Sabina A. Nowak
Uniwersytet Jagielloński

Nowa jakość oceniania: ocenianie kształtujące

Najtrudniejszym testem językowym, jaki przyszło mi kiedykolwiek zdawać, była forma rozmowy o pracę polegająca na wykonaniu 2 zadań. Nie wiedziałam, co napisać w liście formalnym i jak zapytać o ofertę handlową.

Studentka UJ

W szkole wyższej coraz częściej słyszy się o wykorzystaniu wiedzy teoretycznej w praktyce oraz zdobywaniu doświadczenia w trakcie studiów. Pragmatyczny aspekt kształcenia od wielu lat akcentowany jest głównie poprzez refleksyjne podejście Deweya (1988) w procesie uczenia się przez działanie. Mimo iż koniecznością stało się przygotowanie językowe studentów na potrzeby zawodowe (ang. *English for occupational purposes,* EOP), proces kształcenia na lektoratach skupia się w dalszym ciągu na *stricte* językowych, nie zaś ogólnolingwistycznych aspektach oceniania. Oceny nauczycieli oparte na analizie błędów gramatycznych, leksykalnych czy interpunkcyjnych są wartościową informacją dla studentów, jednak nie odnoszą się do szeroko rozumianych kompetencji komunikacyjnych, o których coraz częściej mówi się w edukacji. Nie można więc założyć pełnego kształtowania **Zintegrowanych Kompetencji Osobistych (ZKO), w których skład wchodzą kompetencje językowe, przedmiotowe, kluczowe i ogólne** (**Majewska, 2008).**

Najważniejszym aspektem przemawiającym za odejściem od tradycyjnego podejścia do oceniania jest to, że testy językowe nie do końca „mierzą to, co mają mierzyć" (określenie trafności testu) w odniesieniu do przyszłego kontekstu zawodowego uczących się[1]. Dlatego warto zadać sobie pytanie, czy w nauczaniu poprzez CLIL i w dobie podejścia opartego na działaniu (ESOKJ, 2003), które opiera się na wykonywaniu autentycznych zadań (ang. *real-life task*) ważne jest jedynie sprawdzanie błędów poprzez określenie, co student wie, czy też **jak potrafi wykorzystać swoje umiejętności**.

W literaturze przedmiotu znaleźć można wiele zagadnień związanych z zapewnieniem jakości oceny (Komorowska, 2010). Brak natomiast odniesień do pomiaru zintegrowanych umiejętności językowo-przedmiotowych (ang. *content and language integrated learning,* CLIL) w odniesieniu do kształcenia językowego do celów zawodowych (ang. *English for specific purposes)*. Wprawdzie porfolio polegające na zbieraniu dowodów świadczących o umiejętnościach uczących się może być im pomocne, jednak takie narzędzie samooceny nie dostarcza informacji zwrotnej na temat stopnia realizacji i jakości założonych efektów kształcenia.

[1] Niniejsze argumenty oparto na pozycji Parisa i Ayres (1997) *Stawanie się refleksyjnym uczniem i nauczycielem*. Autorzy podają **cechy rzeczywistej oceny**, która powinna logicznie wynikać z tego, co się dzieje w klasie, być świadectwem aktywności uczących się na wielu płaszczyznach, pobudzać studentów do pragmatycznego wykorzystania nabywanej wiedzy oraz stać się odbiciem ich wartości i przekonań (tamże, s. 16).

Jednym z rozwiązań dydaktycznych może być wprowadzenie **oceniania kształtującego (lub alternatywnego, w skrócie OA**[2]) opartego na ocenie performansów, będących czynnościami podejmowanymi w ramach kształcenia (ang. *performance–based assessment*). Najczęściej stosowanymi tutaj narzędziami w pracy ze studentami są **dzienniki kształcenia** (Brown, 2004, s. 256–278), które bazują na samoocenie uczących się. Ich główną zaletą jest to, że pozwalają dostrzeć własne mocne i słabe strony, gdyż dostarczają różnego rodzaju informacji zwrotnej (ang. *transparency*). W przeciwieństwie do portfolio samoocena oparta jest nie tyle na samookreślaniu się na skali umiejętności, ile na efektywnym wykonywaniu zadań i dokonywaniu metarefleksji (refleksji nad refleksją).

Tworzenie rubryk wchodzących w skład dzienników kształcenia

Ocenianie trudno mierzalnych, gdyż subiektywnie określanych osiągnięć (performansów), wydaje się rzeczą nie tyle trudną, co niemożliwą dla lektorów niebędących specjalistami z przedmiotu kierunkowego studentów. Rozwiązaniem wydaje się tutaj zastosowanie **rubryk**[3], które są wartościowym narzędziem pomagającym w kształtowaniu kompetencji i rozwoju osobistym uczących się. Rubryki[4] opisują wykonanie danego działania, a w kontekście dokonywania pomiaru edukacyjnego są tym, co może wchodzić w skład oceny poszczególnych kryteriów umiejętności. Jednak w przeciwieństwie do kryteriów oceny, których główną cechą jest ocenianie, rubryki mają funkcję opisową, nie zaś ewaluacyjną. Najważniejsze jest, aby sprecyzować „aspekty, które określają, jakie są oczekiwania względem uczących się, którzy mają przystąpić do pisania" (Bachman, 1990, s. 118).

Można je zatem określić jako narzędzia służące uczącym się do samooceny, zawierające szczegółowy opis tego, co stanowi przedmiot zainteresowania i potrzeb przyszłych pracodawców. Obecnie w edukacji rybryki definiuje się jako „narzędzie pomiaru, które określa szczegółowe wymagania w odniesieniu do zadania" (Stevens i Levi, 2005, s. 3). To, co wyróżnia rubryki od oceny (w różnej skali: 1–6, ndst–cel, 0–100%) jest ich deskryptywny charakter składający się z czterech głównych elementów (tamże, s. 5–6). Parametrami włączonymi

[2] Celem **OA** (Komorowska, 2002) jest dostarczanie ciągłej informacji zwrotnej opartej na zadaniach różnego rodzaju.

[3] Potrzeba tworzenia rubryk związana była ze zmianą podejścia od oceniania i nauczania do samooceny zorientowanej na uczenie się. Jak twierdzą Stevens i Levi (2005, s. 14), pierwszym etapem tworzenia rubryk powinna być pogłębiona refleksja nauczyciela, a więc uświadomienie sobie i zwerbalizowanie głównych założeń oraz celów związanych z kształceniem (założonymi efektami kształcenia) oraz pomiarem edukacyjnym. Ważnym aspektem ewaluacji rubryk jest określenie ich trafności (*validity*), rzetelności (*reliability*) i użyteczności (*utility*) (Quinlan, 2012, s. 158). Poprzez trafność mamy na uwadze ujęcie w nich celów edukacyjnych oraz odzwierciedlenie treści przedmiotowych realizowanych w procesie edukacji. Rzetelność odnosi się do niezmienności wyników w różnych okresach czasu, porównywalnych wyników oceny tego samego oceniającego (*intra–rater reliability*), a także zgodności ocen między różnymi osobami (tu: rówieśnikami) (*inter–rater reliability*).

[4] Słowo rubryka/i pochodzi z łacińskiego słowa *rubrica/ruber* oznaczający słowo *czerwony*. Mieszkańcy starożytnego Rzymu słowem „rubryka" określali prawo, gdyż wszelkie ustawy były zapisywane na czerwono (Quinlan, 2012, s. 3). W średniowiecznych rękopisach była to litera, słowo lub zdanie wyróżnione na czerwono w odróżnieniu od tekstu zapisanego czarnym kolorem. W księgach liturgicznych kościoła „rubryką" zaznaczane są na czerwono fragmenty tekstu wskazujące to, co księża mają w danym momencie wykonać.

w objaśnienie jest *opis zadania* (polecenia do wykonania), określona *skala* (poziomy osiągnięć wyrażone w punktach, np. 1–4, lub określeniami np. „Satisfactory/Needs improvement"), *aspekt zadania* (podział na umiejętności/wiedzę) oraz *opis tego, co składa się na każdy poziom osiągnięć* (będący swego rodzaju specyficzną informacją zwrotną).

W polskiej literaturze przedmiotu trudno znaleźć wyczerpujące informacje na temat sposobów tworzenia rubryk i rezultatów ich użycia. Główny problem polega na ich odniesieniu do różnych rodzajów zadań, które nie zawsze można użyć do wszystkich kontekstów edukacyjnych. Innym przykładem specyfiki ich użycia jest trudność w operacjonizacji wskaźników oceny. Jako przykład warto podać za ESOKJ (2003, s. 168):

> *Uwaga na temat rubryki „Inne kategorie": w przykładowych skalach zamieszczonych w Systemie Opisu wskaźniki stopnia powodzenia w realizacji zadania […].
>
> Próba sformułowania, skalibrowania i umieszczenia w tych skalach wskaźników biegłości dla kategorii *Potrzeby wsparcia ze strony rozmówcy* nie powiodła się.

Jak wynika z powyższego zapisu, ustalenie pełnego zakresu dobrze sformułowanych i trafnie dobranych wskaźników do każdego działania edukacyjnego jest przedsięwzięciem czasochłonnym, żmudnym, a czasami niewykonalnym[5]. Jednak z pewnością trud włożony w praktyczne stosowanie (w tym upraszczanie, łączenie i dobieranie) różnych kategorii opisu teoretycznego oceny, jak również ich rozbudowywanie i dostosowywanie do istniejącego kontekstu edukacyjnego (por. ESOKJ, 2003, s. 169) jest działaniem nie tyle potrzebnym, ile koniecznym z dwóch powodów. Po pierwsze, ze strony nauczających, którzy będą stosować je w trakcie pomiaru oraz podczas dostarczania uczącym się informacji o tym, co złożyć się ma na ich poszczególne osiągnięcia. Po drugie, ze strony uczących się, którzy dzięki dobrze określonym rubrykom będą w stanie bardziej świadomie dokonać samooceny i przydatności zdobywanej wiedzy i umiejętności, przez co skuteczniej będą **kierować** własnym uczeniem się.

W badaniu[6] założono udział N = 64 studentów zarządzania w turystyce w specjalistycznym dyskursie i dialogu edukacyjnym dokonujące się za pomocą pisemnych refleksji. Studenci uczęszczali na lektorat z języka angielskiego w biznesie na poziomie B2+ przez okres jednego roku akademickiego (2013/2014). Istotnym elementem analizy danych stała się weryfikacja merytoryczna wypowiedzi w formie pisemnych refleksji oparta na wykonywaniu różnego rodzaju zadań w ramach **refleksyjnych dzienników kształcenia**

[5] Należy zaznaczyć, że choć w Systemie Opisu (2003, s. 190) zrezygnowano z wykorzystania wskaźników odnoszących się do umiejętności związanych z pracą i kompetencjami socjokulturowymi (tamże, s. 189) jednak waga kształcenia tych umiejętności ma ogromne znaczenie w przygotowywaniu studentów do przyszłej pracy zawodowej. Autorka jest zdania, iż zbyt mało uwagi poświęca się kompetencjom prakseologicznym (skutecznego działania) uczących się, co widoczne jest szczególnie w zadaniach wymagających kreatywnego myślenia i rozwiązywania problemów.

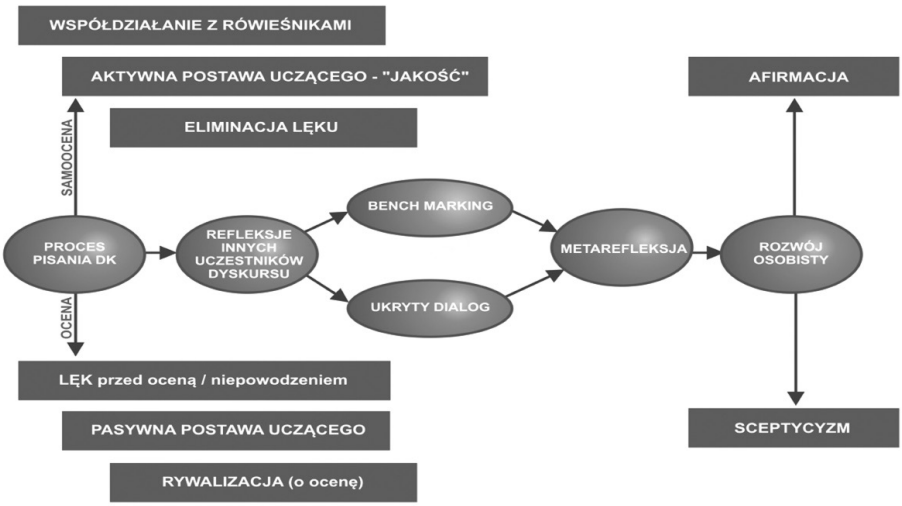
[6] Wyniki badań stanowią część rozprawy doktorskiej autorki pod tytułem *Rola refleksji i samooceny w zintegrowanym kształceniu językowo-przedmiotowym*.

(RDK)[7]. W przeciwieństwie do testów językowych[8] dzienniki prowadzone były na platformie „Pegaz" Uniwersytetu Jagiellońskiego, która zezwalała na dostęp o dogodnej dla studentów porze. Zakładały podjęcie interakcji pomiędzy uczestnikami specyficznego specjalistycznego dyskursu dzięki zamieszczaniu i udostępnianiu wypowiedzi pisemnych na opartej na Moodle uniwersyteckiej platformie (ang. *computer assisted language learning*, CALL). Refleksje nie były oceniane, gdyż ich użyteczność opierała się na innych przesłankach związanych z samooceną opartą na **metarefleksji** (refleksji nad refleksją). Celem było uczenie się przez działanie (ang. *action-oriented learning*), a więc imitowanie określonych sytuacji, w których studenci mogli się kiedyś znaleźć. Proces kształcenia dokonywał się poprzez obserwowanie i porównywanie performansów (ang. *performance*), czyli poprzez wyłonienie najlepszych przykładów efektywnego i skutecznego działania. RDK wykształcały wiele umiejętności poznawczych, afektywnych i prakseologicznych.

Jako jednostki testowe refleksje były krótkie (około 100 słów), jednak ze względu na różnorodne rodzaje zadań ich napisanie wymagało dużego wysiłku ze strony uczących się (rekonstruowanie wiedzy, nadawanie znaczeń, zmiana perspektywy). Dzięki zapoznawaniu się z wypowiedziami innych studentów możliwa była transformacja wiedzy i przekonań, gdyż uczący się zapoznawali się z punktami widzenia innych **uczestników wspólnoty dyskursywnej**. Lęk towarzyszący uczącym się na początku procesu pisania refleksji obniżał się w miarę nabywania wprawy i pewności siebie w ich dokonywaniu. Dzięki prowadzeniu RDK studenci mieli możliwość kształtowania własnej wiedzy metapoznawczej (Anderson i Krathwohl, 2001) poprzez analizę **wiedzy o zadaniu, użytych strategii i określeniu samowiedzy**. Zaobserwowane zależności pomiędzy tradycyjną oceną nauczyciela a samooceną opartą na wykonywaniu zadań i dokonywaniu samooceny metapoznawczej zostały zobrazowane na rysunku 1.

[7] Dzienniki kształcenia mogą przybierać różną formę w zależności od kontekstu CLIL, wspólnoty dyskursywnej oraz celu, jaki ma przyświecać ich prowadzeniu.

[8] Testy językowe nie mają wartości poznawczej (Paris i Ayres, 1997, s. 39), gdyż ocenianie polega na zastosowaniu umiejętności myślenia niższego rzędu (odtworzenia w pamięci, zrozumienia i zastosowania). Nie bierze się pod uwagę analizy, ewaluacji i twórczości, w tym czynników indywidualnych i kontekstowych. Rezultaty ogłaszane są w celu poinformowania o wyniku, nie zaś po to, aby narzędzie, jakim jest test, spełniało funkcję kształtującą, na podstawie której uczący się mogliby wykorzystać je w dalszej pracy. Użyteczność egzaminu nie jest wysoka z uwagi na to, że nacisk kładzie się na formę wypowiedzi (poprawność leksykalno-gramatyczną), nie zaś na treści (umiejętność krytycznego myślenia czy rozwiązywania problemów). Wspomniani autorzy wymieniają siedem negatywnych konsekwencji wprowadzania testów i egzaminów, które mogą mieć wpływ na uczących się:1. studenci nie mają wpływu na rodzaj testu; 2. zdecydowana większość studentów osiąga średnie wyniki; 3. egzamin nie wykształca umiejętności poszukiwania i porządkowania informacji; 4. egzamin opiera się na indywidualnej pracy, nie ma więc mowy o współpracy i kooperacji pomiędzy uczącymi się; 5. jednostki testowe są zazwyczaj krótkie, wyrwane z kontekstu, odtwórcze i zachęcają do udzielania odpowiedzi zamiast formułowania (konstruowania wiedzy) oraz nadawania znaczeń (twórczość); 6. egzaminowi zazwyczaj towarzyszy lęk i brak wiary w siebie oraz własne możliwości uczących się; 7. egzamin nie określa także wiedzy metapoznawczej uczących się, która ma ogromne znaczenie w uczeniu się przez całe życie.



Rysunek 1. Proces kształcenia w CLIL oparty na refleksyjnych dziennikach kształcenia (oprac. własne)

W pisemnych refleksjach starano się dostrzec stopień realizacji tematu, a tym samym spełnione lub nie oczekiwania względem założonych efektów kształcenia. Opracowanie rubryk opierało się na kryterium użyteczności związanej z określeniem ich przydatności w przyszłej pracy zawodowej. W opracowywaniu poszczególnych wskaźników oceny pomocne okazało się określenie **cech prymarnych**[9] wypowiedzi uczących się (tab. 1).

Rubryki opracowywane zostały przez lektorkę i były omawiane ze studentami podczas zajęć. Opisy osiągnięć różniły się nieznacznie w poszczególnych grupach, gdyż oparte były na różnych wypowiedziach uczących się. Różnice polegały głównie na określeniu ważności poruszanych zagadnień i sformułowań, a więc efektach, jakie wypowiedzi wywierały na odbiorcach oraz możliwej interpretacji ich odbiorców. Podjęte działania weryfikacji jakości poszczególnych odpowiedzi stały się doskonałą okazją do udzielania informacji zwrotnej. Rubryki tworzono na podstawie założych efektów kształcenia, analizy tekstów i wyłonionych przykładów dobrej praktyki.

Wykorzystanie rubryk ze studentami nie było rzeczą prostą. Po pierwsze, wymagało wiedzy na temat zadania i natury refleksji. Po drugie, skupiało się na określeniu ich użyteczności jako narzędzia służącego samoocenie dla studentów[10]. Po trzecie, miało odnosić się do ich zastosowania w różnych kontekstach edukacyjnych (uniwersalności). Studenci często nie zgadzali się z opinią lektora, że ich odpowiedzi wymagają wprowadzenia pewnych poprawek (podania uzasadnienia, przykładów, wyjaśnienia, korzyści itp.). Nie rozumieli,

[9] **Cechy prymarne** (ESOKJ, 2003, s. 180) to zestaw wypowiedzi świadczących o efektywnym wykonaniu zadania, uszeregowane od najsłabszych do najlepszych odpowiedzi posiadających pożądane cechy.

[10] Rubryki miały bowiem służyć nie tylko nauczycielowi, lecz przede wszystkim studentom. Dodatkowy problem związany był również z tym, że rubryki nie były uczącym się wcześniej znane.

że w dziennikach chodziło o dokonywanie głębokich refleksji, które dzięki skutecznej komunikacji spowodują osiąganie sukcesów zarówno na gruncie zawodowym, jak i osobistym. Dopiero po jakimś czasie dostrzegali, że analiza wyłonionych opisów osiągnięć będących formą deskryptorów[11] obrazujących najlepsze przykłady skutecznej komunikacji była okazją do rzetelnej nauki.

Tabela 1. Przykładowa rubryka oparta na wypowiedziach studentów zarządzania w turystyce Uniwersytetu Jagiellońskiego

*If you were to work as a manager in a museum, how would **you make it more attractive** to visitors?* (Gdybyś miał pracować w muzeum jako menedżer, jak mógłbyś **uatrakcyjnić** to miejsce dla zwiedzających?)				
SKALA	Lack of task fulfillment	Needs some improvement	Not bad!	Excellent!
PRZYKŁADOWE WYPOWIEDZI STUDENTÓW ZARZĄDZANIA W TURYSTYCE	*It's hard to say if I would like to work there, because I think It could be too boring for me, not only work like in this museum but in all. I don't have any impressions which I don;t like in castle, because I can find there everything which is worth to see for visitors. (*Museum-Castle in Łańcut) --------- I would like to describe a museum that really moved me emotionally. It's Schindler's factory. --------- *It's a great place for all.*	*You will find there many attractions for kids, for example there are interactive games or virtual fitting room so you can try on clothes from the past and feel like a princess or like a knight* --------- *I really liked each part of museum except photo gallery. It could have been organized in audiovisual way.*	*I think that this museum should extend exhibitions in the future."* --------- *The only thing that I do not like during sightseeing was too much things in one place. Of course, rooms are divided into categories, but I had the impression, for example, in one display case is a lot of pictures. If I could change something it otherwise sort out some of these things."* ---------	-
OPIS OSIĄGNIĘĆ	Refleksja napisana jest na inny temat. Refleksja zawiera jedynie opis, wyrażenie opinii, osąd	Refleksja zawiera przykłady możliwych atrakcji muzeum bez odniesienia się do innowacyjności własnego pomysłu. Brak podania konkretnych sposobów/atrakcji lub uzasadnienia	Refleksja zawiera sugestie lub rekomendacje bez podania uzasadnienia lub konkretnych przykładów rozwiązań na uatrakcyjnienie muzeum	Kreatywne/ pomysłowe/twórcze pomysły poparte uzasadnieniem, opinią, osądem wraz z zaznaczeniem korzyści (konsekwencji), jakie przyniesie ich wdrożenie

Pomimo iż w badaniu zakładano użycie rubryk do oceny wszystkich wykonywanych czynności językowo-przedmiotowych uczących się, działanie to nie powiodło się tak, jak planowano na początku badania. Głównym powodem była przede wszystkim czasochłonność związana z ich opracowaniem oraz pewne ograniczenia natury merytorycznej. Pierwszą trudnością był brak dostępnej

[11] **Deskryptor** to ogólne stwierdzenie określające zakładane efekty kształcenia na poszczególnych stopniach kształcenia.

literatury przedmiotu na temat tworzenia i stosowania[12] rubryk. Dodatkowym problemem okazał się fakt, że studenci pisali refleksje na różne tematy, z tego powodu nie było możliwości wyróżnienia wszystkich cech kluczowych istotnych przy tworzeniu rubryk ani poprzez sortowanie, ani ocenę porównawczą (ESOKJ, 2003, s. 180). Działanie to nie pozwoliło również badaczce na stworzenie wielowymiarowej skali, która pozwalałaby na utworzenie jednej uniwersalnej rubryki odpowiadającej wszystkim działaniom językowym.

Wnioski

Podsumowując, warto zauważyć, że w użyciu rubryk najistotniejszy był sam proces ich tworzenia (dokonywany na podstawie wyłonionych cech prymarnych) oraz stosowania wraz ze studentami, co dzięki wyłonionym deskryptorom służyło udzielaniu informacji zwrotnej. Do głównych przesłanek stosowania samooceny w tym procesie należy wymienić rozwijanie strategii metapoznawczych, które pomagały w krytycznym myśleniu, aktywnym procesie konstruowania znaczeń oraz interpretacji doświadczenia opartego na przekonaniach uczących się.

Główną intencją było wywołanie stałej dyspozycji efektywnego uczenia się polegającego na świadomym określaniu własnych potrzeb, wyznaczaniu i kontrolowaniu celów oraz dzieleniu się własnymi doświadczeniami z innymi uczącymi się. Autorce chodziło o pomiar dydaktyczny w znaczeniu uzyskania pozytywnych **konsekwencji** wdrażania procesu uczenia się, który przyniesie pożądane efekty kształcenia. Założono, że dzięki osiąganym wynikom uczący się będą mogli lepiej funkcjonować w różnych sferach życia (prywatnej, zawodowej, publicznej i edukacyjnej).

Najważniejsza kwestia dotyczy jednak tego, kto lub co ma decydować otym, czego student ma się nauczyć, co mu się może przydać, jak i kiedy. Metarefleksje nad wykonanymi zadaniami oparte na pogłębionej refleksji były dowodem na przydatność i słuszność wprowadzenia oceniania kształtującego na szczeblu wyższym. Uzasadnione wydaje się zatem wprowadzenie bardziej autentycznych form oceniania osiągnięć opartych na samodzielnym, refleksyjnym myśleniu i dokonywanej samoocenie.

Bibliografia

Anderson L.,W., Krathwohl D.R., *A Taxonomy for Learning, Teaching and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives.* Longman. Pearson Education, 2001.

Dewey J., *Jak myślimy,* Wydawnictwo Naukowe PWN, Warszawa 1988.

ESOKJ *Europejski System Opisu Kształcenia Językowego: uczenie się, nauczanie, ocenianie,* Wydawnictwo Centralnego Ośrodka Doskonalenia Nauczycieli, Warszawa 2003.

[12] Podobne doświadczenie zdobyte podczas pracy w roli egzaminatora matury podstawowej i rozszerzonej okazało się niewystarczającym doświadczeniem w ich używaniu na poziomie uniwersyteckim w kontekście CLIL.

Komorowska H., *Sprawdzanie umiejętności w nauce języka obcego. Kontrola – Ocena – Testowanie,* Fraszka Edukacyjna, Warszawa 2002.

Majewska R., *Zintegrowane kształcenie przedmiotowo-językowe. Dydaktyka zadaniowa a rozwój osobistej kompetencji komunikacyjnej w warunkach półautonomii*. Praca doktorska. Uniwersytet im. Adama Mickiewicza, 2013.

Paris S.G., Ayres L.R., *Stawanie się refleksyjnym uczniem i nauczycielem,* Wydawnictwa Szkolne i Pedagogiczne, Warszawa 2007.

Qinlan A.M., *A complete guide to rubrics: Assessment made easy for teachers of K-college*. Lanham, Md: Rowman & Littlefield Education, 2012.

Sterna D., Strzemieczny J., *Strategie oceniania kształtującego*. *Organizacja procesów edukacyjnych* [*w*:] *Jakość edukacji. Różnorodne perspektywy,* pod red. G. Mazurkiewicza, Wydawnictwo Uniwersytetu Jagiellońskiego, 2014. http://www.ceo.org.pl/pl/ok/news/strategie-oceniania-ksztaltujacego (dostęp: 27.02.2014).

Stevens, D.D., Levi A., *Introduction to Rubrics. An Assessment Tool to Save Grading Time, Convey Effective Feedback, and Promote Student Learning*. Stylus Publishing, LLC. Sterling. Virginia 2005.