

Małgorzata Brożyna-Reczko

Uniwersytet Komisji Edukacji Narodowej w Krakowie

e-mail: malgorzata.brozyna-reczko@uken.krakow.pl

ORCID: 0000-0002-5373-2171

**KORPUSY JEDNOJĘZYCZNE NA ZAJĘCIACH
Z KULTURY JĘZYKA POLSKIEGO DLA TŁUMACZY****ABSTRAKT**

Artykuł ma na celu zilustrowanie możliwych zastosowań korpusów językowych na zajęciach z kultury języka polskiego, ze szczególnym uwzględnieniem jednojęzycznych korpusów referencyjnych jako źródeł danych kolokacyjnych, które mogą być używane do weryfikacji jednostek leksykalnych w tłumaczeniu. Teoretyczne rozważania na temat korpusów w tłumaczeniach oraz kwestia kolokacji i prozodii semantycznej zostały wzbogacone o część dydaktyczną, obejmującą zestaw ćwiczeń przygotowanych dla studentów przekładoznawstwa w celu rozwijania ich umiejętności przeszukiwania korpusów, zwłaszcza przy weryfikacji połączeń wyrazowych, które są niejednokrotnie źródłem błędów oraz niejasności w tekście docelowym.

Słowa kluczowe: korpusy jednojęzyczne, kolokacje, prozodia semantyczna, błędy tłumaczeniowe, kształcenie tłumaczy

ABSTRACT**TRAINING TRANSLATION STUDENTS TO USE POLISH MONOLINGUAL CORPORA –
PRACTICAL ACTIVITIES**

The article aims to illustrate possible applications of language corpora in classroom translation activities, given the particular interest to monolingual reference corpora as sources of collocation data that may be used to verify lexical chunks in target texts. Theoretical reflection on corpora in translation and the issue of collocation and semantic prosody is followed by the practical part involving a set of classroom activities designed for translation students to develop their corpus searching skills, especially when verifying and editing word combinations.

Keywords: monolingual corpora, collocation, semantic prosody, translation errors, translator training

1. Korpusy w szkoleniu i praktyce tłumaczy

Korpusy jako zbiór autentycznych tekstów wraz z różnymi narzędziami korpusowymi mogą być cennym sposobem oraz źródłem pozyskiwania informacji o wzorcach i prawidłowościach językowych, konwencjach tekstowych¹ oraz praktykach tłumaczeniowych. Mogą one pomóc tłumaczom, gdy słowniki lub glosariusze nie dostarczają wystarczających lub precyzyjnych danych leksykalnych. Służą zatem jako zasoby językowe, szczególnie pomocne w rozwiązywaniu problemów językowych podczas tłumaczenia². Dla początkujących tłumaczy i studentów przekładoznawstwa, którzy są bardziej skłonni do stosowania dosłownych strategii tłumaczeniowych, korpusy mogą być szczególnie przydatne³, na przykład poprzez dostarczanie częściej używanych kombinacji słów, dzięki czemu brzmią one bardziej naturalnie w języku docelowym. Korzystanie z korpusów zwiększa zatem ogólną jakość tłumaczenia⁴.

Korpusy mogą również pomóc tłumaczom i studentom przekładoznawstwa w zapoznaniu się z konwencjami i normami stylistycznymi, frazeologią i terminologią poszczególnych dziedzin tematycznych, np. dokumentów prawnych, medycznych i wielu innych. Tłumacze mogą z łatwością odnieść się do danych leksykalnych, typowych kombinacji słów z porównywalnych korpusów, które mogą służyć jako materiał referencyjny, oraz ekwiwalentów z korpusów równoległych. Korzystając z zaawansowanych narzędzi korpusowych, tłumacze mogą projektować i tworzyć własne korpusy do przyjętych przez siebie celów badawczych, a następnie korzystać z narzędzi korpusowych, takich jak listy słów, listy słów kluczowych i N-gramy do analizy słownictwa i struktur gramatycznych wybranych tekstów⁵.

Liczne artykuły, badania i projekty pokazały, w jaki sposób tłumacze oraz studenci i nauczyciele przekładu mogą stosować korpusy do różnych celów, również w kontekście dydaktycznym. Seria konferencji tematycznych CULT (*Corpus Use and Learning to Translate*), poświęconych włączaniu lingwistyki korpusowej i jej

¹ A. Hardie, T. Mcenery, *Corpus Linguistics: Method, Theory and Practice*, Cambridge 2011.

² M. Brożyna-Reczko, *From text to corpus: The roles and tasks of a translator in the digital age*, [w:] *Time and Temporality*, A. Piskorz (red.), Kraków 2023, s. 223–241.

³ A. Frankenberg-Garcia, *Training translators to use corpora hands-on: Challenges and reactions by a group of thirteen students at a UK University*, „Corpora” 2015, 10 (3), s. 352.

⁴ Warto jednak dodać, że w tekstach zawartych w korpusie mogą pojawiać się błędy językowe, dlatego nie można korpusów traktować jako słowników poprawnościowych.

⁵ M. Brożyna-Reczko, *Narzędzia cyfrowe w dydaktyce przekładu: zasoby leksykalne oraz narzędzia korpusowe do edycji tekstu*, „Roczniki Humanistyczne” 2020, LXVIII, 10, s. 179–191.

narzędzi do praktyki tłumaczeniowej i nauczania, jest jednym z przykładów debaty naukowej na temat zastosowania korpusów w tłumaczeniu i dydaktyce przekładu⁶.

Warto również wspomnieć, że nauka korzystania z korpusów jako zasobów tłumaczeniowych powinna być częścią programu nauczania przyszłych tłumaczy i elementem ich kompetencji zawodowych. Umiejętność korzystania z korpusów jest uwzględniona w ramach kompetencji EMT⁷ jako element kompetencji technologicznych. Studenci powinni posiadać umiejętność korzystania z wyszukiwarek, narzędzi opartych na korpusach, narzędzi do analizy tekstu, tłumaczeń wspomaganych komputerowo (CAT) i narzędzi zapewniania jakości (QA).

Jednak, jak pokazują badania, tłumacze zazwyczaj nie korzystają z korpusów zbyt często w swojej codziennej praktyce. Silvia Bernardini⁸ przeprowadziła ankietę na temat używania korpusów przez profesjonalnych tłumaczy i okazało się, że 60,2% z 623 respondentów nie używa korpusów w swojej praktyce tłumaczeniowej, a 41,9% nigdy nawet nie słyszało o korpusach.

Podobne badanie zostało niedawno przeprowadzone w Hiszpanii. Daniel Gallego-Hernandez⁹ miał na celu zbadanie, w jaki sposób profesjonalni tłumacze pracują z różnymi zasobami tłumaczeniowymi. Okazało się, że 50% nie korzystało z korpusów nigdy lub używało ich rzadko, 30% korzystało z nich czasami, a tylko 18% – często lub bardzo często. Powstaje zatem pytanie, dlaczego tłumacze nie używają korpusów w swojej codziennej praktyce. Wydaje się, że brakuje im wiedzy i doświadczenia w wykorzystywaniu korpusów do rozwiązywania problemów tłumaczeniowych. Co więcej, nie są zaznajomieni ze strategiami wyszukiwania, co powoduje trudności w skutecznym korzystaniu z narzędzi korpusowych, wyszukiwarek¹⁰. Jednak, jak wskazują badacze¹¹, programy studiów magisterskich w zakresie tłumaczeń pisemnych zawierają specjalne moduły dotyczące korpusów tłumaczeniowych. Poza tym, korpusy powinny być integralną częścią innych

⁶ P. Rodríguez-Inés, *Corpus Use and Learning to Translate, almost 20 years on*, „Cadernos de Tradução” 2016, 36 (SPE), s. 9–13.

⁷ EMT Competence Framework, 2022, [online], https://commission.europa.eu/news/updated-version-emt-competence-framework-now-available-2022-10-21_en [dostęp: 12.03.2025].

⁸ S. Bernardini, *Corpora for translation education and translation practice: achievements and challenges*, [w:] *Proceedings of the Third International Workshop on Language Resources for Translation Research and Training (LR4Trans-III)*, Genua 2006.

⁹ D. Gallego-Hernandez, *The use of corpora as translation resources: a study based on a survey of Spanish professional translators*, „Perspectives: Studies in Translatology” 2015, 23 (3), s. 375–391.

¹⁰ A. Frankenberg-Garcia, *Training translators to use corpora hands-on...*, s. 351–380.

¹¹ Zob. np. N. Kübler, *Working with corpora for translation teaching in a French-speaking setting*, [w:] *New Trends in Corpora and Language Learning*, A. Frankenberg-Garcia, L. Flowerdew, G. Aston (red.), London 2011, s. 62–80.

modułów i kursów tłumaczeniowych i językowych, aby studenci mogli rozwinąć nawyk korzystania z korpusów i narzędzi korpusowych za każdym razem, gdy napotykają problemy tłumaczeniowe lub językowe.

W literaturze przedmiotu wyróżnia się dwa uzupełniające się podejścia do wykorzystania korpusów w dydaktyce przekładu: wykorzystanie korpusu do nauki tłumaczenia (*corpus use for learning to translate*) oraz nauka wykorzystywania korpusu do tłumaczenia (*learning corpus use to translate*)¹². W pierwszym podejściu korpus jest wykorzystywany do rozwiązywania konkretnych problemów tłumaczeniowych. Studenci otrzymują dostęp do danych korpusowych, a nauczyciel instruuje ich, jak przeszukiwać korpus i analizować dane. W drugim podejściu studenci są projektantami korpusów. Aby rozwiązać problemy tłumaczeniowe i językowe, muszą zebrać odpowiednie źródła, użyć właściwego oprogramowania do ich kompilacji i zastosować adekwatne strategie wyszukiwania. Uznaje się, że oba podejścia powinny być stosowane w procesie nauczania, aby studenci mogli zapoznać się z różnymi typami zadań, które wymagają pracy z korpusami.

Obecnie można zauważyć coraz większą liczbę gotowych korpusów łatwo dostępnych online. Poniżej zamieszczono kilka przykładów korpusów i narzędzi do przeszukiwania danych korpusowych:

- witryna COCA¹³ zapewnia bezpłatny dostęp online do szerokiej gamy korpusów języka angielskiego, np. korpusu języka angielskiego w wariantach amerykańskim (*Corpus of Contemporary American English*), korpusu współczesnych wiadomości internetowych *Now*¹⁴, do którego codziennie dodawane są nowe teksty, korpusu tekstów prasowych np. *Time*¹⁵, korpusu tekstów prawnych *Supreme Court*¹⁶;
- BNCweb – narzędzie do przeszukiwania korpusu języka angielskiego w wariantach brytyjskim¹⁷,

¹² *Corpus use for learning to translate and learning corpus use to translate*, A. Beeby, P. Rodríguez-Inés, P. Sánchez-Gijón (red.), Philadelphia–Amsterdam 2009.

¹³ M. Davies, *The Corpus of Contemporary American English (COCA)*, 2008, [online], <https://www.english-corpora.org/coca> [dostęp: 28.06.2025].

¹⁴ M. Davies, *Corpus of News on the Web (NOW)*, 2016, [online], <https://www.english-corpora.org/now> [dostęp: 28.06.2025].

¹⁵ M. Davies, *TIME Magazine Corpus*, 2007, [online], <https://www.english-corpora.org/time> [dostęp: 28.06.2025].

¹⁶ M. Davies, *Corpus of US Supreme Court Opinions*, 2017, [online], <https://www.english-corpora.org/scotus> [dostęp: 28.06.2025].

¹⁷ S. Hoffmann, Y. Berglund, D. Lee, N. Smith, *BNCweb Manual – A detailed description of BNCweb, with examples and methodological tips*, 2002, [online], <http://bncweb.lancs.ac.uk/cgibinbncXML/BNCquery.pl?theQuery=search&urlTest=yes> [dostęp: 28.06.2025].

- BNClab – narzędzie umożliwiające prowadzenie badań językowych i socjologicznych dotyczących języka angielskiego w wariantach brytyjskim¹⁸,
- platforma Sketch Engine z wieloma dostępnymi jednojęzycznymi i równoległymi korpusami, np. *OPUS*, *Europarl*, *EuroLex*, *OpenSubtitles*¹⁹,
- NKJP – *Narodowy Korpus Języka Polskiego* z trzema dostępnymi wyszukiwarkami²⁰,
- Monco PL – wyszukiwarka korpusowa, której indeks jest stale aktualizowany o ponad milion słów dziennie²¹.

Warto jeszcze wspomnieć o narzędziach do przetwarzania języka naturalnego, które są dostępne na stronie konsorcjum clarin.pl²².

2. Jednojęzyczne korpusy w praktyce tłumaczeniowej i kształceniu tłumaczy

Zgodnie z typologią EAGLES opracowaną przez Johna Sinclaira²³ korpusy dzielą się na jednojęzyczne i wielojęzyczne. Te ostatnie obejmują głównie porównywalne i równoległe korpusy w dwóch lub więcej językach. Korpusy jednojęzyczne zawierają dane tylko w jednym języku. W zależności od wyboru tekstu jednojęzyczne korpusy mogą być ogólne lub specjalistyczne.

Korpus ogólny obejmuje różne rodzaje tekstów (mówionych i pisanych) z różnych rejestrów, na różne tematy, tak aby korpus mógł reprezentować próbkę języka ogólnego. Terminy „korpus referencyjny” i „korpus ogólny” są zwykle używane w literaturze zamiennie. Geoffrey Leech twierdzi jednak, że „korpus referencyjny ma na celu dostarczenie wyczerpujących informacji o języku [...], musi to być korpus ogólny o szerokim zasięgu języka dla danego języka i będzie

¹⁸ V. Brezina, D. Gablasova, S. Reichelt, *BNClab*, Lancaster University, 2018, [online] <http://corpora.lancs.ac.uk/bnclab> [dostęp: 28.06.2025].

¹⁹ A. Kilgarriff, V. Baisa, J. Bušta, M. Jakubíček, V. Kovář, J. Michelfeit, P. Rychlý, V. Suchomel, *The Sketch Engine: ten years on*, „*Lexicography*” 2014, 1, s. 7–36.

²⁰ *Narodowy Korpus Języka Polskiego*, A. Przepiórkowski, M. Bańko, R.L. Górski, B. Lewandowska-Tomaszczyk (red.), Warszawa 2012, [online], <https://nkjp.pl> [dostęp: 28.06.2025].

²¹ P. Pęzik, *Budowa i zastosowania korpusu monitorującego MoncoPL*, „*Forum Lingwistyczne*” 2020, 7, s. 133–150.

²² *Clarin*, [online], <https://clarin-pl.eu> [dostęp: 28.06.2025].

²³ J. Sinclair, *EAGLES Preliminary recommendations on Corpus Typology*, 1996, [online], <http://www.ilc.pi.cnr.it/EAGLES96/corpusyp/corpusyp.html> [dostęp: 12.03.2025].

on traktowany przez społeczność użytkowników jako pewnego rodzaju standard dla tego języka”²⁴. Kluczową cechą każdego korpusu jest reprezentatywność, co oznacza, że korpus musi być reprezentatywny dla jakiegoś języka lub typu tekstu²⁵. Reprezentatywność ogólnego korpusu zależy od tego, jak zrównoważony jest korpus, dlatego też uwzględniona jest różnorodność typów tekstów. Główną trudnością w projektowaniu i tworzeniu zrównoważonego korpusu jest to, że należy starannie wybierać teksty, aby uwzględnić różne gatunki i wystarczającą liczbę każdej kategorii tekstów, tak aby korpus był reprezentacją języka ogólnego.

Duże korpusy referencyjne są zwykle opatrzone adnotacjami, co oznacza, że do korpusu dodano informacje językowe. Każde słowo jest powiązane z jego kategorią gramatyczną. Pozwala to na bardziej zaawansowane wyszukiwanie, które obejmuje znaczniki POS i inne kategorie morfologiczne.

Oprócz korpusów referencyjnych lub ogólnych korpusów jednojęzycznych istnieje również rodzaj korpusów, o których należy wspomnieć jako o użytecznym narzędziu w badaniach i praktyce tłumaczeniowej, a mianowicie korpusy monitorujące, które są stale aktualizowane w celu śledzenia szybkich zmian językowych. Korpusy monitorujące, jak wskazuje Piotr Pęzik, ułatwiają badanie właściwości dystrybucyjnych rzadkich słów, neologizmów i jednostek frazeologicznych. Ponadto umożliwiają badaczom i tłumaczom obserwację trendów frekwencyjnych²⁶.

Jednojęzyczne korpusy są również często wykorzystywane w nauce i nauczaniu drugiego języka. Istnieje wiele materiałów, zestawów zadań i źródeł przeznaczonych dla uczniów i nauczycieli, w których korpusy są wykorzystywane do różnych działań edukacyjnych w klasie. Artykuły w tym obszarze badań omawiają rolę i skuteczność korpusów i technik korpusowo-lingwistycznych w nauczaniu drugiego języka²⁷.

Federico Zanettin argumentuje, że zasoby korpusowe mogą być wykorzystywane podczas nauczania pisania i tłumaczenia w drugim języku. Opisuje on jedno z zadań, w którym studenci zostali poproszeni o weryfikację tłumaczenia maszynowego krótkiego tekstu, a następnie napisanie eseju wyjaśniającego, w jaki sposób wykorzystali korpusy i narzędzia korpusowe do oceny i weryfikacji tłumaczenia.

²⁴ G. Leech, *Recent grammatical changes in British English*, Paper presented at ICAME 2002, Göteborg, 22–26 May 2002.

²⁵ G. Leech, *Corpora and Theories of Linguistic Performance*, [w:] *Directions in Corpus Linguistics*, J. Star-tvik (red.), Berlin 1992, s. 116.

²⁶ P. Pęzik, *Budowa i zastosowanie korpusu monitorującego MoncoPL...*, s. 133–150.

²⁷ *Corpora and Language Teaching*, K. Aijmer (red.), Amsterdam 2009.

Zadanie i kurs miały głównie na celu rozwijanie kompetencji językowych. Dowodzi to jednak, że tłumaczenie i korzystanie z korpusów mogą być wykorzystywane na zajęciach dla osób uczących się języka²⁸.

Istnieją tysiące artykułów, które podkreślają rolę korpusów w uczeniu się i nauczaniu drugiego języka oraz niewielka liczba artykułów, które podkreślają potencjał zasobów opartych na korpusach w nauczaniu L1²⁹. Jednym z powodów jest to, że tłumacze pracujący w języku ojczystym prawdopodobnie potrzebują mniej materiałów referencyjnych niż tłumacze pracujący w języku drugim³⁰. Narzędzia korpusowe mogą jednak okazać się szczególnie przydatne w obu sytuacjach.

Frederico Zanettin porównuje jednojęzyczne korpusy docelowe do jednojęzycznych słowników docelowych, podkreślając jednak różnicę, że słowniki reprezentują syntetyczne podejście do znaczenia leksykalnego (poprzez definicję), podczas gdy korpusy – podejście analityczne (poprzez wiele kontekstów). Jeśli chodzi o duże korpusy referencyjne, Zanettin wskazuje również, że mogą one funkcjonować jako słowniki ogólne. W ten sposób można je wykorzystać do sprawdzenia znaczenia i adekwatności kandydatów do tłumaczenia w kontekstach docelowych. Wybierając ekwiwalent tłumaczeniowy, tłumacz może ocenić, czy jest on odpowiedni w danym kontekście. Co więcej, jednojęzyczne korpusy mogą służyć jako pomoc w analizie i zrozumieniu tekstu źródłowego³¹.

Zanettin nazwał również duże korpusy referencyjne „solidnymi”, gdyż są one gotowe, stworzone przez społeczność badawczą i branżę językową i są zwykle dostępne za pośrednictwem Internetu³². Prototypowymi przykładami są korpusy narodowe, takie jak BNC dla brytyjskiego angielskiego lub NKJP dla języka polskiego. Ich budowa wymaga wiele pracy i czasu, i nie są one zbyt znane wśród profesjonalnych tłumaczy, choć są dostępne dla szerszej publiczności. Potrzeba również wysiłku ze strony potencjalnych użytkowników, ponieważ korzystanie z tych korpusów wymaga wiedzy na temat tworzenia zapytań w celu uzyskania odpowiednich wyników.

Zalety korzystania z korpusów referencyjnych zostały również omówione w artykule Piotra Pęczika na temat roli *Narodowego Korpusu Języka Polskiego* w praktyce tłumaczeniowej. Wskazuje on, że korpusy referencyjne mogą służyć jako

²⁸ F. Zanettin, *Corpus-based translation activities for language learners*, „The Interpreter and Translator Training (ITT)” 2009, 3 (2), s. 209–224.

²⁹ *Handbook on Digital Learning for K-12 Schools*, A. Marcus-Quinn, T. Hourigan (red.), Springer 2017.

³⁰ A. Frankenberg-Garcia, *Training translators to use corpora hands-on...*

³¹ F. Zanettin, *Corpora in Translation Practice*, „Language Resources for Translation Work and Research” 2002, s. 10–14.

³² Ibidem.

skuteczne narzędzie do oceny ekwiwalencji frazeologicznej. Argumentuje również, że przydatność referencyjnych korpusów jednojęzycznych jest widoczna w weryfikacji poprawności jednostek frazeologicznych i leksykalno-gramatycznych. Na poparcie swoich twierdzeń podaje kilka przykładów z zadań studentów specjalności translatorskiej i wyjaśnia, w jaki sposób można wykorzystać korpusy do wyeliminowania błędów leksykalnych w każdym przypadku. Ponadto Pęzik wskazuje, że prozodia semantyczna danych fraz nie jest taka sama w tekście źródłowym i docelowym³³. Prozodia semantyczna jest kolejnym komponentem leksykalnym, którego nie należy ignorować przy wyborze odpowiedniego ekwiwalentu tłumaczeniowego. Prozodia semantyczna odnosi się do znaczenia kolokacyjnego ukrytego między słowami³⁴.

Łukasz Grabowski również przedstawia błędy tłumaczeniowe wynikające z niedopasowanej prozodii semantycznej. Jako przykład przytoczył fragment prawodawstwa Unii Europejskiej wraz z jego polskim tłumaczeniem, w którym fraza *encourage improvements* (pol. *zachęcać do poprawek*) została przetłumaczona jako *sprovokować poprawę*. Zwraca uwagę to, że *sprovokować* w języku polskim niesie ze sobą negatywne konotacje. W związku z tym kolokacja *sprovokować poprawę* jest nietypowa i idiosynkratyczna i może być niezrozumiała ze względu na zderzenie negatywnego *sprovokować* z pozytywnym *poprawa*³⁵.

Badanie to dowodzi, że błędnie zinterpretowana prozodia semantyczna może być źródłem błędów tłumaczeniowych i językowych, a w konsekwencji tekst docelowy może brzmieć nienaturalnie. W pracy tłumaczeniowej konieczne jest ustalenie, czy pewne słowa mają ogólną tendencję do pozytywnej lub negatywnej prozodii semantycznej. Prozodia semantyczna ekwiwalentów może być różna³⁶.

3. Kolokacja w praktyce tłumaczeniowej i w kształceniu tłumaczy

Omówione przykłady dowodzą, że kolokacja i powiązana z nią prozodia semantyczna mogą być istotnymi źródłami błędów tłumaczeniowych. Zofia Kozłowska i Anna Szczesny twierdzą, że studenci automatycznie tłumaczą połączenie wyra-

³³ P. Pęzik, *NKJP w warsztacie tłumacza*, [w:] *Narodowy Korpus Języka Polskiego*, A. Przepiórkowski, M. Bańko, R.L. Górski, B. Lewandowska-Tomaszczyk (red.), Warszawa 2012, s. 301–311.

³⁴ R. Xiao, T. McEnery, *Collocation, Semantic Prosody, and Near Synonyms: A Cross-Linguistic Perspective*, „Applied Linguistics” 2006, 27, s. 106.

³⁵ Ł. Grabowski, *Provoke or encourage improvements? On semantic prosody in English-to-Polish translation*, „Perspectives: Studies in Translation Theory and Practice” 2022, 30, s. 120–136.

³⁶ N. Kübler, *Working with corpora...*

zowe z języka źródłowego na język docelowy i nie mają nawyku weryfikowania kolokacji par słów w tekstach (korpusach) języka docelowego czy słownikach i polegają nadmiernie na własnej kompetencji językowej. Problematyczne dla studentów wydają się nie stałe wyrażenia i idiomy, jak można by się spodziewać, ale luźne kolokacje, w których jeden element może być wymieniany przez ograniczoną grupę słów. Jak podkreślają Kozłowska i Anna Szczęsny, tych błędów leksykalnych, które najczęściej popełniają studenci, można uniknąć, odwołując się do danych korpusowych³⁷. Piotr Pęzik twierdzi również, że trudniej jest przekonać studentów o błędach frazeologicznych w tekstach docelowych, zwłaszcza gdy tłumaczą na swój język ojczysty. Jako rodzimi użytkownicy języka czują się uprawnieni do osądów na temat frazeologii w swoim języku ojczystym. Dlatego proponuje on, aby dyskusja na temat wyborów tłumaczeniowych oparta na dowodach była elementem zajęć dla tłumaczy³⁸. Oznacza to, że nauczyciel powinien poprzeć swoje komentarze lub osądy danymi korpusowymi.

Drugi powód, o którym warto wspomnieć, odnosi się do problemu interferencji językowej. Studiując języki i kultury obce, studenci mają tendencję do używania obcych struktur gramatycznych i nietypowych kombinacji słów w swoim języku ojczystym. Przez większość czasu czytają w języku obcym, piszą w nim, analizują i ćwiczą słownictwo swojego drugiego języka, zwłaszcza na początku swojej edukacji. Pod wpływem języka obcego regularnie popełniają błędy kolokacji w nieprzetłumaczonych tekstach napisanych w języku L1.

Po zaobserwowaniu takiej tendencji i trudności studentów w wyrażaniu swoich myśli w L1 na zajęciach z kultury języka polskiego dla studentów studiów licencjackich wprowadzono specjalne zadania i ćwiczenia mające na celu zidentyfikowanie problemu interferencji i uświadomienie im pewnych braków w kompetencji w zakresie języka L1. Innym celem dydaktycznym było pokazanie studentom kilku strategii i metod, które mogą wykorzystać do rozwijania swoich umiejętności w języku ojczystym. Zaprezentowano różne narzędzia korpusowe, aby studenci mogli z nich korzystać, gdy mają wątpliwości, czy dana kombinacja słów jest typowa w języku polskim.

W jednym z zadań studenci zostali poproszeni o sparafrazowanie tekstu (z języka polskiego na polski) obejmującego zmianę stylu tekstu z nieformalnego na formalny. Oryginalny tekst został napisany w stylu nieformalnym, a zadaniem studentów było zachowanie ogólnego znaczenia tekstu, ale zmodyfikowanie jego stylu. Czas przeznaczony na wykonanie zadania wynosił jedną godzinę. Studenci

³⁷ Z. Kozłowska, A. Szczęsny, *Tłumaczenie pisemne na język polski*, Warszawa 2019, s. 48.

³⁸ P. Pęzik, *Providing corpus feedback for translators with the PELCRA Search Engine*, [w:] *Explorations Across Languages and Corpora: PALC 2009*, S. Goźdz-Roszkowski (red.), Łódź 2011, s. 135–144.

mogli korzystać ze wszystkich znanych im i potrzebnych zasobów leksykalnych, w tym słowników i korpusów.

Wybrany do tego ćwiczenia tekst został opublikowany na forum *Forum-Kulturystryka.pl* i opisuje bójkę na dyskotecie między dwiema osobami z różnych polskich subkultur³⁹. Tekst wydaje się dynamiczny ze względu na dużą liczbę użytych w nim czasowników. Zawiera również elementy leksykalne charakterystyczne dla stylu nieformalnego i slangu.

Sparafrasowane teksty (w sumie 57) zostały ocenione, błędy zostały zidentyfikowane i omówione ze studentami. W sumie okazało się, że 52% wszystkich błędów zostało sklasyfikowanych jako leksykalne i stylistyczne. Grupa ta obejmuje błędy kolokacji. Poniżej znajduje się kilka przykładów z tekstów studentów.

Zmiana strategii szybko pokazała skutki i moja dziewczyna zastosowała profesjonalne chwytty tj. „Suplex” lub „Dźwignię” przy użyciu nóg.

Moja dziewczyna, Iwona, ma 19 lat, 175 cm wzrostu i wybitną kondycję.

W pierwszym zdaniu połączenie *zmiana strategii pokazała skutki* wymaga modyfikacji. Nie występuje ono w *Narodowym Korpusie Języka Polskiego* (rysunek 1), nawet jeśli wybierze się cały korpus i zwiększy odległość między wyrazami do dwóch.

Wyszukiwarka korpusowa **PELCRA** dla danych **NKJP** (Cytowanie)

zmiana strategia** pokazać** skutek****

Maks. odstęp: Zachowaj szyk: ☒ Wyniki:

Zaawansowane Podkorpus:

Przeszukiwany zbiór zawiera 1,524,696,745 słów. Znalaziono 0 akapitów pasujących do zapytania w 0.117s.

Rysunek 1. Zgodność: *zmiana strategii pokazała skutek* (NKJP)

Analiza kolokacji dowodzi, że istnieją czasowniki, które częściej występują z rzeczownikiem *skutek*, w szczególności w lewym kontekście, np. *przynosić skutek*, *odnosić skutek*.

³⁹ *Forum kulturystryka*, [online], <http://forum-kulturystryka.pl/topic/56309-b%C3%B3jka-z-dre-siar%C4%85/> [dostęp: 12.03.2025].

Badanie to pokazuje pewne trudności studentów w poprawnym wyborze kolokacji w języku polskim. Dowodzi również, że mają oni ograniczoną wiedzę na temat zasobów leksykalnych, które mogliby wykorzystać do wyjaśnienia swoich wątpliwości językowych, oraz że stosują wzorce kolokacji i zwrotów, które znają z języka L2 (głównie języka angielskiego) w tekstach, które tworzą w języku L1. Poniżej znajduje się więcej przykładów takiej interferencji językowej z tego samego zadania.

Jako rezultat, obie kobiety miały uszkodzone nosy, podbite oczy, a agresywna kobieta uszkodzone żebra.

Dziewczyna w dresach próbowała uwolnić się od chwytu, kiedy w międzyczasie jej oddech stawał się coraz krótszy, jednak końcowo zrezygnowała i prosiła o poddanie.

W tych dwóch zdaniach użyto dwóch kalk z języka angielskiego: *Jako rezultat* [*As a result*] i *końcowo* [*finally*]. Ponadto w pierwszym zdaniu zasady interpunkcji zostały zapożyczone z języka angielskiego. W języku polskim po początkowej frazie nie stawia się przecinka.

Pojęcie interferencji językowej było szeroko dyskutowane w kontekście przyswajania języka drugiego. W literaturze przedmiotu można znaleźć opracowania dotyczące transferu pozytywnego i negatywnego, który może stać się źródłem błędów w L2. Przedstawione przykłady dowodzą, że L2 może być również źródłem błędów w L1, a studenci powinni być częściej angażowani w zadania wymagające pracy z kolokacjami, ich weryfikowania i edytowania w ich L1. Zmotywuje ich to również do sprawdzania połączeń wyrazowych w jednojęzycznych korpusach i poszukiwania odpowiednich kombinacji słów.

4. Zadania wymagające pracy z korpusami

Ponieważ studenci specjalności translatorskiej bądź przekładoznawczej na poziomie licencjackim popełniają błędy w zakresie kolokacji w języku ojczystym i mają tendencję do automatycznego tłumaczenia kombinacji słów bez korzystania z korpusów i słowników, opracowano serię praktycznych ćwiczeń, które wymagają od studentów korzystania z narzędzi korpusowych, w tym konkordancji i kolokacji. Głównym celem zadań było podniesienie świadomości studentów na temat kolokacji, zwłaszcza w procesie weryfikacji tłumaczenia, co może okazać się niezwykle przydatne w pracy postedytora⁴⁰. Ogólnym celem tych zadań jest to, aby studenci zauważyli, że korpusy są niezbędne w ocenie i wyborze kombinacji słów. Ponad-

⁴⁰ M. Brożyna-Reczek, *From text to corpus...*

to taki rodzaj zadań pomoże im zdobyć kompetencje w zakresie wyszukiwania w korpusie. Należy dodać, że zajęcia praktyczne zostały poprzedzone rozszerzoną prezentacją różnych rodzajów korpusów, w szczególności polskich i angielskich korpusów jednojęzycznych i równoległych.

ZADANIE 1

Wybierz jedno- lub dwustronicową próbkę tekstu w ogólnym, niespecjalistycznym języku angielskim (na przykład z codziennych gazet), a następnie przetłumacz go na język polski, używając narzędzi tłumaczenia maszynowego. Następnie zidentyfikuj kombinacje słów, które wymagają weryfikacji. Użyj referencyjnych jednojęzycznych i równoległych korpusów, aby przeanalizować frazy i ostatecznie je poprawić. Opisz swoją strategię wyszukiwania.

W tym zadaniu studenci muszą najpierw zidentyfikować niezgodności kolokacyjne i poprzeć swoje twierdzenia dowodami z korpusu. Na tym etapie studenci powinni użyć narzędzia konkordancji w jednojęzycznym korpusie referencyjnym, aby ustalić, czy analizowane połączenie istnieje w korpusie języka polskiego. Następnie mogą przejść do narzędzia kolokacji, aby znaleźć bardziej odpowiednie kombinacje słów w badanym kontekście. Może się okazać, że studenci będą również korzystać z innych zasobów leksykalnych, w tym słowników jedno- i dwujęzycznych lub słowników kolokacji. Poniżej znajduje się jeden z przykładów z zadań studentów wraz z komentarzem na temat strategii wyszukiwania.

ENG: It was more of an aspirational rather than a hard-and-fast goal.

PL: To był raczej cel aspiracyjny niż twardy i szybki.

W tym przykładzie student zidentyfikował dwa problemy. Pierwszy odnosi się do *twardy* i *szybki*, ponieważ *hard and fast* jest idiomem i może być przetłumaczone jako *szttywny*, *niezmienny*. Drugi problem to *cel aspiracyjny*. Według danych z NKJP słowo *cel* nie łączy się wyrazem *aspiracyjny*. W MoncoPL można znaleźć kilka przypadków, w których te dwa słowa występują razem. Warto zauważyć, że jest to stosunkowo nowa kolokacja, ponieważ najwcześniejszy przykład pochodzi z 2019 roku.

W tym przykładzie student dostarczył dowody korpusowe i pracował z dwoma rodzajami korpusów: referencyjnym i monitorującym. Co najważniejsze, był w stanie wykryć błędy leksykalne i podjąć decyzję na podstawie danych korpusowych.

Opisane zadanie może również zostać zmodyfikowane. Nauczyciel może wykorzystać dane z korpusów równoległych. Istnieje prawdopodobieństwo znalezienia kalk językowych w równoległych korpusach i poproszenia studentów o ich analizę i weryfikację.

Zadanie to jest również dobrym wprowadzeniem do bardziej zaawansowanych zadań postedycji, w których uczniowie pracują głównie ze specjalistycznymi tekstami.

Kolejne zadanie dotyczy prozodii semantycznej, która jest powiązana z kolokacją i okazuje się źródłem błędów leksykalnych. Prozodia semantyczna powinna być również omawiana ze studentami podczas zajęć dotyczących kolokacji.

ZADANIE 2

Wybierz parę synonimów w języku angielskim lub polskim i określ ich prozodię semantyczną. Jakie są ich odpowiedniki w korpusach równoległych?

W tym zadaniu studenci najpierw szukają pary synonimów, analizują je, pracują z korpusami referencyjnymi, zwłaszcza z funkcją kolokacji, zbierają dane, interpretują je i ostatecznie identyfikują podobieństwa i różnice między synonimami pod względem prozodii semantycznej. Zadanie to pomaga studentom lepiej zrozumieć kwestię synonimii w języku i tłumaczeniu.

W kolejnym kroku studenci pracują z równoległymi korpusami, szukają odpowiedników synonimów. Dzięki temu mogą rozpoznać, że w zależności od kontekstu i otaczających kolokatów synonimy mogą być tłumaczone inaczej, a tym samym mieć taką samą lub inną prozodię semantyczną. Poniżej znajduje się kilka przykładów synonimów, które studenci postanowili przeanalizować pod kątem prozodii semantycznej: *działać, czynić* / *do, work, act* / *recover, reclaim* / *znamienny, symptomatyczny* / *czuć, darzyć*.

W tej grupie znaleźli się studenci, którzy do analizy wybrali słowa ogólne, w większości polisemiczne, produktywne, czyli używane w różnych wyrażeniach i często używane w języku potocznym. Jak się okazało, studenci zebrali ogromną ilość danych, których nie potrafili odpowiednio zinterpretować. Pozostali studenci, którzy zdecydowali się analizować słowa o bardziej ograniczonym znaczeniu, byli bardziej zadowoleni z wyników i uznali zadanie za pomocne.

Ostatnie zadanie poświęcone jest jednemu z aspektów kolokacji, zwykle pojmianemu w kształceniu tłumaczy – dotyczy zmian w zakresie użycia kolokacji.

ZADANIE 3

Wybierz wyrażenie (używane we współczesnej polszczyźnie), które niedawno zyskało nowe znaczenie. Z łatwością zidentyfikujesz takie połączenia wyrazowe w tekstach medialnych lub politycznych oraz w języku mówionym (np. na forach internetowych). Następnie przeanalizuj jego znaczenie, frekwencję, kolokacje i dystrybucję. Opisz każdy etap swoich badań.

W tym zadaniu studenci powinni przeanalizować dane, np. z korpusu monitorującego, aby zidentyfikować wyrażenie, które zyskało nowe znaczenie.

Studenci byli w stanie zidentyfikować kilka słów i zwrotów w języku polskim, na przykład:

Narodowy Program Szczepień jako eksperyment kryminalny.

„gen sprzeciwu”, który tkwi w Polakach.

Zadanie to miało również na celu zilustrowanie konkretnymi przykładami, że tłumacz powinien stale obserwować zmiany zachodzące w językach.

5. Wnioski

Przeprowadzone badania dowodzą, że kolokacje nadal stanowią powtarzający się problem tłumaczeniowy i językowy z powodu interferencji językowej i braku umiejętności wyszukiwania danych w korpusach. Co więcej, studenci nie mają nawyku sprawdzania, weryfikowania lub korzystania z różnych zasobów językowych, jeśli napotkają problem językowy. Wydaje się, że ufają oni swoim kompetencjom językowym, zwłaszcza jako rodzimi użytkownicy języka. Dlatego nauczyciel powinien częściej odwoływać się do danych korpusowych, aby uargumentować swoje wybory i oceny. Studenci wówczas sami częściej będą sięgać do korpusów. Oparta na dowodach językowych dyskusja na temat wyborów tłumaczeniowych powinna być częścią każdego modułu lub kursu.

Z tych powodów studenci powinni pracować z przetłumaczonymi tekstami zawierającymi kalki językowe, które mają zweryfikować i poprawić. Co więcej, studenci powinni być świadomi własnych nawyków interferencyjnych i starać się je eliminować. Te zagadnienia powinny być również częścią zajęć z kultury języka polskiego dla tłumaczy.

Jedną z przeszkód w korzystaniu z materiałów pochodzących z korpusów jest to, że mogą one zawierać teksty z różnymi błędami językowymi. W takim przypadku studenci potrzebują pomocy nauczyciela, ponieważ takie teksty nie mogą służyć jako stylistycznie i gramatycznie poprawna próbka językowa dla tłumaczy. Z drugiej jednak strony, korpusy są zbiorami rzeczywistego użycia języka – zawierają różne warianty stylistyczne i stanowią doskonały materiał do obserwowania współczesnych zmian językowych.

BIBLIOGRAFIA

Bernardini S., *Corpora for translation education and translation practice: achievements and challenges*, [w:] *Proceedings of the Third International Workshop on Language Resources for Translation Research and Training (LR4Trans-III)*, Genua 2006.

- Brezina V., Gablasova D., Reichelt S., *BNClab*, Lancaster University 2018, [online], <http://corpora.lancs.ac.uk/bnclab> [dostęp: 28.06.2025].
- Brożyna-Reczko M., *Narzędzia cyfrowe w dydaktyce przekładu: zasoby leksykalne oraz narzędzia korpusowe do edycji tekstu*, „Roczniki Humanistyczne” 2020, LXVIII, 10, s. 179–191.
- Brożyna-Reczko M., *From text to corpus: The roles and tasks of a translator in the digital age*, [w:] *Time and Temporality*, A. Piskorz (red.), Kraków 2023, s. 223–241.
- Clarín, [online], <https://clarin-pl.eu> [dostęp: 28.06.2025].
- Corpora and Language Teaching*, A. Karin (red.), Amsterdam 2009.
- Corpus use for learning to translate and learning corpus use to translate*, A. Beeby, P. Rodríguez -Inés, P. Sánchez-Gijón (red.), Philadelphia–Amsterdam 2009.
- Davies M., *Corpus of News on the Web (NOW)*, 2016, [online], <https://www.english-corpora.org/now> [dostęp: 28.06.2025].
- Davies M., *Corpus of US Supreme Court Opinions*, 2017, [online], <https://www.english-corpora.org/scotus> [dostęp: 28.06.2025].
- Davies M., *TIME Magazine Corpus*, 2007, [online], <https://www.english-corpora.org/time> [dostęp: 28.06.2025].
- Davies M., *The Corpus of Contemporary American English (COCA)*, 2008, [online], <https://www.english-corpora.org/coca> [dostęp: 28.06.2025].
- EMT Competence Framework, 2022, [online], https://commission.europa.eu/news/updated-version-emt-competence-framework-now-available-2022-10-21_en [dostęp: 12.03.2025].
- Forum kulturystyka, [online], <http://forum-kulturystyka.pl/topic/56309-b%C3%B3jka-z-dresiar%C4%85> [dostęp: 12.03.2025].
- Frankenberg-Garcia A., *Training translators to use corpora hands-on: Challenges and reactions by a group of thirteen students at a UK University*, „Corpora” 2015, 10 (3), s. 351–380.
- Gallego-Hernandez D., *The use of corpora as translation resources: a study based on a survey of Spanish professional translators*, „Perspectives: Studies in Translatology” 2015, 23 (3), s. 375–391.
- Grabowski Ł., *Provoke or encourage improvements? On semantic prosody in English-to-Polish translation*, „Perspectives: Studies in Translation Theory and Practice” 2022, 30, s. 120–136.
- Handbook on Digital Learning for K-12 Schools*, A. Marcus-Quinn, T. Hourigan (red.), Springer 2017.
- Hardie A., Mcenery T., *Corpus Linguistics: Method, Theory and Practice*, Cambridge 2011.
- Hoffmann S., Berglund Y., Lee D., Smith N., *BNCweb Manual – A detailed description of BNCweb, with examples and methodological tips*, 2002, [online], <http://bncweb.lancs.ac.uk/cgibinbncXML/BNCquery.pl?theQuery=search&urlTest=yes> [dostęp: 28.06.2025].
- Kilgarrieff A., Baisa V., Bušta J., Jakubíček M., Kovář V., Michelfeit J., Rychlý P., Suchomel V., *The Sketch Engine: ten years on*, „Lexicography” 2014, 1, s. 7–36.
- Kozłowska Z., Szczęsny A., *Tłumaczenie pisemne na język polski*, Warszawa 2019.

- Kübler N., *Working with corpora for translation teaching in a French-speaking setting*, [w:] *New Trends in Corpora and Language Learning*, A. Frankenberg-Garcia, L. Flowerdew, G. Aston (red.), London 2011, s. 62–80.
- Leech G., *Corpora and Theories of Linguistic Performance*, [w:] *Directions in Corpus Linguistics*, J. Startvik (red.), Berlin 1992, s. 105–122.
- Leech G., *Recent grammatical changes in British English*, paper presented at ICAME 2002, Göteborg, 22–26 May 2002.
- Narodowy Korpus Języka Polskiego, A. Przepiórkowski, M. Bańko, R.L. Górski, B. Lewandowska-Tomaszczyk (red.), Warszawa 2012, [online], <https://nkjp.pl> [dostęp: 28.06.2025].
- Pęzik P., *Budowa i zastosowanie korpusu monitorującego MoncoPL*, „Forum Lingwistyczne” 2010, 7, s. 133–150.
- Pęzik P., *NKJP w warsztacie tłumacza*, [w:] *Narodowy Korpus Języka Polskiego*, A. Przepiórkowski, M. Bańko, R.L. Górski, B. Lewandowska-Tomaszczyk (red.), Warszawa 2012, s. 301–311.
- Pęzik P., *Providing corpus feedback for translators with the PELCRA Search Engine*, [w:] *Explorations Across Languages and Corpora: PALC 2009*, S. Gozdz-Roszkowski (red.), Łódź 2011, s. 135–144.
- Rodríguez-Inés P., *Corpus Use and Learning to Translate, almost 20 years on*, „Cadernos de Tradução” 2016, 36 (SPE), s. 9–13.
- Sinclair J., *EAGLES Preliminary recommendations on Corpus Typology*, 1996, [online], <http://www.ilc.pi.cnr.it/EAGLES96/corpustyp/corpustyp.html> [dostęp: 12.03.2025].
- Xiao R., McEnery T., *Collocation, Semantic Prosody, and Near Synonymy: A Cross-Linguistic Perspective*, „Applied Linguistics” 2006, 27, s. 103–129.
- Zanettin F., *Corpora in Translation Practice*, „Language Resources for Translation Work and Research” 2002, s. 10–14.
- Zanettin F., *Corpus-based translation activities for language learners*, „The Interpreter and Translator Trainer (ITT)” 2009, 3 (2), s. 209–224.