, jak i umiejętności oraz znajomości języka wybrane do ocenienia, a także stopień specjalistyczności zadania powinny wynikać ze szczegółowego zbadania sytuacji TLU i określenia umiejętności i znajomości języka wymaganych do skutecznej komunikacji w tym konkretnym kontekście.

Trafność i autentyczność testu

Wśród językoznawców panuje zgoda co do tego, że umiejętności językowe zmieniają się zależnie od kontekstu i zadania testowego, a zatem ich interpretacje będą się także odpowiednio zmieniać. Na przykład, jeśli zdający otrzymaliby test na czytanie ze zrozumieniem, który zawierałby dwa artykuły – jeden o próbkowaniu i kwantyzacji w elektronice cyfrowej, a drugi o różnych metodach kontroli jakości produkcji, prawdopodobnie uzyskaliby dość różne wyniki dla każdego z tekstów, zwłaszcza jeśli byłiby to studenci zarządzania. Z drugiej strony, gdyby dwie grupy zdających – jedna studiująca elektronikę, a druga zarządzanie – otrzymały zadanie testowe, w którym należałoby opisać wyniki eksperymentu w swojej dziedzinie studiów, interpretacje wyników obu grup mogłyby być porównywalne. Można zatem argumentować, że kryteria oceny i poziom specjalistyczności testu biegłości powinny być ściśle skorelowane z cechami TLU: jeśli sytuacja TLU wymaga czytania tekstów i podręczników w określonej dziedzinie, wówczas jako materiały testowe należy wybrać treści specjalistyczne. Jeśli jednak sytuacja TLU wiąże się z podjęciem studiów uniwersyteckich w dziedzinie inżynierii, a test jest wykorzystywany do celów rekrutacji, preferowane powinno być podejście koncentrujące się na ogólnych akademickich umiejętnościach językowych.

Co więcej, nie wystarczy po prostu dostarczyć zdającym tematy i zadania związane z kierunkiem ich studiów lub pracą. Egzaminowanie powinno bowiem angażować uczestników w taki sposób, aby „zarówno umiejętności językowe, jak i wiedza z danej dziedziny wchodziły w interakcję z treścią testu w sposób podobny do sytuacji TLU” (Douglas 2002: 6–7 – tłum. własne). Innymi słowy, zadania muszą być autentyczne, reprezentatywne dla określonej dziedziny i powinny wymuszać realistyczną komunikację językową.

Autentyczność, określana przez pryzmat reakcji użytkownika języka na zadanie testowe, zawiera dwa aspekty: sytuacyjny i interakcyjny (Bachman 1990: 10). Pierwszy odnosi się do cech zadania wynikających z analizy zadań rzeczywistych w TLU. Drugi wiąże się z zaangażowaniem umiejętności językowych zdającego podczas wykonywania specjalistycznego zadania testowego. Zakres, w jakim zdający angażuje konkretną wiedzę merytoryczną do wykonania zadania testowego i reaguje na sytuację użycia języka docelowego, ma kluczowe znaczenie w ocenianiu ESP, ponieważ określa stopień specjalistyczności testu. W związku z tym testów nie można po prostu klasyfikować jako ogólnego lub specjalistycznego przeznaczenia, ale należy je raczej umieścić, jak to ujął Douglas, na „kontinuum specjalistycznego przeznaczenia od ogólnego do wysoce specjalistycznego” (2002: 1).

W przypadku specjalistycznych testów językowych ważne jest, aby były w nich obecne zarówno autentyczność sytuacyjna, jak i interakcyjna, ponieważ jest całkiem możliwe, że zadanie testowe będzie postrzegane przez zdającego jako niezwiązane z jego dziedziną nauki, ale mimo to dość interesujące, a tym samym angażujące jego komunikacyjne umiejętności językowe. Takie zadanie może prowadzić do wyników, które da się zinterpretować jako dowód komunikacyjnych umiejętności językowych zdającego, ale nie umiejętności językowych w dziedzinie specjalistycznej. Innymi słowy, można je umieścić bliżej ogólnego końca kontinuum specjalistyczności. Zadanie testowe może też jednak doskonale przypominać to wykonywane w sytuacji użycia języka docelowego TLU, a mimo to nie angażować zdającego w znaczący sposób w komunikacyjne użycie języka. Na przykład zestaw instrukcji do przeprowadzenia eksperymentu chemicznego może być autentycznym materiałem, ale gdy jest wykorzystywany do opracowania testu językowego wielokrotnego wyboru w celu zmierzenia znajomości słownictwa lub czasowników w trybie rozkazującym, nie jest

używany autentycznie, to znaczy nie w sposób, w jaki użyłby go nauczyciel w laboratorium chemicznym. Tak więc dla studenta chemii, który przystępuje do testu EAP, poziom autentyczności sytuacyjnej byłby dość wysoki, podczas gdy poziom autentyczności interakcyjnej – niski (Douglas 2002: 18). Dlatego też, aby test znalazł się wysoko na skali spektrum specjalistyczności, konieczne jest najpierw określenie krytycznych cech sytuacji TLU pod względem kontekstu i rodzaju zadania, a następnie opisanie, w jaki sposób wpłynie to na wyniki zdającego, tak aby zaangażować jego umiejętności językowe i specjalistyczną wiedzę językową w celu przekazywania znaczeń. Skuteczne działanie w takim przypadku może być interpretowane jako dowód komunikacyjnych umiejętności językowych dotyczących sytuacji docelowej, tj. w zakresie języka obcego specjalistycznego.

Testy ESP mogą być przygotowywane zarówno jako kryterialne, jak i normatywne. Niemniej jednak te pierwsze pozwalają spojrzeć na testowanie ESP z ważnej perspektywy, tj. wymogu precyzyjnego określenia poziomu umiejętności językowych i dziedziny treści, które posłużyłyby jako kryteria dla oceny wykonania zadania (Douglas 2002: 17). Z tego powodu opracowywanie testu opartego na kryteriach, które obejmuje szczegółową analizę i specyfikację nie tylko cech konkretnej sytuacji użycia języka docelowego, lecz także kryteriów oceny wykonania zadania, jest niezwykle korzystne dla rozwoju testów ESP. Tradycyjnie skale oceny odzwierciedlają orientację językową, koncentrując się na takich aspektach, jak: gramatyka, spójność lub płynność. Jednak przy opracowywaniu kryteriów oceny ESP przydatne jest zastosowanie etnografii ugruntowanej i zbadanie autentycznych (indygenicznym) kryteriów oceny (Douglas 2002: 10).

Etnografia ugruntowana to badanie służące zrozumieniu i opisaniu sytuacji TLU z perspektywy użytkowników języka (Douglas 2002: 93). Technika ta polega na filmowaniu wybranych osób w rzeczywistych sytuacjach komunikacyjnych w obszarze TLU, a następnie nagrywaniu ich komentarzy na temat przebiegu zdarzeń komunikacyjnych i ich refleksji po obejrzeniu i wysłuchaniu materiału. Nagrania wideo są również komentowane przez ekspertów z badanej dziedziny oraz przez innych badaczy, np. językoznawców lub antropologów, którzy mogą zapewnić inne perspektywy analizy. Z drugiej strony ocena indygeniczna jest opracowywana i stosowana przez specjalistów z danej dziedziny w sytuacjach TLU w celu zmierzenia sukcesu komunikacyjnego praktykantów w dziedzinach akademickich i zawodowych (Jacoby 1998, za: Douglas 2002: 68). Badanie różnych rodzajów działań w środowisku zawodowym i profesjonalnym, a także towarzyszących im technik i kryteriów oceny za pomocą etnografii ugruntowanej lub badania oceny indygenicznej może zapewnić cenny wgląd w naturę wymogów TLU, pokazać ich postrzeganie przez użytkowników języka i pomóc w ustaleniu odpowiednich kryteriów dla specjalistycznych testów językowych.

Na koniec, podchodząc do konstruowania testu ESP, warto określić rodzaj wniosków, które zostaną wyciągnięte na podstawie wyników testów osób egzaminowanych. Według Tima McNamary (1996: 8) mogą one dotyczyć zdolności do wykonywania przyszłych zadań lub pracy w sytuacji użycia języka docelowego (mocna hipoteza dot. działania językowego) bądź zdolności do używania języka w określonych przyszłych zadaniach lub pracy (słaba hipoteza dot. działania językowego). Rozróżnienie to ma kluczowe znaczenie dla teoretycznych podstaw testów ESP, ponieważ, jak ostrzega McNamara, na pierwszy rodzaj wnioskowania, tj. wydajność pracy, może mieć wpływ kilka czynników niezależnych od zdolności językowych, np. osobowość lub odpowiedzialność. Z pewnością testy ESP powinny być zgodne z założeniami hipotezy słabej i umożliwiać egzaminatorom wyciąganie wniosków na temat umiejętności językowych, a nie wydajności pracy. Trzeba jednak podkreślić, że gdy twórca definiuje konstrukt testu, powinien wziąć pod uwagę wymogi dotyczące sytuacji użycia języka docelowego, znajomości języka i specjalistycznej wiedzy językowej oraz wzajemnych powiązań między nimi.

ESP w kontekście akademickim – przykłady

Ocena ESP opiera się na trzech elementach: celu, znaczeniu i autentyczności. W związku z tym doskonale nadaje się do oceny opartej na zadaniach, w której nacisk kładziony jest na symulację rzeczywistej sytuacji użycia języka i mobilizację wiedzy językowej w celu wyrażenia znaczenia, a nie manipulowania formą (Nunan 2004: 4). Z tego powodu testy i ocenianie ESP nie powinny się skupiać na sprawdzaniu wiedzy językowej i znajomości odosobnionych elementów języka. W istocie powinny raczej badać określone, zintegrowane umiejętności i działania komunikacyjne, ponieważ w ten sposób zadania językowe są przeprowadzane w rzeczywistości pracy i studiów akademickich. Egzaminy powinny symulować rzeczywiste działania, jakie zdający będą musieli podjąć w miejscu, gdzie będą oceniani na podstawie efektów swojej pracy, a nie umiejętności czytania lub słuchania. Aby egzamin był autentyczny, powinno się oceniać głównie sprawności produktywne, wspierane tekstami pisаныmi i słuchanymi (Fischer i in. 2011: 27, za: Sendur 2020: 27).

W tej części artykułu przedstawię **dwa przykłady zadań** służących do oceniania umiejętności posługiwania się językiem angielskim dla celów akademickich (EAP), które zostały przygotowane dla studentów kierunku Elektroenergetyka na Politechnice Poznańskiej. Zadania te mają na celu pomiar sprawności produktywnych w zakresie pisania (przykład 1) i mówienia (przykład 2). Zostały one wykorzystane na egzaminie certyfikатовym Academic Certificate in English Language (ACERT) na poziomie B2 w skali *Europejskiego systemu opisu kształcenia językowego* (ESOKJ). Egzamin ACERT Stowarzyszenia Akademickich Ośrodków Nauczania Języków Obcych w Polsce (SERMO) jest testem osiągnięć, który przeprowadza się na zakończenie kursu języka obcego, zwykle na studiach pierwszego stopnia. ACERT został opracowany w latach 2011–2014 w odpowiedzi na przepisy Ministerstwa Nauki i Szkolnictwa Wyższego (2011). ACERT jest obecnie przeprowadzany na siedmiu uczelniach: Politechnice Gdańskiej, Politechnice Koszalińskiej, Politechnice Łódzkiej, Uniwersytecie Medycznym w Łodzi, Politechnice Poznańskiej, Politechnice Warszawskiej i Politechnice Wrocławskiej.

Egzamin może być opracowany dla poziomów od A1 do C1 w skali ESOKJ. Niemniej jednak większość testów ACERT jest na poziomie B2, ponieważ zgodnie z Polską Ramą Kwalifikacji dla szkolnictwa wyższego w Polsce egzaminy językowe kończące studia pierwszego stopnia muszą być zdawane na poziomie B2. Egzamin, niezależnie od poziomu, składa się z pięciu części: rozumienie ze słuchu, rozumienie tekstu czytanego, słownictwo i gramatyka, pisanie oraz mówienie. Za każdą część można uzyskać maksymalnie 20 punktów, a za cały test 100 punktów. Części rozumienia ze słuchu i rozumienia tekstu czytanego zawierają po dwa teksty, z których każdy zawiera dziesięć elementów testowych. Część sprawdzająca znajomość słownictwa i gramatyki zawiera dwadzieścia elementów. W części z pisania znajdują się dwa zagadnienia, a zdający mogą wybrać, które chcą wykonać. Część z mówienia podzielona jest na dwa zadania: jedno wymagające interakcji i drugie wymagające wypowiedzi monologicznej. Czas przeznaczony na egzamin jest różny na różnych poziomach: A1, A2 i B1 – 90 minut, B2 i C1 – 120 minut.

Uczelnie, które chcą przeprowadzać egzamin ACERT i stać się akredytowanym centrum ACERT, muszą przyjąć strukturę, procedury administracyjne, system punktacji i kryteria oceny egzaminu opracowane w ramach SERMO. Zawartość testu i elementy testowe muszą zostać opracowane lokalnie przez komisje egzaminacyjne w akredytowanych ośrodkach, aby spełniały wymagania programowe na danych uczelniach, ponieważ ACERT weryfikuje osiągnięcia i powinien być dostosowany do nauczanych treści. Testy są analizowane przez centralną komisję akredytacyjną, która składa się z ekspertów w dziedzinie uczenia i testowania języków obcych. Przed przeprowadzeniem egzaminu komisja weryfikuje jakość zadań egzaminacyjnych i dostosowanie ich poziomu do skali ESOKJ. Jeśli jakość testu jest wystarczająca, zostaje on dopuszczony do użytku, a akredytacja utrzymana. Ponadto w Centrum Języków i Komunikacji Politechniki Poznańskiej przeprowadzane są badania walidacyjne (jakościowe i ilościowe) *a posteriori* w celu dalszego ulepszania testów i ich elementów.

Egzamin ACERT B2 na Politechnice Poznańskiej przeznaczono dla studentów, którzy ukończyli 120-godzinny kurs języka obcego na studiach pierwszego stopnia. Jego celem jest sprawdzenie, czy osiągnęli oni poziom B2 kompetencji językowych w zakresie słuchania, czytania, pisania i mówienia, tak aby skutecznie komunikować się w kontekście akademickim i zawodowym. Politechnika Poznańska prowadzi 9 wydziałów z 33 kierunkami studiów. Centrum Języków i Komunikacji prowadzi kursy z języka specjalistycznego, a następnie przeprowadza egzaminy dostosowane do każdego kierunku studiów we współpracy z władzami wydziałów. Egzamin składa się zarówno ze wspólnych zadań ogólnoakademickich, jak i zadań specjalistycznych związanych z każdym kierunkiem studiów. Zadania specjalistyczne są tworzone zgodnie z wewnętrznymi wytycznymi i specyfikacjami.

Od zdających oczekuje się umiejętności interakcji ze studentami i wykładowcami z innych krajów, czytania literatury przedmiotu, uczestnictwa w wykładach i projektach międzynarodowych oraz przygotowywania prezentacji w języku obcym. Przyszli inżynierowie z różnych dziedzin powinni być w stanie skutecznie komunikować się przy użyciu odpowiedniego rejestru językowego i przedstawiać profesjonalne prezentacje. Kurs językowy i egzamin mają na celu poprawę umiejętności komunikacyjnych studentów w obszarach akademickich i dziedzinowych, zwiększenie mobilności i przygotowanie ich do międzynarodowego rynku pracy.

W artykule przedstawiłam definicję testów ESP wraz z omówieniem warunków ich realizacji. Specjalistyczne egzaminy językowe w kontekście pracy lub studiów są tworzone dla konkretnej grupy odbiorców, na określony temat związany z dyscypliną i z myślą o określonym celu, nie ma zatem konfliktu między oceną języka a treścią, co zwiększa trafność i rzetelność oceny znajomości języka.

BIBLIOGRAFIA

- Alderson, J.C. (1981), *Report of the Discussion on Testing English for Specific Purposes*, [w:] J.C. Alderson, A. Hughes (red.), *ELT documents 111: Issues in Language Testing*, London: The British Council, s. 123–134.
- Bachman, L.F. (1990), *Fundamental Considerations in Language Testing*, Oxford: Oxford University Press.
- Douglas, D. (2002), *Assessing Languages for Specific Purposes*, New York: Cambridge University Press.
- Dudley-Evans, T., St John, M. (1998), *Developments in English for Specific Purposes: A Multi-Disciplinary Approach*, Cambridge: Cambridge University Press.
- MNiSW (2011), Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 2 listopada 2011 r. w sprawie Krajowych Ram Kwalifikacji dla Szkolnictwa Wyższego (Dz. U nr 253, poz. 1520).
- Fulcher, G. [2003] (2014), *Testing Second Language Speaking*, wyd 2, New York: Routledge.
- Hutchinson, T., Waters, A. (1987), *English for Specific Purposes*, Cambridge: Cambridge University Press.
- Jordan, R.R. (2001), *A History of English for Academic Purposes (EAP): From small Beginnings*, [w:] E.G. Sanchez (red.), *Present and Future Trends in TEFL*, Almeria: Universidad de Almeria, s. 169–178.
- McNamara, T.F. (1996), *Measuring Second Language Performance*, Harlow: Longman.
- Munby, J. (1978), *Communicative Syllabus Design*, Cambridge: Cambridge University Press.
- Nunan, D. (2004), *Task-Based Language Teaching*, Cambridge: Cambridge University Press.
- Rea-Dickins, P. (1987), *Testing Doctors' Written Communicative Competence: An Experimental Technique in English for Specialist Purposes*, „Quantitative Linguistics”, nr 34, s. 185–218.

- Sendur, A.M. (2020), *Testowanie i certyfikacja w glottodydaktyce języków ogólnych i specjalistycznych*, Kraków: Oficyna Wydawnicza AFM.
- Weir, C.J. (2003), *Continuity and Innovation: Revising the Cambridge Proficiency in English Examination 1913–2002 Cambridge Local Examinations Syndicate*, [w:] C.J. Weir, M. Milanovic (red.), *Studies in Language Testing*, Cambridge: Cambridge University Press, s. 305–309.

DR BARBARA SAWICKA Absolwentka filologii angielskiej Uniwersytetu im. Adama Mickiewicza w Poznaniu i doktor nauk humanistycznych w zakresie oceniania umiejętności językowych. Uczestniczyła w wielu projektach krajowych i międzynarodowych, takich jak projekt ECML VITbox i EUNICE. Ma bogate doświadczenie w nauczaniu języka angielskiego dla celów specjalistycznych, jest autorem kilku artykułów i podręcznika do nauki języka angielskiego dla studentów elektrotechniki i energetyki oraz opracowała wiele kursów online dla studentów kierunków technicznych.

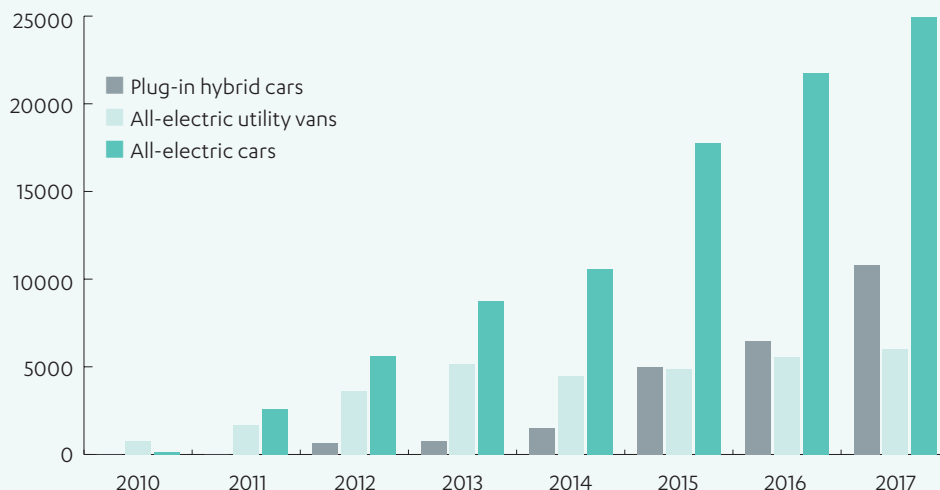
Artykuł został pozytywnie zaopiniowany przez recenzenta zewnętrznego „JOWS” w procedurze *double-blind review*.

Sprawności produktywne: wypowiedź pisemna

Przykład 1. Opis i interpretacja wykresu

Imagine you participate in a research project which examines trends in electric vehicle sales. You have been asked to write an analysis of the registrations of plug-in electric vehicles in EU from 2010 to 2017. You have already gathered information and plotted a graph. In about 150-200 words describe and interpret the graph shown below. Use appropriate vocabulary and different expressions. Use tenses consistently. In your composition:

- ➔ describe the trends shown below
- ➔ compare the trends
- ➔ give possible reasons for the changes in the trends, e.g. charging time



Źródło: www.wikipedia.com.

Kryteria oceny:

- wykonanie zadania (spełnienie wymagań określonych w poleceniu i powiązanych z typem tekstu: przeanalizowanie wykresu, porównanie trendów, wyciągnięcie wniosków);
- spójność i logiczność wypowiedzi;
- zakres i poprawność słownictwa ogólnego i specjalistycznego;
- zakres i poprawność gramatyczna.

Sprawności produktywne: wypowiedź ustna

Przykład 2. Klasyfikowanie, opisywanie i przewidywanie przyszłych trendów w wytwarzaniu energii elektrycznej

Imagine that as a part of project work, you were asked by your tutor to prepare an overview and an evaluation of energy resources used in electricity generation. During your presentation cover the following aspects:

- ➔ classification of resources (renewables vs. non-renewables);
- ➔ advantages and disadvantages of using renewable energy sources;
- ➔ advantages and disadvantages of using non-renewable energy sources;
- ➔ expected future developments in electricity generation.

Kryteria oceny:

- treść (wykonanie zadania i wiedza specjalistyczna – zgodnie z kryteriami: umiejętność opisanie i sklasyfikowania zasobów energetycznych: przedstawienie zalet i wad oraz wypowiedź na temat trendów w rozwoju energetyki);
- płynność;
- zakres i poprawność słownictwa ogólnego i specjalistycznego;
- zakres i poprawność gramatyczna.