

*Hanna Komorowska*

Uniwersytet SWPS, Warszawa  
<https://orcid.org/0000-0002-9395-2734>  
[hkomorowska@swps.edu.pl](mailto:hkomorowska@swps.edu.pl)

## OCENIANIE W NAUCZANIU JĘZYKÓW OBCYCH. FAKTY, MITY I TRUDNOŚCI

Facts, myths and difficulties in the evaluation  
of educational attainment

The article presents selected psychological and sociological aspects of the process of evaluating educational attainment in language education. Attention is given to stereotypical and ahistorical opinions connected with the role of assessment in grammar-translation and audiolingual methods as well as in the communicative approach. Distortions are also discussed such as introducing testing techniques of a practical value in mass education as solutions responding to authentic communicative needs of the learners or presenting formative evaluation as one of the recent educational innovations. Several myths which unify the academic community and facilitate interaction are examined and distinguished from complex issues and difficulties such as assessing intercultural competence or objectivizing the evaluation of interactive skills. The text suggests that misrepresentation of the past serves to create an image of continuous progress in SLA/FLT as well as to paint a more optimistic picture of the present educational context with assessment as its integral element.

Keywords: history of ELT, language learning, language teaching, assessment in SLA/FLT, testing techniques, fact vs. myth, assessment difficulties

**Słowa kluczowe:** historia nauczania języków obcych, przyswajanie języka, nauczanie języka, ocenianie wyników, fakty, mity, trudności w ocenianiu

## 1. Wstęp

Wagę oceniania trudno przecenić, szczególnie że nauczyciele regularnie wymieniają je jako jeden z dwu – obok zarządzania klasą – najistotniejszych, ale i najtrudniejszych elementów swej pracy (Czetwertyńska, 2015).

Ocenianie to problem, który ma – być może – najbardziej janusowe oblicze ze wszystkich aspektów procesu nauczania i uczenia się języków obcych.

Wszystkim znane jest jego *oblicze praktyczne* widoczne w pracy dydaktycznej. Oceniać – czy chce, czy nie chce – musi każdy nauczyciel: nauczyciele szkolni i lektorzy, gdyż tego wymaga od nich administracja, a nauczyciele prywatni i korepetytorzy – z własnej inicjatywy, gdyż inaczej nie mogliby zaplanować kolejnych lekcji języka. Obie grupy nauczycieli różni jedynie stopień formalizacji tego procesu, ale owa formalizacja jest kwestią o znacznej istotności.

Ocenianie ma też *oblicze teoretyczne*. Teoria ewaluacji wyników nauczania ma już długą historię, historia ta jednak – co publikacje na temat testowania zazwyczaj pomijają – jest w ogromnej mierze oparta na teorii programu nauczania. Teoretyczne rozważania o ocenianiu pojawiły się w naukach pedagogicznych w latach 60. ubiegłego wieku, kiedy to powstały pierwsze obszernie prace traktujące o taksonomii celów nauczania (Bloom i in., 1956, 1964). Wiązały się one ściśle z nieco wcześniej już wyraźnie zauważalnym, ale bardziej rozciągniętym w czasie, rozwojem teorii programu nauczania, a działało się tak dlatego, że zasady konstrukcji programu już od czasu ich sformułowania przez badaczy takich jak Tyler (1949), Taba (1963), Taylor i Richard (1985) czy Ornstein i Hunkins (1988) wymagały nie tylko klarownej prezentacji celów nauczania, ale także określenia sposobów sprawdzania, w jakim stopniu cele te zostały zrealizowane. Co więcej, zróżnicowane kategorie celów wymagały zróżnicowanych technik kontroli i oceny (Bloom i in., 1971). Prace nad tym zagadnieniem rozwijają się do dziś, czego dowodem prace autorów takich jak Alismail i McGuire (2015), Gordon i Rajagopalan (2016), Pinar (2004), Tedesco i in. (2014) czy Tsagari i Banerjee (2016).

Ocenianie ma również *oblicze socjologiczne*. Wielki jest jego wpływ społeczny, szczególnie widoczny w przypadku tak zwanych egzaminów doniosłych (*ang. high stakes examinations*), a więc tych, od których zależą dalsze losy edukacyjne i zawodowe osób przystępujących do tego rodzaju oceny. Przykładem może tu być matura decydująca o możliwości dostania się na uczelnię wyższą i pożądany kierunek studiów, a warto dodać, że wpływ tego egzaminu sięga również budżetu rodzinnego, gdyż słabiej zdany egzamin oznacza zazwyczaj konieczność płacenia za studia, czy to na uczelni państwowej, czy prywatnej. Nie bez znaczenia jest też ogromna rola stopy życiowej rodziny uczniów, co wiąże się z ich pozaszkolnym kontaktem z językiem obcym w toku dodatkowych lekcji

lub wyjazdów zagranicznych – pokazuje to dowodnie wskaźnik HOMEPOS (ang. *home possessions*) silnie różnicujący pod względem socjoekonomicznym 15-latków z poziomu A1 i B1/B2 w europejskim badaniu *SurveyLang* (2012).

Nie sposób też zapomnieć, że ocenianie ma również *oblicze psychologiczne*. Ogromny jest przecież wpływ etykietowania uczniów jako lepszych lub gorszych na ich samoocenę i poczucie sprawczości, a zatem i na podejmowanie przez nich wysiłku oraz na typ zadań, jakie zdecydują się podjąć w przyszłości, co od dawna określane jest mianem efektu Pigmaliona i efektu Go-lema (Rosenthal i Jacobsen, 1968; Rubie-Davies, 2009).

W kształceniu językowym ocenianie zyskało jeszcze jeden aspekt, nieobecny w innych dziedzinach edukacji, a mianowicie aspekt, który omawiam poniżej jako *aspekt mitologiczny*. Jego cele i techniki były na przestrzeni historii nauczania i uczenia się języków obcych niezwykle zróżnicowane, gdyż w pełni zależne od metod nauczania powszechnie stosowanych w danym okresie. Dlatego też szczególnie w obrębie kształcenia językowego panuje wśród badaczy i nauczycieli wiele poglądów i opinii na temat oceniania, jego historii i jego przydatności. Niektóre z nich funkcjonują już jako utarte przekonania, które warto chyba określić jako *mity*. Określenie to wymaga jednak chwili uwagi.

Mit to opowieść, która od najdawniejszych czasów ma wyjaśniać sens istotnych ludzkich doświadczeń. W definicjach etnoreligijnych opowieść ta jest związana ze zjawiskami nadprzyrodzonymi, lecz w definicjach antropologicznych, psychologicznych i socjologicznych niekoniecznie odnosi się do tej tematyki. W antropologii istnieje pogląd, iż skoro mit to opowieść o poszukiwaniu sensu, to owo poszukiwanie może, ale nie zawsze musi okazać się udane (Campbell, 2013). Nawet gdy nie jest udane, ma szansę okazać się przekonujące i stać się popularne. Być może dlatego termin ten w ostatnim stuleciu był używany częściej w odniesieniu do fabuł i opowieści nieprawdziwych. Dziś, użyty potocznie, oznacza przeważnie przesąd, zazwyczaj dość powszechny, ale też – jak w znanej definicji Rolanda Barthesa – prezentację pewnej strony zagadnienia, czy poglądu tak, jakby była jedyną możliwą perspektywą, a to zbliża się już do pojęcia błędnego przypuszczenia, odzwierciedlającego intencję, a niekiedy wręcz manipulację narratora (Barthes, 1957/2013). W takim znaczeniu będzie ono używane w dalszych częściach niniejszego tekstu.

Jaka jest zatem funkcja mitów w takim właśnie znaczeniu w środowiskach akademickich?

Do odpowiedzi na to pytanie przydatne okaże się ujęcie socjologiczne. W tym podejściu mit to narracja, która ukazuje pewną formę poznania, nadaje sens, ale taki, który jest zgodny z etosem danego środowiska. Umożliwia on więc łatwą komunikację w grupie, integruje społeczność poprzez swą bezkonfliktowość i ukazuje wartości wspólne oraz normy, które powinny być w tej

społeczności przestrzegane. Poprzez wyraźne wskazanie norm pełni tym samym funkcję kontrolną wobec rzeczywistości społecznej, ukazuje cele działań i eliminuje, czy choćby deprecjonuje, działania niezgodne z obowiązującą narracją. Efektem jest bezkrytyczne powtarzanie twierdzeń i opinii.

Dlaczego tak się dzieje? Na to pytanie odpowiedzieć można znanym cytatem z rozmowy francuskich pisarzy: Flaubert mówi do Maupassanta: "Jakaś część wszystkiego pozostaje niezbadana, ponieważ wpadliśmy w nawyk przypominania sobie, gdy tylko używamy oczu, co inni ludzie przed nami myśleli o rzeczy, na którą akurat patrzymy" (Steegmuller, 1949: 60). To, co powtarzamy na temat przeszłości, ma jednak pewną dodatkową funkcję – zawsze w jakiś szczególny sposób służy teraźniejszemu potrzebom. W swej książce *Harmonia Caelestis* Esterhazy mówi: "...to nie jest tak, że spokojnie wspominamy przeszłość, spacerujemy po niej, rzeczowo oceniamy.... Nie, teraźniejszość jest zawsze agresywna... i dlatego zanurza się w chaosie czasów, żeby z nich wyłowić to, czego jej potrzeba do jeszcze lepszego wypełnienia obecnej formy" (Esterhazy, 2008: 351). Teraźniejszość glottodydaktyki ma również pewne potrzeby i także nie jest pozbawiona pewnych mitów.

Jakie zatem mity funkcjonują w odniesieniu do oceniania w kształceniu językowym?

## 2. Popularne poglądy na temat oceny – utarte, ale czy uzasadnione. Kika mitów glottodydaktycznych

Wydaje się, że siedem utartych sądów, które określiłabym jako mity to przekonania szczególnie rozpowszechnione, a tym samym warte tu omówienia. Są one następujące:

- I. Dopiero współczesna ocena sformalizowała ocenianie i wprowadziła kryterium skuteczności.
- II. Ocenianie w metodzie gramatyczno-tłumaczeniowej było szkodliwe i bezużyteczne.
- III. Ocenianie ukierunkowane na żywy język zawdzięczamy metodzie audiolingwalnej.
- IV. Ocenianie jest dziś ukierunkowane na potrzeby ucznia.
- V. W podejściu komunikacyjnym oceniamy umiejętności przydatne w autentycznych sytuacjach.
- VI. Ocenianie kształtujące to innowacja w dydaktyce.
- VII. Kryterium native speaker już nie istnieje.

Przeanalizuję je tu kolejno.

*Mit I – współczesna ocena wprowadziła kryterium skuteczności komunikacyjnej*

Historia oceniania zaczęła się paradoksalnie w niezwykle współczesny sposób. Dziś mawiamy, że celem opanowywania języka jest realizowanie celów pozajęzykowych, a miarą umiejętności komunikacyjnych jest skuteczność. Wymagania, jakim należało sprostać w sytuacji starożytnych czy średniowiecznych kontaktów handlowych, by zyskać, a nie stracić; oczekiwania, jakie należało spełnić w położeniu niewolnika, by zachować życie; autopromocja, jaką należało zaplanować, by ze zbędnego jeńca stać się przydatnym tłumaczem – to wszystko pokazuje, że życie weryfikowało poziom umiejętności językowych względem efektywności w bardzo konkretnej pozajęzykowej codzienności. A jednak te niezwykle długie okresy w historii ludzkiej komunikacji – jak się teraz modnie mówi – egzolingwalnej, określamy mianem przestarzałych, wyłącznie ze względu na brak systematycznych i sformalizowanych rozwiązań. Pełne dane o sformalizowanych systemach egzaminacyjnych pochodzą jednak z bardzo wczesnych okresów, na przykład z Chin, gdzie w VII wieku naszej ery Cesarska Służba Administracyjna drobniawczo skodyfikowała system oceniania przyszłych urzędników państwowych tak skutecznie, że przetrwał do 1905 roku (Bachman, 2015). Skodyfikowane były także już w XVI wieku egzaminy uniwersyteckie na stopień doktora prawa kanonicznego; w roku 1503 zdawał je w Ferrarze Mikołaj Kopernik, gdy uzyskanie doktoratu w Padwie okazało się dla niego zbyt kosztowne (Goddu, 2010). Jako że o sformalizowanym ocenianiu umiejętności językowych wiemy istotnie niewiele aż do pierwszych etapów upowszechniania edukacji w XVIII wieku, o ocenianiu w dawnych czasach mówimy niepocholebnie. Jednak skuteczność komunikacji językowej w danej sytuacji i instrumentalny charakter umiejętności językowych wobec niejęzykowej rzeczywistości to kryteria, które – choć wówczas najczęściej subiektywne – rzeczywiście istniały i to do nich dziś w ewaluacji wracamy.

*Mit II – ocenianie w metodzie gramatyczno-tłumaczeniowej to bezużyteczny relikw przesłości, jak i sama ta metoda*

Wzorem nauki łaciny ocenianie znajomości języka pozostawało w zgodzie z wymogami metody gramatyczno-tłumaczeniowej i spoczywało na dwóch filarach – tłumaczenia i dłuższej wypowiedzi pisemnej, co zresztą w oryginale było skierowane do dobrze wykształconych studentów filologii klasycznej (Hovatt i Widdowson, 2004). Warto jednak zwrócić uwagę na fakt, iż tłumaczenie wcześniej sprawdzało nie tyle umiejętności tłumaczeniowe tak jak je dziś rozumiemy, ile inne kwestie: mianowicie czytanie ze zrozumieniem, a także umiejętność pisanie w języku ojczystym, a ta stanowiła trzon programowy ówczesnej

nauki, gdyż alfabetyzacja w społeczeństwie XIX wieku była jeszcze w stanie embrionalnym. Funkcja translacji była zatem zupełnie inna niż to się dziś uważa, a że teksty wybierane do tłumaczenia były wysokiej klasy tekstami literackimi, uczono więc zarazem kultury wysokiej, nieznannej społeczeństwu w ogromnej mierze niepiśmiennemu. Owe teksty literackie, niezwykle wartościowe, pisane były przez wykształconych rodzimych użytkowników języka bynajmniej nie dla celów pedagogicznych, a tym samym w dzisiejszej terminologii paradoksalnie nie sposób odmówić im autentyczności. Dlatego odsądzanie dziś od czci i wiary tłumaczeniowego oceniania w metodzie gramatyczno-tłumaczeniowej tego okresu, to dowód oceny ahistorycznej polegającej na przykładaniu dzisiejszej miary do ówczesnej sytuacji. Wówczas był to ogromny krok naprzód, a to z uwagi na stopień pewnego już sformalizowania oceny, społeczną wartość walki z powszechnym analfabetyzmem i promocję kultury literackiej. Drugi człon oceniania, jakim była dłuższa wypowiedź pisemna, to w dzisiejszej terminologii przykład oceniania holistycznego, nadal obecnego w edukacji, choć teraz już wraz z analitycznie budowaną oceną kryterialną.

Zarysowany tu obraz nie ma na celu, jak łatwo się zorientować, postulowania powrotu do metody gramatyczno-tłumaczeniowej w powszechnej nauce szkolnej. Zamiar tej analizy jest inny: widać z niej jak chętnie – nie tylko w polityce – stosujemy podwójne standardy tylko po to, by wykazać, że to, co niegdyś istniało, uznać należy za bezsensowne, a to co wprowadzamy dziś, jest bezdyskusyjnie lepsze.

### *Mit III – ocenianie ukierunkowane na żywy język zawdzięczamy metodzie audiolingwalnej*

Przyjrzyjmy się dalszemu rozwojowi wypadków. Po II wojnie światowej w krajach zachodnich upowszechniła się metoda audiolingwalna, zwana też strukturalną lub metodą Mim-mem, która – z uwagi na sytuację społeczno-polityczną tego okresu – dotarła do Polski już w złagodzonej formie dopiero ponad dwadzieścia lat później, w drugiej połowie lat 60. Wbrew powszechnym przekonaniom dopiero w tej właśnie metodzie audiolingwalnej zrealizowało się w pełni i znalazło w centrum uwagi oceniających to, o co zazwyczaj oskarża się metodę gramatyczno-tłumaczeniową, czyli koncentracja na nauce słownictwa i gramatyki. Stało się tak pomimo zapewnień, iż jej celem – jakoby w odróżnieniu od celów metody gramatyczno-tłumaczeniowej – jest nauka żywego języka i rozwijanie sprawności, a przyczyną tego zjawiska stał się ideał odejścia od subiektywizmu oceny. Pozytywne następstwo tych tendencji to narodziny testowania (Lado, 1961). Jednak dążenie do obiektywizacji oceny testowej w zrozumiałym sposobie ukierunkowano na mierzalność, a ta najlepiej sprawdzała

się w zero-jedynkowej punktacji stosownej dla testowania elementów języka. Pojawiło się zatem także następstwo negatywne: testowano przede wszystkim znajomość materiału leksykalnego i morfosyntaktycznego, a to postawiło poprawność językową wyżej niż to się kiedykolwiek wcześniej działo w czasach panowania metody gramatyczno-tłumaczeniowej.

Pomimo owej niezgodności stawianych sobie celów z zawartością testów językowych, wkład metody audiolingwalnej w rozwój metodologii ewaluacji jest nie do przecenienia. To tej bowiem metodzie zawdzięczamy powstanie w pełni sformalizowanych systemów oceniania, które później dopiero wzbogacono o nowe techniki dostosowane do bardziej interakcyjnych celów nauczania języka.

#### *Mit IV – nauczanie i testowanie ukierunkowane są dziś na potrzeby uczącego się*

Nauczanie zorientowane na ucznia to jedno z głównych założeń współczesnej glottodydaktyki. Postuluje ono, jak wiadomo, z jednej strony dobór celów i treści nauczania zgodnych z potrzebami komunikacyjnymi ucznia, z drugiej zaś – dostosowanie metod i technik nauczania do możliwości, stylów i strategii uczącego się, co powinno znaleźć także odzwierciedlenie w kontroli i ocenie wyników nauki (Nunan, 2012).

Dobór celów i treści programowych, a tym samym przedmiot oceniania, jest jednak zgodny z potrzebami komunikacyjnymi uczących się jedynie w przypadku nauczania języka obcego dla celów specjalistycznych, o czym świadczą liczne egzaminy branżowe. O takiej zgodności w przypadku nauki szkolnej trudno jednak mówić, a to ten kontekst stanowi ogromny procent kształcenia językowego na świecie. Jest zrozumiałe, że potrzeby komunikacyjne dzieci i młodzieży będą się kształtować w przyszłości, jednak bieżące zainteresowania z trudem przekładają się na dobór treści programowych, a w ogóle nie przekładają się na proces oceniania wyników nauki. Zawarte w podręcznikach i testach treści są jedynie odbiciem przekonania autorów o potrzebach i zainteresowaniach uczniów, dla których program wraz z zaplanowanymi narzędziami kontroli musi zostać przygotowany jeszcze przed rozpoczęciem kursu językowego.

Co więcej, egzaminy całosciowe, szczególnie zaś egzaminy doniośle, a także liczne egzaminy międzynarodowe nie uwzględniają potrzeb uczących się w odniesieniu do poszczególnych sprawności i kompetencji, a udział w nich wymaga poddaniu się przekrojowej ocenie w zakresie wszystkich podstawowych sprawności na wybranym poziomie. Owa wielosprawnościowość systemów egzaminacyjnych i związana z nią certyfikacja powoduje, iż nie sposób mówić w tej sytuacji o jakimkolwiek ukierunkowaniu na osobę uczącego się.

I tu, jak widać, ukierunkowanie nauczania i oceniania na osobę ucznia okazuje się raczej wartościowym postulatem niż cechą charakterystyczną współczesnego kształcenia.

*Mit V – podejście komunikacyjne ocenia użycie języka w autentycznych sytuacjach*

Podejście komunikacyjne, zapoczątkowane we wczesnych latach 70. zestawem programów funkcjonalno-komunikacyjnych przygotowanych przez zespół Uniwersytetu w Reading dla Rady Europy w Strasbourgu, a upowszechnione w Polsce we wczesnych latach 90. zmieniło nastawienie do oceniania umiejętności językowych. W podejściu tym zaczęto głosić autentyczność sytuacji komunikacyjnych, w których odbywa się ocena i autentyczność technik testowania jakoby silnie związanych z realiami aktów komunikacyjnych. Mimo to, bywa, iż analiza typowych testów określanych jako komunikacyjne nasuwa całkiem inne wnioski (Bachman i Palmer, 2010).

Jednym z typowych sposobów testowania, jakie przyniosło podejście komunikacyjne, jest na przykład technika „cloze”, polegająca na uzupełnianiu pewnej liczby wyrazów wcześniej usuniętych z tekstu przez jego konstruktörów. Technikę tę i zasady konstruowania za jej pomocą odpowiednich zadań testowych zapożyczono z badań nad uszkodzonymi starożytnymi manuskryptami, przede wszystkim pochodzącymi z Qumran. Badacze owych tekstów stwierdzili, że staranna rekonstrukcja jest możliwa, gdy w tekście brakuje nie więcej niż co siódmeo wyrazu, inaczej nie sposób zyskać pewności co do wierności oryginałowi. W teście językowym luk musi więc być mniej. Jak często jednak uczący się znajdzie się w autentycznej sytuacji, w której będzie rekonstruował uszkodzony tekst w obcym języku (Shohamy, 2015)? Widać wyraźnie, że technika ta z autentyzmem komunikacyjnym nie ma wiele wspólnego, a powodem jej częstego stosowania jest praktyczność i skuteczne różnicowanie poziomu językowego testowanych. Mówiąc o testach komunikacyjnych, chwalimy jednak zazwyczaj – nie zawsze słusznie – ich autentyzm. Inny przykład to technika porządkowania rozsypanych paragrafów w celu uzyskania logicznego tekstu. I znów, kiedy uczący się w codziennej sytuacji będzie wykonywał podobne zadanie? Być może jako redaktor naczelny odpowiedzialny za tekst przygotowany przez nierozgarniętego pracownika; prędzej jednak powierzy mu prostsze zadania lub wręcz zwolni go raczej niż wykona tego rodzaju porządkowanie po raz drugi. Równie popularna technika to dopasowywanie tytułów do poszczególnych paragrafów – także przydatna wyłącznie w pracy edytorskiej. Autentyzm sytuacji będących kontekstem oceniania i zadań stojących przed osobą testowaną jest więc w najwyższym stopniu wątpliwy.



I tu jednak, pomimo przewagi praktyczności nad autentyzmem, podejście komunikacyjne wniosło do metodologii ewaluacji bardzo wiele. Zmienił się przede wszystkim typ tekstów, na których testowany wykonuje badane operacje, pojawiło się testowanie ustne w postaci interakcji z egzaminatorem lub z innym testowanym, wprowadzono jasne kryteria oceny i klucze odpowiedzi dla oceniających pytania otwarte i zadania typu „cloze”, zróżnicowano też badane umiejętności (Sterna, 2016).

### *Mit VI – Ocenianie kształtujące to innowacja dydaktyczna*

Interesującym zjawiskiem w dyscyplinie, jaką jest glottodydaktyka, jest pewna idiosynkrazja dotycząca poczucia upływu czasu. W kwestiach cywilizacyjnych, a szczególnie technologicznych – inaczej niż w sferze emocji i moralności – świat niestychnie przyspieszył. W odniesieniu do wielu zagadnień z zakresu przyswajania języków obcych tak się nie stało. Przyczyny takiego zjawiska mogą być dwojakie. Jedną z nich to brak wiedzy o historii dyscypliny – opcja prawdopodobna w czasach, gdy wypada cytować tylko najnowsze publikacje. Drugą może być z kolei inicjatywa chwalebna w intencjach, lecz manipulacyjna – a mianowicie chęć skutecznego promowania rozwiązania, które pomimo swej znacznej wartości nie przyjęło się dotąd na szerszą skalę.

Przykładem jest tu właśnie ocenianie, a konkretnie ocenianie kształtujące, o którym twierdzi się w naszej dziedzinie dość powszechnie, iż jest nowe, a zatem stanowi wartą upowszechniania innowację. Ocenianie kształtujące jest jednak jeszcze o kilka lat starsze niż blisko dziś pięćdziesięcioletnie podejście komunikacyjne.

Ocenianie kształtujące zostało po raz pierwszy zaprezentowane pod nazwą *formative evaluation* przez Scrivena blisko 50 lat temu, choć odnosiło się wtedy nie do oceny osiągnięć indywidualnego ucznia, a do oceny kursów, programów i materiałów nauczania (Scriven, 1967). Jednak fakt, iż ten – jak się potem okazało – bardzo znaczący dla dyscypliny tekst pojawił się niedługo po dwóch wspomniany wyżej tomach Benjamina Blooma o taksonomii celów nauczania (Bloom i in., 1956, 1964), skupił na sobie niezwykle zainteresowanie i wywarł wpływ na dalsze istotne prace. Niedługo, bo cztery lata później, opublikowany został ogromny tom poświęcony ocenie procesu uczenia się i wyników nauczania, noszący tytuł *Handbook of Formative and Summative Evaluation of Student Learning* (Bloom, Hastings i Madaus, 1971), który na wiele lat stał się biblią dydaktyków różnych dyscyplin.

Ocenianie kształtujące nie było przy tym zjawiskiem charakterystycznym tylko dla świata anglosaskiego. Piśmiennictwo francuskie i niemieckie przyniosło obfitą literaturę poświęconą tej formie oceny już w latach 70. XX

wieku, a od tego czasu prace teoretyczne poświęcone ewaluacji formatywnej – poprzez lata 80. i 90. (Cardinet, 1977; De Ketele, 1986; Herrmann i Höfer, 1999; Weiss, 1991) – ukazują się systematycznie do dziś (Allal, 2015; Earl, 2013; Köller, 2015). Już dwadzieścia lat temu badacze zainteresowani tym tematem wyszukali 580 tekstów na temat oceniania kształtującego tylko z poprzedzającej tę datę dekady i przeanalizowali 250 z nich (Black i William, 1998). Imponująca liczba aż 800 studiów przypadku zgromadzonych na podstawie literatury omawiającej konkretne zastosowania oceniania kształtującego w szkolnictwie pokazuje dowodnie długość i intensywność obecności oceniania kształtującego w pedagogice, a w tym także w pedagogice językowej (Hattie, 2008; Heritage, 2013).

Jakkolwiek zatem oceniania kształtującego nie sposób traktować jako innowacji, to jednak nadal widać poważne trudności w promowaniu tej formy oceniania wśród nauczycieli, zapewne z uwagi na jego czasochłonność i subiektywizm. Jednak wspomniana wyżej metaanaliza przypadków wykazuje wyraźnie, że stosowanie oceniania kształtującego podnosi wyniki nauki i zwiększa zaangażowanie uczniów, jego rola jest więc nie do przecenienia. Pomimo zmitologizowania innowacyjności zjawiska jako forma ewaluacji jest więc ono ze wszech miar godne dalszego upowszechniania.

### *Mit VII – We współczesnym ocenianiu kryterium native-speaker nie jest już punktem odniesienia*

Dziś, znając liczne geograficzne, regionalne i społeczne warianty języka obcego, staramy się odejść, lub wręcz uciec, od myślenia o poziomie umiejętności językowych tego, kto jest rodzimym użytkownikiem języka obcego (*native-speaker*) i preferujemy pojęcie biegłego czy sprawnego użytkownika (*proficient speaker, proficient user*). Jednak, konstruując jego pojęcie, pozostajemy i tak w zależności od modelu rodzimego użytkownika, jakkolwiek wykształconego lub wyidealizowanego, tyle że nie jest to już rodzimy użytkownik jednej odmiany języka uznanej za prestiżową. Innej, dogodnej miary biegłości nie mamy. Spór o te pojęcia toczył się w dużej mierze dlatego, że część badaczy nie przyjmowała do wiadomości idealnego konstruktów *native-speakera* i jego kompetencji, a tylko jego konkretne regionalne upostaciowania i równie konkretne wypowiedzi (Kramersch, 2003). Znaczenie punktu odniesienia, jakim jest model rodzimego użytkownika i brak sensownych alternatyw, przyznają nawet wybitni specjaliści z zakresu testowania (Davies, 2015), jakkolwiek przyznają też, że model sprawnego użytkownika przynosi pewne zmiany. Są to głównie modyfikacje podsystemowych kryteriów oceny odnoszące się przede wszystkim do trzech aspektów: wymowy, poziomu poprawności gramatycznej i idiomatyki,

o czym dowodnie świadczą analizy języka angielskiego jako *lingua franca* (Jenkins, 2007; Seidlhofer, 2002). W nauczaniu różnice odnoszą się także do roli kompetencji kulturowej, która ustępuje miejsca kompetencji interkulturowej, obie jednak prawie nigdy nie stają się przedmiotem ewaluacji – vide trudność nr 1 omawiana w dalszej części niniejszego tekstu.

Modyfikacje kryteriów oceny mają wielką wartość z punktu widzenia celów nauczania, jakim jest skuteczna interakcja i komunikacja, w tym komunikacja za pomocą języka obcego w jego funkcji *lingua franca*. Nie oznacza to jednak, że realne jest pominięcie biegłości idealnego native-speakera jako punktu odniesienia wysokiej sprawności językowej, choć nie jest już on modelem do bezkrytycznego odwzorowywania.

Kwestia ta nie musi, oczywiście, odnosić się do wartości nauczyciela, który jest native-speakerem danego języka. Tu decyduje nie tylko biegła znajomość języka, ale i przygotowanie pedagogiczne. Obie grupy nauczycieli charakteryzują się określonymi, choć odmiennymi zaletami z punktu widzenia za-trudniających ich instytucji, a także z punktu widzenia samych uczących się (Medgyes, 1994/2017; Paradowski, 2011).

### 3. Trudności i kontrowersje w sferze ewaluacji

Oprócz mitów, spośród których omówiłam powyżej jedynie kilka najpowszechniejszych, widać w glottodydaktyce także pewne wyraźne trudności pojawiające się w obszarze oceniania. Znajdują one różne, nie zawsze trafne, ale często przeciwstawne rozwiązania, toteż mogą stanowić pożywkę dla powstawania kolejnych mitów. Z tej racji warte są – jak się wydaje – rozważenia na równi z siedmioma mitami omówionymi powyżej. Zajmę się tu trzema z nich.

#### Trudność 1 – Ocenianie sumatywne niektórych kompetencji językowych

Istnieją kompetencje, które z trudem poddają się lub wręcz wcale nie poddają się całościowej ocenie. Stałą i nieprzemijającą trudnością jest na przykład w obszarze kontroli i oceny sprawa oceniania kompetencji interkulturowej. Kompetencja ta jest od przynajmniej dwudziestu lat stale wymieniana jako jeden z istotnych celów nauczania w dokumentach programowych wielu krajów, jednak do dziś jest celem, którego realizacja konsekwentnie nie jest sprawdzana. I jakkolwiek sugeruje się liczne techniki dydaktyczne i samokształceniowe, a w szczególności metody etnograficzne i projekty uczniowskie, to jednak oprócz wiedzy kulturoznawczej, mierzonej testami o faktograficznym nachyleniu, do dziś trudno o narzędzia ewaluujące kompetencję interkulturową uczniów. Próby tego rodzaju są o wiele częstsze w szkolnictwie wyższym (Griffith i in., 2016). Dążenie

do zachowania poprawności metodologicznej wymagałoby nacisku na to, by każdy cel programowy był starannie oceniany w końcowych etapach kursu językowego. Im silniejsze są jednak płynące z różnych stron ideologiczne naciski na systemy edukacyjne w Europie i poza nią, tym trudniej nalegać, by w tym przypadku tak się właśnie stało. Kontrowersyjne jest bowiem samo wprowadzanie tej kompetencji do procesu ewaluacji dziś, gdy postawy wobec odmienności i wybory poszczególnych wartości tak silnie się polaryzują.

Istnieją także inne kompetencje, które – choć z odmiennych względów – sprawiają podobne trudności (Lenz i Berthele, 2010). Niektórzy badacze zaliczają do nich na przykład mediację, tak zwany dialog konwergentny czy inaczej wielojęzyczny, jak również interkomprehensję, twierdząc, że poddają się one nieźle ocenianiu kształtującemu, ale już nie całościowemu (Downing i Vela Delfa, 2011).

## Trudność 2 – Ocenianie sprawności interakcyjnych

Problemem trudnym do pominięcia, ale i do rozwiązania, jest ocena sprawności interakcyjnych. Dzieje się tak dlatego, że sprawność mówienia jednego z interlokutorów i jego osobowość wpływają na poziom rozumienia drugiego z nich, a zatem też i na kolejną wypowiedź rozmówcy. Ponadto różne są poziomy umiejętności poszczególnych osób w odniesieniu do umiejętności rozszyfrowywania intencji rozmówców i odczytywania nadawanych przez nich sygnałów niewerbalnych, które mogłyby znacznie ułatwić rozumienie (Szmerdt, 2002). Jak wskazywał Thornbury w dyskusjach eksperckich Cambridge University Press (*personal communication*) – często nie rozumiemy przecież nie tyle słyszanej wypowiedzi, ile danego rozmówcy. Oceny nie ułatwia przy tym subiektywizm oceniającego i typ szkolenia, jakiemu podlegał, przygotowując się do roli egzaminatora. Próby rozwiązania tych problemów poprzez ujednoczenie cech interlokutora były prowadzone już w latach 80. w Stanach Zjednoczonych. W miejsce *Oral Proficiency Interview – OPI* wprowadzono wtedy na rynek tak zwany test pośredni (*Simulated Oral Proficiency Interview – SOPi*), w którym w miejsce egzaminatora- interlokutora wprowadzono kasetę z nagraniem jego wypowiedzi – pytań i poleceń, na której to kasecie, w przerwach o długości zróżnicowanej w zależności od typu zadania, egzaminowany nagrywał własne wypowiedzi (Stansfield, 1996). Choć test w tej postaci okazał się porównywalny co do rzetelności ze zwykłym wywiadem typu OPI, to jednak wywoływał większy stres i charakteryzował się niższą trafnością fasadową (ang. *face validity*) niż jego tradycyjne odpowiedniki OPI (Kenyon i Tschirner, 2000). Ujednoczenie interlokutora tym sposobem nie okazało się więc dobrym rozwiązaniem, a sam czynnik interlokutora nadal nie poddaje się rzetelnej analizie. Kontrowersjom próbuje się zapobiec poprzez coraz staranniejszy dobór i coraz precyzyjniejsze uszczegółowienie kryteriów oceny – problem jednak pozostaje.

### Trudność 3 – Zgodność technik testowania z procesami poznawczymi młodszych uczniów

Testowanie receptywnych umiejętności językowych odbywa się zawsze – siłą rzeczy – na pewnych treściach, a te mają swój ładunek poznawczy, który znajduje swe odzwierciedlenie w doborze wyrazów, struktur składniowych, długości poszczególnych zdań, ale i w rozmiarach całego tekstu. Ładunek ten powinien być zgodny z poziomem procesów poznawczych osób uczących się w danej grupie wieku. I tu nowe tendencje w kształceniu językowym i nowe sytuacje społeczne piętrzą kolejne trudności. Widać to na przykład w grupach uczniów młodszych.

I tak na przykład tendencje w europejskiej polityce językowej, które promują wczesny start, przynoszą w efekcie rozwiązania organizacyjne, dzięki którym nauka rozpoczyna się już w przedszkolu. Mobilność społeczna umożliwia wielu dzieciom intensywną ekspozycję na język w trakcie licznych wyjazdów, a to powoduje u części uczących się bardzo znaczny wzrost umiejętności językowych. Wyższe poziomy egzaminów językowych adresowane są jednak do uczniów starszych, zatem ich intelektualny poziom utrudnia lub wręcz uniemożliwia przystąpienie do nich uczniów młodszych.

Trudności pojawiają się także w egzaminach ogólnie dobrze dostosowanych koncepcyjnie do poziomu rozwoju poznawczego uczniów (Herman i Ripley, 2012; Lewkowicz, 2015; Shohamy, 2011). Problem widoczny jest wtedy, gdy ładunek poznawczy zadań okazuje się zawyżony bynajmniej nie wskutek złej konstrukcji zadań testowych dla uczniów młodszych, ale z uwagi na wadliwe relacje pomiędzy tekstem a materiałem ilustracyjnym stanowiącym jego element. Lekceważenie procesów poznawczych dziecka poprzez nietrafny dobór materiału ilustracyjnego widać na przykład wtedy, gdy rozumienie przyimka „za” sprawdza się rysunkiem, na którym łódkę umieszczono obok głowy plażującego dziecka licząc, że oczywista dla dorosłego symbolika przedstawiania trójwymiarowej przestrzeni na dwuwymiarowym rysunku okaże się dla dziecka równie zrozumiała.

Z jeszcze innymi trudnościami spotykamy się w nawet najlepszych egzaminach międzynarodowych przeznaczonych dla niższych grup wieku, gdy mamy w nich do czynienia z testowaniem sprawności słuchania poprzez sprawność czytania. Interpretacja wyników wskazuje wtedy na niedomagania nie w tej sprawności, w której one istotnie występują. Nie wszystko jednak – z praktycznych względów – poddaje się testowaniu z pominięciem tekstu pisanego.

Trudność 3 wydaje się stosunkowo najłatwiejsza do usunięcia lub przynajmniej ograniczenia, wiąże się bowiem – w większym stopniu niż poprzednie – ze starannością procesu konstrukcji narzędzi oceniania i współpracą poszczególnych autorów.

#### 4. Przyczyny i skutki mitologizacji. Wnioski

W grupie spraw, którą określam tu jako mity, widać dwie charakterystyczne tendencje. Pierwsza to pewnego rodzaju gra z przeszłością, by wzmocnić wiarę w postęp i utwierdzić się w przekonaniu, że linia rozwojowa stale prowadzi od gorszych do lepszych rozwiązań. Druga – to gra z teraźniejszością, byśmy sami sobie przedstawiali ją w możliwie najlepszym świetle. Obie te tendencje realizują się poprzez szereg świadomych lub podświadomych zabiegów. Są to:

- deprecjonowanie przeszłości, np. starożytności i średniowiecza, o czym świadczy mit I;
- stosowanie podwójnych standardów w ocenie metod, np. metody grammatyczno-tłumaczeniowej, jak to widać w micie II;
- idealizowanie przeszłości, by zarysować linię postępu, jak to widać w przesunięciach ocen metody audiolingwalnej (mit III);
- swego rodzaju „lakierowanie” teraźniejszości i jej „lifting mentalny”, co czynią mity IV i V w kwestii potrzeb uczącego się i oceniania kształtującego;
- ukrywanie przewagi kryterium praktyczności nad kryterium autentyczności, jak widać w micie V, ale też
- upraszczanie, co charakteryzuje mit VII dotyczący kryterium rodzimego użytkownika.

Wszystko to ma ukazać, iż było źle, a jest dobrze, choć o postępie – acz nielinearnym – świadczy dostatecznie wiele pozytywów, by glottodydaktycy musieli uciekać się do zniekształcania faktów. Wśród oczywistych sukcesów możemy wymienić choćby następujące: sformalizowano i skodyfikowano testowanie, wprowadzono jasne kryteria oceny i klucze odpowiedzi, zmieniono typ tekstów, na których pracują zdający, wprowadzono interakcyjność w tok egzaminu ustnego, zróżnicowano badane umiejętności i ich poziomy oraz zbadano efektywność kształcenia ocenianego (Hattie, 2008; Heritage, 2013).

Znaczna część owego mitologizowania, a w konsekwencji i zniekształcania faktów, bierze się również – jak się wydaje – z powierzchowności w formułowaniu opinii metodycznych. Po części powierzchowność ta wynika z niezbyt starannego zapoznania się z historią dydaktyki języków obcych; to, na szczęście, wada uleczalna, o ile starczy chęci i pracowitości. Po części jednak – i to trudniejszy problem – powierzchowność ta jest zamierzona, stanowiąc odbicie coraz silniejszych związków dydaktyki z rynkiem. Uderza dziś bowiem liczba wydawnictw, wielość tytułów podręczników, materiałów uzupełniających, programów komputerowych, funkcjonujących jako towar, którego producenci poszukują kolejnego klienta. Ogromna jest też liczba szkół językowych, z których każda chciałaby promować

się, deklarując ofertę kursów prowadzonych własną „metodą”. Wszystko to owocuje potrzebą reklamy, a ta jest skuteczniejsza wtedy, kiedy ukazuje dany produkt czy zjawisko jako innowacyjne, a zatem odmienne od wszystkiego, co dotąd znane. Nie dziwi, wobec tego, że, aby kontrast był bardziej przekonujący, wszelkie wcześniejsze rozwiązania ukazywane są jako przestarzałe i niewiele warte. By tak je zaprezentować, przemilcza się zalety, a cechy związane z właściwą im funkcją i kontekstem – przedstawia się jako już od zarania pozbawione sensu. Tych błędów pozwoliłoby uniknąć stosowanie kryteriów etycznych w promowaniu oferty edukacyjnej, a skoro w świecie pieniądza trudno na to liczyć, jedyną nadzieją może być zdolność adresatów reklam do krytycznego osądu i – co za tym idzie – eliminowania nieuzasadnionych twierdzeń.

Nieco inny rodzaj problemu kryje się w grupie kwestii omówionych wyżej jako trudności i kontrowersje. Otóż nie wszystkie bariery w procesie oceniania da się usunąć, choć wciąż czynione są starania zmierzające do ich ograniczenia. Spory dotyczące wartości, by przypomnieć prawdę oczywistą, choć wciąż aktualną, są jednak na stałe wpisane w losy ludzkości, a nie tylko w jej mikroskopijny fragment, jakim jest ocenianie w nauce języków obcych. Można zatem mieć tylko nadzieję, że częściej niż dziś będą one prowadzić do dialogu, a nie do konfliktu. Warto pamiętać, że każda zmiana kontekstu niesie nowe problemy, a to, pomimo pojawiających się trudności, motywuje do nowych działań i tworzy nowe szanse.

## BIBLIOGRAFIA

- Allal L., Lopez L.M. (2015), *Formative assessment of learning. A review of publications in French* (w) OECD, *Formative assessment. Improving learning in secondary classrooms*. Paris: OECD, str. 241-264.
- Alismail H.A., McGuire P. (2015), *21st century standards and curriculum. Current research and practice* (w) „Journal of Education and Practice”, nr 6, str. 150-154.
- Bachman L. (2015), *Language assessment. Its development and future* (w) Kunnan A.J. (red.), *Talking about language assessment*. London: Routledge, str. 199-216.
- Bachman L., Palmer A. (2010), *Language assessment in practice. Developing language tests and justifying their use in the real world*. Oxford: Oxford University Press.
- Barthes R. (1957/ 2013), *Mythologies*. New York: Hill and Wang.
- Black P., William D. (1998), *Assessment and classroom learning* (w) „Assessment in Education”, nr 5, str. 7-71.
- Bloom B.S., Englehart M.D., Furst H. J., Hill W.H., Krathwohl D.R. (1956), *Taxonomy of educational objectives. The classification of educational goals. Handbook I. Cognitive domain*. New York: David MacKay.

- Bloom B.S., Englehart M.D., Furst H. J., Hill W.H., Krathwohl D.R. (1964), *Taxonomy of educational objectives. The classification of educational goals. Handbook II. Affective domain*. New York: David MacKay.
- Bloom B.S., Hasting J.T., Madaus G.F. (1971), *Handbook of formative and summative evaluation of student learning*. New York: McGrawHill Book Co.
- Campbell J. (2013), *Potęga mitu*. Kraków: Znak.
- Cardinet J. (1977), *Objectifs éducatifs et évaluation individualisée*. Neuchâtel: Institut Roman Recherches et de Documentation Pédagogique.
- Czetwertyńska G. (2015), *Ocenianie jako nauczanie i uczenie się* (w) „Języki Obce w Szkole”, nr 3, str. 4-9.
- Davies A. (2015), *40 Years in Applied Linguistics* (w) Kunnan A.J. (red), Talking about language assessment. London: Routledge, str. 35-50.
- De Ketele J.M. (red), (1986). *L'évaluation: approche descriptive ou prescriptive?* Brussels: De Boeck.
- Downing R.H., Vela Delfa C. (2011), *Intercomprehension between neighboring languages. From language learning to teacher training* (w) Bataller S.M., Clavel-Arroitia B. (red.), Multiple voices in academic and professional discourse. Current issues in specialized language research, teaching and new technologies. Newcastle on Thyne: Cambridge Scholars Publishing, str. 219-229.
- Earl L. M. (2013), *Assessment as learning. Using classroom assessment to maximize learning*. London: Corvin Press.
- Esterhazy P. (2008), *Harmonia caelestis*. Warszawa: Czytelnik.
- Goddu A. (2010), *Copernicus and the Aristotelian tradition*. Leiden-Boston: Brill.
- Gordon E.W., Rajagopalan K. (2016), *The testing and learning revolution. The future of assessment in education*. New York: Palgrave Macmillan.
- Griffith R.L., Wolfeld L., Armon B.K., Rios J., Liu O.L. (2016), *Assessing intercultural competence in higher education. Existing research and future directions* (w) Research Report ETS RR -16-25, <https://onlinelibrary.wiley.com/doi/epdf/10.1002/ets2.12112> [DW 15.09.2018].
- Hattie J. (2008), *Visible learning. A synthesis of over 800 meta-analysis relating to achievement*. London: Routledge.
- Heritage M. (2013), *Formative assessment in practice. A process of inquiry and action*. Cambridge: Harvard Education Press.
- Herrmann J., Höfer C. (1999), *Evaluation in der Schule – Unterrichtsevaluation*. Gütersloh: Bertelsmann Stiftung.
- Herman J., Ripley M. (2012), *Assessment and teaching of the 21st century*. New York: Springer Publishers.
- Howatt A., Widdowson H. (2004), *The history of ELT*. Oxford: Oxford University Press.
- Jenkins A. (2007), *English as a lingua franca. Attitude and identity*. Oxford: Oxford University Press.



- Kenyon D., Tschirner E. (2000), *The rating of direct and semi-direct oral proficiency interviews. Comparing performance at lower proficiency levels* (w) „The Modern Language Journal”, nr 84, str. 85-101.
- Köller O. (2015), *Formative assessment in classrooms. A review of the empirical German literature* (w) *Formative assessment. Improving learning in secondary classrooms*. Paris: OECD, str. 265-279.
- Kramsch C. (2003), *The privilege of the non-native speaker* (w) Blyth C. (red.), *The sociolinguistics of foreign language classrooms. Contributions of the native, the near-native, and the non-native speaker*. Boston, MA: Heinle, str. 251-62.
- Kucharczyk R. (2015), *Jak definiować poprawność językową w epoce wielojęzyczności? Kryteria oceny wypowiedzi ustnej na egzaminie z języka obcego nowożytnego (na przykładzie egzaminów maturalnych)* (w) Sujecka-Zajac J. (red.), *Ewaluacja biegłości językowej*. Warszawa: Rada Koordynacyjna ds. Certyfikacji Biegłości Językowej UW, str. 119-132.
- Lado R. (1961), *Language testing. The construction and use of foreign language tests*. London: Longman.
- Lenz P., Berthele R. (2010), *Assessment in plurilingual and intercultural education*. Strasbourg: Council of Europe.
- Lewkowicz A. J. (2015), *Developing appropriate monolingual assessment criteria for tests in many languages* (w) Sujecka-Zajac J. (red.), *Ewaluacja biegłości językowej*. Warszawa: Rada Koordynacyjna ds. Certyfikacji Biegłości Językowej UW, str. 49-57.
- Medgyes P. (1994/2017), *The Non-native Teacher*. Exeter: Swan Communication.
- Nunan D. (2012), *Learner-centered English language education*. New York: Routledge.
- Ornstein A.C., Hunkins F. P. (1988), *Curriculum. Foundations, Principles and issues*. Boston: Allyn and Bacon.
- Paradowski M.B. (2011), *Multilingualism – Assessing Benefits* (Komorowska H.(red.), Warszawa: FRSE, str. 335-354.
- Pinar W.F. (2004), *Understanding curriculum*. New York: Peter Lang.
- Rosenthal R., Jacobsen L. (1968), *Pygmalion in the classroom. Teacher expectation and pupils' intellectual development*. New York: Holt, Rinehart and Winston.
- Rubie-Davies Ch. (2009), *Teacher expectations and labeling* (w) Saha J, Dworkin A.G (red.), *International Handbook of Research on Teachers and Teaching*, str. 695-707.
- Scriven M. (1967), *The methodology of evaluation* (w) „AERA Monograph Series on Evaluation”, nr 1, str. 39-83.
- Seidlhofer B. (2002), *A Concept of international English and related issues: From “Real English” to “Realistic English”*. Strasbourg: Council of Europe.

- Shohamy E. (2015), *From cloze to consequences and beyond. Interview by Anne Lazaraton* (w) Kunnan A.J. (red.), *Talking about language assessment*. Routledge, London, str.150 -183.
- Shohamy E. (2011), *The Power of tests. A critical perspective on the uses of language tests*. Singapore: Longman.
- Stansfield C. W. (1996), *Test development handbook. Simulated Oral Proficiency Interview*. Washington, DC: Center for Applied Linguistics.
- Stegmuller F. (1949), *Maupassant. A Lion in the path*. London: Macmillan.
- Sterna D. (2016), *Uczyć się uczyć. Ocenianie kształtujące w praktyce*. Warszawa: Centrum Edukacji Obywatelskiej.
- SurveyLang*. (2012). Brussels: European Union.
- Szmerdt D. (2002). *Ocenianie sprawności mówienia* (w) *Ewaluacja w nauce języka obcego* (red. Komorowska H). Białystok: Wydawnictwo UWB, s. 37-52.
- Taba H. (1963), *Teaching strategies and cognitive functioning in elementary school children*. Cooperative research project 2404. San Francisco: San Francisco.
- Taylor P.H., Richards M. (1985), *An introduction to curriculum studies*. Milton Keynes: NFER Nelson.
- Tedesco J.C., Opertii R., Amadio M. (2014), *The curriculum debate. Why it is important today* (w) „Prospects”, nr 44, str. 527-546.
- Tsagari D., Banerjee J. (2016), *Handbook of second language assessment*. Boston: De Gruyter Mouton.
- Tyler R. (1949), *Basic principles of curriculum and instruction*. Chicago: The University of Chicago Press.
- Weiss J. (red). (1991), *Évaluation – problème de communication*. Cousset: Del Vals.