Metody ustalania ekwiwalentów czasowników wyrażających stany emocjonalne w przekładzie czesko-polskim na materiale z korpusu równoległego InterCorp

Book · April 2019

CITATIONS
5

READS
269

1 author:

Elzbieta Kaczmarska-Zglejszewska
University of Warsaw
69 PUBLICATIONS **169** CITATIONS

SEE PROFILE

WYDZIAŁ POLONISTYKI UNIWERSYTETU WARSZAWSKIEGO

Elżbieta Kaczmarska

Metody ustalania ekwiwalentów czasowników wyrażających stany emocjonalne w przekładzie czesko-polskim na materiale z korpusu równoległego InterCorp

Metody ustalania ekwiwalentów czasowników wyrażających stany emocjonalne w przekładzie czesko-polskim na materiale z korpusu równoległego InterCorp

WYDZIAŁ POLONISTYKI UNIWERSYTETU WARSZAWSKIEGO

Elżbieta Kaczmarska

Metody ustalania ekwiwalentów czasowników wyrażających stany emocjonalne w przekładzie czesko-polskim na materiale z korpusu równoległego InterCorp

Warszawa 2019

Wydano nakładem Wydziału Polonistyki Uniwersytetu Warszawskiego
e-mail: wyd.polon@uw.edu.pl

Dystrybucja: www.ksiegarniapolon.uw.edu.pl

Recenzenci

doc. PhDr. Ivana Dobrotová, Ph.D.
doc. Mgr. Martina Ivanová, CSc.

Redakcja językowa:

Agnieszka Ogrodowczyk

Korekta:

Zespół

Skład, łamanie, projekt okładki:

Agnieszka Kownacka

Zdjęcie:

Czeski Cieszyn – most. Data historyczna: 1939.
Za zgodą Ing. Ivo Petra, właściciela portalu „Fotohistorie" – http://www.fotohistorie.cz

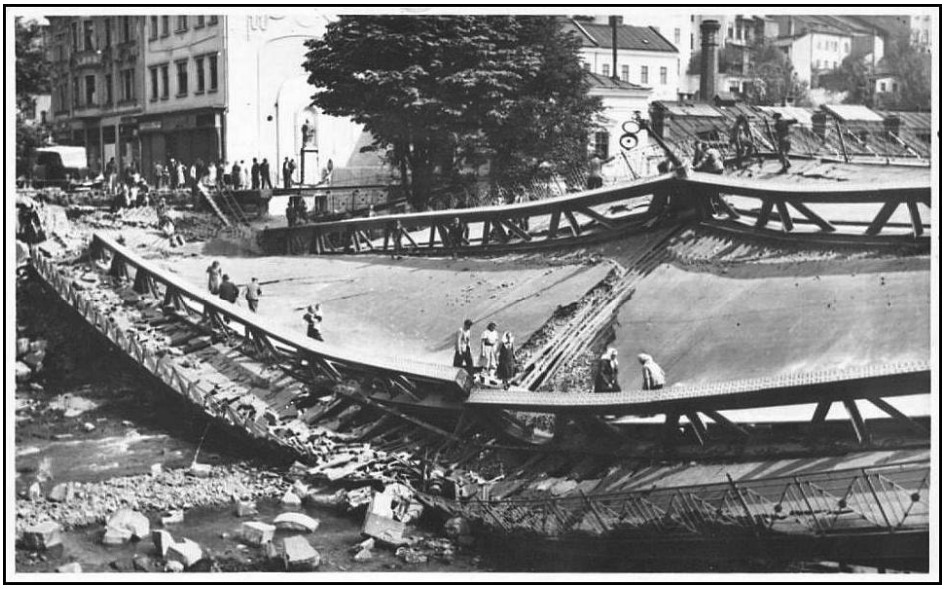
Wydanie I, elektroniczne, Warszawa 2019

Druk i oprawa:

Mazowieckie Centrum Poligrafii
ul. Ciurlionisa 4, 05-270 Marki
e-mail: w.hunkiewicz@mcpdruk.pl

ISBN 978-83-65667-93-9

Spis treści



Wstęp

Tematyka pracy

Prezentowane opracowanie koncentruje się na problemie ustalania polskich ekwiwalentów czeskich czasowników. Celem analizy jest sprawdzenie, jaka metoda potrafi najlepiej poradzić sobie z tym zadaniem oraz jakie faktory determinują wybór ekwiwalentu. Przedmiotem analizy są czeskie jednostki odnoszące się do różnych stanów psychicznych czy emocjonalnych[1]. Badanie prowadzone jest na materiale z korpusu równoległego InterCorp.

Ustalanie trafnych ekwiwalentów to często wyzwanie nie tylko w przypadku języków odległych geograficznie, kulturowo i będących dla siebie wzajemnie egzotycznych; nieoczekiwane trudności pojawiają się również na styku języków blisko spokrewnionych[2]. Niemożność oddania w języku docelowym dokładnie tego samego, co wyrażone jest w języku wyjściowym, wiąże się często z problemem precyzyjnego zrozumienia przekładanej jednostki (Kaczmarska i Rosen 2014A). Szczególnie problemowymi w tym kontekście jawią się czasowniki dotyczące

[1] Klasa czasowników dotyczących różnych stanów psychicznych czy emocjonalnych (znanych jako *psych verbs*) zawiera czasowniki percepcji, poznania, emocji. W tym opracowaniu nie będzie poddana analizie natura *psych verbs*; są tu one tylko przykładem grupy czasowników sprawiających szczególne problemy w procesie przekładu. Fenomen tych jednostek został szeroko opisany w literaturze przedmiotu; do znanych prac z tego obszaru należą opracowania (również kontrastywne) m.in. autorów (wymienione tylko niektóre z ich prac): Adriana Belletti i Luigi Rizzi (Belletti i Rizzi 1988), Agnieszka Będkowska-Kopczyk (Będkowska-Kopczyk 2014), Adam Biały (Biały 2005), William Croft (Croft 1993), Stefan Engelberg (Engelberg 2018), Barbara Lewandowska-Tomaszczyk i Katarzyna Dziwirek (Dziwirek i Lewandowska-Tomaszczyk 2009, 2010), Iwona Nowakowska-Kępna (Nowakowska-Kempna 1986, 1995, 2000), Karel Pala i Zdena Šachová (Pala i Šachová 1986), Agnieszka Spagińska-Pruszak (Spagińska-Pruszak 1994), Bożena Rozwadowska (Rozwadowska 2012), Alexandros Tantos (Tantos 2006).

[2] Oczekiwanymi trudnościami można nazwać pojawienie się „fałszywych przyjaciół", leksyki bezekwiwalentnej (Warchał 2011; Kaczmarska 2014B), czy nazw własnych, których tłumaczenie jest sztuką samą w sobie; szerzej o przekładzie nazw własnych (Cieślikowa 1996, 1998; Lewicki 2000; Hejwowski 2009; Zakharkevich 2009).

różnych stanów psychicznych czy emocjonalnych, które same w sobie mogą być źródłem niejasności i nieporozumień (Kaczmarska 2016: 228), bowiem emocje i uczucia są fenomenem trudno wyrażalnym[3]. Pisała o tym m.in. A. Wierzbicka, wspominając o relacji między przeżywaniem uczuć i mówieniem o nich (Wierzbicka 1971: 30):

> *Uczucie to jest coś, co się czuje – a nie coś, co się przeżywa w słowach. W słowach można zapisać myśli – nie można zapisać w słowach uczuć. Myśl jest czymś, co ma strukturę dającą się odtworzyć słowami. Uczucie z natury rzeczy jest pozbawione struktury, a więc niewyrażalne.*

Myśl zawarta w tym cytacie wyraża niemożliwość oddania w języku uczuć; fakt ten presuponuje również niemożliwość oddania w języku docelowym tego, co niewyrażalne w języku wyjściowym: nierealny jest bowiem trafny przekład tego, co niewyrażalne. Niniejsze opracowanie nie ma jednak na celu tego, co niemożliwe; proces ustalania ekwiwalentów czasowników wyrażających różne stany emocjonalne koncentruje się na tym, co już zostało wyrażone, czyli na znaczeniu konkretnej jednostki, a nie na tym, co przeżywa nadawca wypowiedzi. Zadanie to komplikuje jednakże wieloznaczność analizowanych jednostek; często niełatwo ustalić ich znaczenie w danym kontekście[4], w związku z tym utrudniony jest też ich przekład. Dodatkowo w procesie przekładu zatarciu ulec może część znaczenia (Kaczmarska 2015A, 2015B; Kaczmarska i Rosen 2015B, 2016). Tłumacz jest w stanie podać cały klaster ekwiwalentów, ale żaden z nich nie jest wymarzonym stuprocentowym odpowiednikiem (Lewandowska-Tomaszczyk 1984, 2010A, 2010B, 2013; Kaczmarska 2017A; Lewandowska-Tomaszczyk i Pęzik 2018).

W poszukiwaniu trafnego ekwiwalentu mógłby pomóc słownik dwujęzyczny[5], jednak ze względu na swoje ograniczenia najczęściej nie podaje on odpowiedników wraz z przykładami użycia (Siatkowski i Basaj 2002); poza tym nie rozwiązywałoby to problemu ustalania ekwiwalentów przekładowych dla konkretnego użycia słowa – tu – czasownika. Większymi możliwościami dysponują teoretycznie słowniki i translatory elektroniczne oraz słowniki umieszczone w Internecie w postaci

[3] Problem wyrażania i przekazywania emocji (również w ujęciu konfrontatywnym) poruszały także inne prace: (Wierzbicka 1991, 1997; Harkins i Wierzbicka 2001; Lüdtke 2015). Przegląd literatury na ten temat również w pracach Ulrike Oster (Oster 2010).

[4] Przekład jest procesem złożonym z kilku etapów, z których podstawowym wydaje się odkodowanie znaczenia jednostki w języku wyjściowym; już ten etap może powodować wiele trudności (Kaczmarska, Rosen 2015A).

[5] Na temat ekwiwalencji w słownikach dwujęzycznych pisał m.in. Tadeusz Piotrowski (Piotrowski 2011). Temat ten powróci w kolejnych rozdziałach niniejszego opracowania.

plików[6], jednakże dla języka czeskiego i polskiego nie są one całkowicie wiarygodne. Na przykład w przypadku analizy czeskiej jednostki *mít rád*, tłumaczonej na język polski jako *lubić* i *kochać* (Siatkowski i Basaj 2002), najpopularniejszy translator online – Google Tłumacz – jeszcze wciąż nie jest dobrym źródłem ekwiwalentów, mimo iż od pewnego czasu korzysta z rewolucyjnej metody przekładu wykorzystującej sieci neuronowe (Wu i inni 2016; Wu i inni 2018; Zennaki, Semmar i Besacier 2019). Czeska fraza *Mám rád Alenu*. jest przełożona na język polski jako *Lubię Alinę. / Kocham Alena.* Sama zaś fraza *mít rád* otrzymuje zaskakujące tłumaczenie „jak ty"[7]. Popularny portal Glosbe[8] proponuje natomiast dwa odpowiedniki – *lubić* i *podobać się*, czyniąc tę jednostkę tylko pozornie bardziej jednoznaczną, ponieważ obok ekwiwalentów przedstawia synonim – *milovat* ('kochać'). Natomiast w zamieszczonych tam przykładach znajdujemy również takie ekwiwalenty jak: *być czułym na czyimś punkcie*, *kochać* i in. Podobnie nazywany słownikiem, a działający na zasadzie translatora, portal Tri-Lite[9] tłumaczy jednostkę *mít rád* jako *jak*, natomiast czasownik *milovat* – jako *miłość.* Analogicznie funkcjonuje portal Etranslator[10], który zarówno dla *milovat*, jak i *mít rád* proponuje ekwiwalent *miłość* (Kaczmarska 2015B: 140, 2016: 229–230).

Niniejsza analiza dotyczy opisu i badania materiału dwujęzycznego, nie pretenduje jednak do miana typowego badania konfrontatywnego, w ramach którego porównywane są np. jednostki leksykalne i ich znaczenia wcześniej opisane oddzielnie w ten sam sposób dla każdego języka W prezentowanej pracy analiza nie przebiega symetrycznie; opracowanie koncentruje się na czeskich jednostkach oraz na tym, jak mogą być rozumiane przez osobę polskojęzyczną i w konsekwencji, jak są przekładane na język polski (Kaczmarska 2015B: 140, 2017B: 226). Sytuuje to tę analizę pomiędzy translatoryką a lingwistyką[11]. Pewnym ograniczeniem jest też materiał, na którym bazuje proponowana praca; analiza nie wychodzi w zasadzie poza materiał wyekscerpowany z korpusu równoległego InterCorp[12].

---

[6] Na przykład: *Czesko-polski słownik tematyczny dla uczniów i studentów* autorstwa Dariusza Sieczkowskiego i Aleny Wolfovej – http://stodolni.org/wp-content/uploads/2013/09/slowniktematyczny.pdf

[7] Dostęp 6.12.2018.

[8] Dostęp online: https://pl.glosbe.com/cs/pl/ (5.12.2018).

[9] Dostęp online: http://www.cz-pl.tri-lite.pl (5.12.2018).

[10] Dostęp online: http://www.etranslator.ro/pl/czeski-polski-online-tlumacz.php (6.12.2018).

[11] Na temat powiązań lingwistyki i translatoryki oraz badań z tego zakresu powstała bogata literatura, m.in. (Toury 1980B; Krzeszowski 1990; Ramón García 2002; Granger 2003; Malmkjær 2005; Chesterman 2007; Johansson 2007). Do zagadnień tych nawiązuje rozdział 3.

[12] W pracy pojawia się kilka przykładów z innych źródeł, przy których każdorazowo podane są dane bibliograficzne.

Struktura opisu

Opracowanie podzielone jest na trzynaście rozdziałów. Po krótkim wstępie następuje rozdział 2 omawiający zagadnienia teoretyczne dotyczące ekwiwalencji i przekładu, a także przedstawiający różne definicje ekwiwalencji oraz relacje pomiędzy przekładem tradycyjnym a wiązaniem automatycznym. Dalszą część stanowią uwagi na temat korelacji przekładu i studiów konfrontatywnych (rozdział 3). Charakterystyka korpusu równoległego InterCorp, z którego pochodzi badany materiał, stanowi rozdział czwarty. W kolejnej części autorka przedstawia metodę opisu poszczególnych leksemów (rozdział 5). Po części teoretycznej następują rozdziały prezentujące pilotażowe studia przypadków: *zdát se* (rozdział 6), *trápit* (rozdział 7), *soucítit* (rozdział 8), *závidět* i *žárlit* (rozdział 9), *mít rád* i *milovat* (rozdział 10). W dalszej części przedstawiona zostaje koncepcja projektu algorytmu ułatwiającego ustalanie ekwiwalentów (rozdział 11) oraz studium kolejnego przypadka – *toužit* – opracowanego według opisanego algorytmu (rozdział 12). Końcowa część obejmuje podsumowanie całej pracy i wnioski oraz wskazuje dalsze perspektywy badawcze (rozdział 13).

Pracę uzupełnia bibliografia oraz lista skrótów.

Cytaty

Pojawiające się w tekście cytaty w języku angielskim pozostawione są, z uwagi na upowszechnienie się tego języka, w wersji oryginalnej z odpowiednią informacją o źródle.

Przykłady z korpusu InterCorp

Cytowane przykłady dwujęzyczne opatrzone są skrótem odnoszącym się do źródła, z którego pochodzą, odrębnym dla czeskiej i polskiej wersji. Skrót zawiera informacje o autorze, tytule oraz roku wydania książki, np.:

MKZ1991cz – Milan Kundera, Žert, 1991

MKZ1999pl – Milan Kundera, Żart, 1999

Lista skrótów umieszczona jest w końcowej części pracy (Źródła).

Podziękowania

Za cenne uwagi do całej monografii chciałabym podziękować przede wszystkim recenzentkom wydawniczym – prof. Martinie Ivanovej i prof. Ivanie Dobrotovej; ich szczegółowe uwagi wpłynęły na ostateczny kształt tej pracy. Za trafne komentarze do poszczególnych rozdziałów oraz pomoc techniczną serdecznie dziękuję także dr. inż. Alexandrowi Rosenowi. Wdzięczność chciałabym wyrazić również dr. inż. Maciejowi Piaseckiemu i mgr. inż. Pawłowi Kędzi oraz dr. Jirkowi Hanie za pomoc w przeprowadzeniu badań eksperymentalnych, a także dr. Pawłowi Wolffowi za wielogodzinne objaśnianie meandrów analizy stochastycznej. Dziękuję również zespołowi Czeskiego Korpusu Narodowego za uwierzenie w pomysł automatycznego generowania listy ekwiwalentów z InterCorpu i jego realizację. Za pomoc i wieloletnie wsparcie chciałabym także podziękować prof. Zofii Zaron oraz Koleżankom i Kolegom z Instytutu Slawistyki Zachodniej i Południowej UW, w szczególności dr Milenie Hebal-Jezierskiej. Ogromną wdzięczność pragnęłabym wyrazić również moim Przyjaciołom i mojej rodzinie – Rodzicom, Siostrze i Synowi za wsparcie na każdym etapie mojej pracy, niesłychaną cierpliwość i wyrozumiałość.

Problemy ustalania ekwiwalencji

Czym jest przekład?

Pragnąc dotrzeć do istoty ekwiwalencji i ekwiwalentu, warto się zastanowić, co właściwie oznacza słowo *tłumaczenie* czy *przekład*[13]? Teoretycy postrzegają te terminy jako wieloznaczne. Jerzy Pieńkos pisze, iż tłumaczenie może wskazywać na czynność lub zespół czynności intelektualnych, polegających na formułowaniu komunikatu zawartego w tekście źródłowym (Pieńkos 1993). Następnie rozwija swoją definicję i jako kolejne ze znaczeń przytacza rezultat czynności, kiedy to tłumaczenie jest konkretnym przekładem, tekstem przetłumaczonym z innego lub na inny język w wyniku działania tłumacza. Terminy te – tłumaczenie i przekład – według J. Pieńkosa mogą oznaczać również porównywanie dwóch języków czy praktykę translatorską. Propozycje J. Pieńkosa są niejako ramami rozumienia tych terminów, w które możemy wpisać konkretną definicję. Tych bowiem – szczególnie dotyczących tłumaczenia jako procesu i rezultatu czynności – jest pełen wachlarz[14].

Olgierd Wojtasiewicz proces tłumaczenia postrzega w kategorii formułowania odpowiednika w języku docelowym:

> (…) *operacja, która polega na sformułowaniu w pewnym języku odpowiednika wypowiedzenia sformułowanego uprzednio w innym języku* (Wojtasiewicz 1957: 15).

Inne definicje konkretyzują charakter owego odpowiednika (czyli tekstu w języku docelowym) jako najbliższego i najbardziej naturalnego wobec tekstu w języku źródłowym, podnosząc m.in. kwestię zachowania odpowiedniości semantycznej i stylistycznej (Dubois 1973). Bywa też, iż znawcy przedmiotu dokonują gradacji wagi niektórych elementów, wskazując np. na prymat zachowania

[13] W tym opracowaniu przekład i tłumaczenie traktujemy jako określenia synonimiczne.

[14] Szeroko na temat pojęcia *translation* oraz jego definicji w książce Tomasza Pawła Krzeszowskiego (Krzeszowski 2016: 249–264).

sensu nad wiernością stylu (Taber i Nida 1971: 11; Pieńkos 1993: 13). Wśród elementów niezbędnych do dochowania wierności oryginalnej wypowiedzi wymienia się również reakcję odbiorcy tłumaczenia, która w założeniu ma być tożsama z reakcją odbiorcy na tekst w języku wyjściowym[15].

Jak zauważa J. Pieńkos (Pieńkos 1993: 35), najprostsza definicja określa tłumaczenie jako formułowanie w pewnym języku odpowiednika uprzednio sformułowanego w innym języku, z czego wynika, iż proces tłumaczenia jest sztuką poszukiwania ekwiwalencji dwóch wypowiedzi.

Różne typy tekstów i różne rodzaje tłumaczeń

Wpływ na kształt wypowiedzi w języku docelowym (a co za tym idzie także na kształt konkretnych ekwiwalentów) mają także inne czynniki: cel przekładu, rodzaj tekstu, sam adresat, czy stopień twórczy[16]. Analizując różne rodzaje tekstów, A. Švejcer wyróżnia kilka typów tłumaczeń [literackie (artystyczne) i nieliterackie (naukowe, techniczne[17], publicystyczne, prawno-sądowe)], w których w różnych proporcjach dostrzec można elementy twórcze i nietwórcze (Švejcer 1973: 8). Podział tekstów do tłumaczenia na różne typy nie oznacza automatycznie, jak zauważa J. Pieńkos (Pieńkos 1993), iż niektórzy tłumacze są zwolnieni z dbałości o uzyskanie jedności treści i formy, czy z wierności wobec stylu autora oryginału. Różne typy tekstów– jak należy się spodziewać –wymagają różnych strategii tłumaczenia, metod poszukiwania i uzyskiwania ekwiwalencji.

[15] Nie trzeba odwoływać się do teorii denotacji czy konotacji, aby móc stwierdzić, iż postulat zadbania o to, aby reakcja odbiorcy tekstu docelowego była tożsama z reakcją odbiorcy na tekst w języku wyjściowym, jest niezwykle trudny do zrealizowania. Każda grupa ludzi (naród, czy nawet jakieś określone pokolenie, mieszkańcy jakiegoś regionu, wykonawcy zawodu) może w jakiś szczególny sposób reagować na pewne sformułowania czy nazwy – często pochodzące z języka standardowego. W języku polskim dla pewnych pokoleń takim słowem może być pospolity wyraz „wrona" nawiązujący do WRON – Wojskowej Rady Ocalenia Narodowego, w języku czeskim – fraza „Pražské jaro" będąca nie tylko nazwą muzycznego festiwalu odbywającego się w Pradze, ale także czasu nadziei na uwolnienie się od reżimu politycznego w Czechosłowacji w 1968 roku i następującym po nim wejściu wojsk Układu Warszawskiego. Pojawienie się takich jednostek w odpowiednim kontekście może być dla innych odbiorców (również tłumacza) nieczytelne; por. aluzje erudycyjne i językowe (Wojtasiewicz 1957; Hejwowski 2009). Warto jednak podkreślić, iż reakcje czy asocjacje nie zawsze są związane z konkretnym językiem; czasem są to określenia internacjonalne, ale dostępne np. tylko dla grupy czytelników jakiejś powieści (np. ten, którego imienia nie wolno wymieniać – Voldemort).

[16] Cel, rodzaj tekstu, jego forma, styl, adresat, stopień twórczy to elementy, które związane są z ekwiwalencją i do nich w tym rozdziale będzie jeszcze mowa.

[17] Przyjmuje się, iż przekład naukowo-techniczny jest mniej wrażliwy na formę niż tekst literacki, co ma swoje konsekwencje zarówno w wyborze strategii tłumaczenia, jak i w odbiorze.

Dla tekstu artystycznego już sam przekład, jak twierdzą teoretycy przedmiotu, jest niczym „zło konieczne”[18]. Tłumacze, chcąc zbliżyć swój przekład do tekstu oryginału, przeprowadzają różne zabiegi na wielu płaszczyznach, odnosząc się m.in. do jego warstwy stylistycznej, semantycznej, kulturowej, czy estetycznej. Tekst w języku docelowym pozostaje jednak przekładem, pewnego rodzaju interpretacją (Warchał 2011: 97).

Jak zauważa Zofia Kozłowska, przekład tekstów naukowych jest obok przekładu tekstów artystycznych i przekładu tekstów nieliterackich trzecim rodzajem tłumaczenia (Kozłowska 2007). Teksty naukowe czy naukowo-techniczne różnią się od innych tekstów nie tylko na poziomie słownictwa (zawierają terminy fachowe, często dużą liczbę wyrazów obcych); są one już pisane w specyficznym dla danej dziedziny kodzie i w tym kodzie są również tłumaczone (Pieńkos 2003: 96). Należy jednak zwrócić uwagę na fakt, iż kody danych dyscyplin w różnych językach mogą się różnić (Voellnagel 2014); terminologia danej dziedziny w języku źródłowym i docelowym może być bowiem różna czy niesymetryczna, dlatego niesłychanie ważna jest dbałość o to, by terminy były precyzyjne i jednoznaczne (Kozłowska 2007).

Na sposób i metodę poszukiwania ekwiwalencji oraz w efekcie ekwiwalentu może mieć wpływ również forma tłumaczenia. Tłumaczenia pisemne to interpretacja znaczenia zapisanego w języku wyjściowym oraz stworzenie odpowiednika w języku docelowym, zawierającego w sobie tę samą *jakość* obejmującą znaczenie, emocje, styl, przekaz, kontekst kulturowy. Tłumaczenie pisemne każe też brać pod uwagę wpływ czynników gramatyczno-ortograficznych na kształtowanie się ekwiwalentu; chodzi o różnice systemowe obu języków, różnorodność alfabetów, systemów zapisu. Nie bez znaczenia dla efektu tłumaczenia jest także wykonawca tłumaczenia pisemnego; możemy oczekiwać różnic pomiędzy tłumaczeniem wykonanym przez człowieka a tłumaczeniem maszynowym, czy tłumaczeniem wspomaganym komputerowo[19] (Bowker 2002; Bogucki 2009; Kornacki 2018).

Tłumaczenia ustne[20] (Tryuk 2006, 2007) zdeterminowane są dodatkowo również przez czas. Podczas przekładu symultanicznego tłumacz siedzi w kabinie i nie ma kontaktu z osobą, której wypowiedź tłumaczy. Tłumaczenie pojawia się niemal równocześnie z tekstem oryginalnym i słuchacze (odbiorcy tłumaczenia) mają możliwość jego odbioru przez słuchawki w trakcie trwania przemówienia. W tłumaczeniu konsekutywnym tłumacz zazwyczaj znajduje się niedaleko

[18] Szerzej na temat przekładu tekstu artystycznego m.in. w pracach Anny Legeżyńskiej (Legeżyńska 1985), Edwarda Balcerzana (Balcerzan 1998), Stanisława Barańczaka (Barańczak 2004), Bożeny Tokarz (Tokarz 2010), Jiří Levý 2013; Tomasza Pawła Krzeszowskiego (Krzeszowski 2016).

[19] (CAT – ang. *computer-assisted translation*).

[20] Przekład ustny występuje w kilku formach: symultaniczny, konsekutywny, szeptany i a vista.

mówcy i ma możliwość sporządzania notatek w czasie jego przemówienia. Po wypowiedzeniu fragmentu tekstu oryginalnego, tłumacz wypowiada jego tłumaczenie. Charakter tłumaczeń ustnych i ich tempo wpływają na proces przekładu, czyniąc treść (znaczenia) elementem priorytetowym.

Wartym uwagi jest również przekład audiowizualny (Tomaszkiewicz 2006; Tryuk 2009; Tomaszkiewicz 2010; Baños, Bruti i Zanotti 2013; Deckert 2018), zawierający cechy zarówno tłumaczenia ustnego, jak i pisemnego. Z jednej strony tłumacz zmierza się z wersją lektorską (ang. *over-voice*) i z dubbingiem[21], który wymaga umiejętności dopasowania tekstu do ruchu warg wypowiadającej kwestię postaci (np. osoby, zwierzątka czy postaci baśniowej, a nawet ożywionej rzeczy). Z drugiej strony wyzwaniem są napisy filmowe, które są tłumaczeniem pisemnym, acz nietypowym, bo z góry zakładającym skróty i dopasowanie do ilości miejsca na ekranie. Przekład audiowizualny to też kwestia tłumaczenia w języku migowym i wspomniane już napisy dla niesłyszących[22] (Künstler 2009). Ekwiwalencja w tym przypadku może dotyczyć nie tylko wypowiadanych

[21] Znając tendencje filmowego (zwłaszcza telewizyjnego) przekładu choćby w Polsce (popularna wersja lektorska – ang. *voice-over* i coraz częściej pojawiające się wersje dubbingowane filmów) czy w Czechach (tradycyjnie dubbing), możemy zakwestionować podział zaproponowany przez Henrika Gottlieba (Gottlieb 1998). Gottlieb podzielił bowiem kraje na cztery grupy ze względu na dominujący w nich rodzaj przekładu audiowizualnego: 1) *source-languagecountries* (głównie kraje anglojęzyczne, w których dość rzadko spotyka się produkcje obcojęzyczne; 2) *dubbing countries* – kraje wykorzystujące technikę dubbingu, francusko-, włosko-, hiszpańsko-i niemieckojęzyczne; 3) *voice-overcountries* – kraje wykorzystujące wersję lektorską, np. Rosja czy Polska oraz kraje średniej wielkości, których nie stać na dubbing; 4) *subtitlingcountries* – kraje stosujące głównie napisy – małe kraje europejskie i pozaeuropejskie (Szarkowska 2009: 12). Podział ten przytacza bez krytycznego komentarza Agnieszka Palion-Musioł (Palion-Musioł 2012: 99). Wspominając o Czechach, warto jednak wspomnieć, że w tym kraju dubbing ma swoją długą tradycję, a wiele aktorek i aktorów zrobiło ogromną karierę, dubbingując przez cały okres aktywności jedną postać (głos konkretnej aktorki czy konkretnego aktora). Na zdezaktualizowanie się tego podziału zwróciła uwagę właśnie Agnieszka Szarkowska (Szarkowska 2009: 12–13); klasyfikacja Gottlieba powstała bowiem przed serią filmów *Harry Potter* czy *Shrek*, które w Polsce były z powodzeniem dubbingowane. Coraz częściej mamy też w kinie do wyboru wersję dubbingowaną lub z napisami; zdarza się, że wybór wersji jest związany techniką projekcji 2D lub 3D. Pod znakiem zapytania jest też umieszczenie w tym podziale krajów skandynawskich; powierzchniowo duże, jednak ze stosunkowo małą liczbą mieszkańców najczęściej wykorzystują napisy filmowe (również ze względów edukacyjnych); w przypadku tych krajów nietrafnym jest raczej stwierdzenie, iż są to kraje, których nie stać na dubbing.

[22] *Napisy dla niesłyszących różnią się od napisów dla słyszących przede wszystkim tym, że zawierają oprócz dialogów informacje o wszelkiego rodzaju odgłosach pochodzących z kanału dźwiękowego niewerbalnego* (Szarkowska 2009: 17).

Napisy dla niesłyszących to zapis ścieżki dźwiękowej przekazu audiowizualnego, zawierający treść wypowiedzi: dialogów, komentarzy, i opis wszelkich dźwięków istotnych dla zrozumienia sensu przekazu (Künstler 2009: 115).

– tłumaczonych – kwestii, ale także muzyki, dźwięków, odgłosów ulicy, czy atmosfery.

Różne ujęcia procesu tłumaczenia, jak i samego jego rezultatu zbliżają nas do istoty zjawiska ekwiwalencji.

Ekwiwalencja – propozycje definicji

Ekwiwalencja jako przedmiot badań pojawia się w wielu publikacjach i różne są jej definicje. Jak zauważa Anna Krzyżanowska, już sam termin jest nieprecyzyjny i różnie rozumiany (Krzyżanowska 2013: 36). Ta część pracy koncentruje się jednak tylko na niektórych definicjach, istotnych z punktu widzenia problemu postawionego w temacie[23].

Literatura przedmiotu przyjmuje, iż tradycyjny sposób rozumienia tłumaczenia jako zastąpienia przekazu w języku wyjściowym[24] (dalej – JW) przez równoznaczny przekaz w innym języku (języku docelowym – dalej JD), wywołuje przekonanie, iż przekład musi być ekwiwalentny i że ekwiwalencja między tekstem docelowym (dalej – TD) i tekstem wyjściowym (dalej – TW) jest koniecznym choć niewystarczającym warunkiem identyfikacji i określenia pewnego procesu i / lub jego produktu jako „tłumaczenie" (Toury 1980A: 63; Dąmbska-Prokop 2000: 68–75). W związku z tym niektóre opracowania teoretyczne dotyczące przekładu wyróżniają dwa typy ekwiwalencji: ekwiwalencję w ujęciu wąskimi ekwiwalencję w ujęciu szerokim.

Ekwiwalencja w ujęciu wąskim

Ekwiwalencja w ujęciu wąskim widziana jest jako równoważność semantyczna na poziomie języka pomiędzy tekstem wyjściowym (TW) a tekstem docelowym (TD). Innymi słowy, ekwiwalencja w tym ujęciu polega na zastępowaniu jednostek języka wyjściowego (JW) odpowiednimi jednostkami języka docelowego (JD) (Pisarska i Tomaszkiewicz 1996: 172). Urszula Dąmbska-Prokop przedstawia jedną ze strategii tłumaczenia „nie wprost", kiedy tłumacz oddaje

[23] Definicja ekwiwalencji mogłaby być tematem odrębnego, obszernego opracowania; prace dotyczące przekładu przedstawiają „szeregi" rożnych definicji ekwiwalencji, prezentowanych najczęściej chronologicznie lub pod względem następujących po sobie teorii językowych (Chlumská 2017: 11–30).

[24] W niniejszym opracowaniu pojawia się również synonimiczne określenie – język źródłowy i odpowiednio – tekst źródłowy.

sytuację z języka wyjściowego poprzez inne sformułowanie (Dąmbska-Prokop 2000: 68–75). Np. zastąpienie materiału w języku wyjściowym (JW) przez materiał w innym języku (JD) – oddanie czeskiej frazy *zabít dvě mouchy jednou ranou* jako *upiec dwie pieczenie na jednym ogniu*, zamiast dosłownego – zabić dwie muchy jednym uderzeniem.

Ekwiwalencja w znaczeniu wąskim to też poszukiwanie analogii pomiędzy dwiema jednostkami językowymi, które posiadają (prawie) tę samą denotację[25] i konotację[26], czyli poszukiwanie w języku docelowym (JD) odpowiedniej formy dla przekazania treści i wartości, które wyrażał oryginał w języku wyjściowym (JW). Tłumacz, bazując na swych kompetencjach w obu językach, porównuje znaczenie denotatywne faktów językowych i stara się zachować elementy przekazu i wypowiedzi – przenieść je z TW do TD (Dąmbska-Prokop 2000: 68–69). Nie sposób przecenić w tym miejscu porównanie znaczenia konotacyjnego branych pod uwagę faktów językowych w języku wyjściowym (JW) i języku docelowym (JD); konotacja bowiem wykracza poza poziom językowy i wyzwala skojarzenia intersubiektywne, wspólne dla pewnej grupy, a jednocześnie skojarzenia indywidualne – subiektywne.

Ekwiwalencja w znaczeniu szerokim

Ekwiwalencja w ujęciu szerokim pojmowana jest jako właściwość tekstu docelowego (TD) i zakłada istnienie pewnej relacji pomiędzy tekstem wyjściowym (TW) a tekstem docelowym (TD), niezależnie od zależności, które istnieją pomiędzy dwoma językami – pomiędzy JW a JD (Toury 1980A: 63; Dąmbska-Prokop 2000: 68–75).

Ekwiwalencja w ujęciu szerokim wiąże się też z rozróżnieniem wprowadzonym przez amerykańskiego lingwistę E. Nidę, który wyróżnił *ekwiwalencję formalną* i *ekwiwalencję dynamiczną* (Nida 1964; Nida i Taber 1974: 200; Hejwowski 2009: 38–47; Kielar 1988: 60–84). Teoretycy przekładu ukazują, iż rozróżnienie to dzieli teksty na bardziej podatne na przekład w oparciu o jeden bądź drugi typ ekwiwalencji (Kielar 1988).

[25] Denotacja – zakres danej nazwy, zbiór jej desygnatów. Denotacją nazwy „królik" jest zbiór wszystkich (przeszłych, obecnych i przyszłych) królików.

[26] Konotacja tworzy warstwę skojarzeniową słowa. To cechy pojawiające się wspólnie z nazwą, tworzące jej treść i sens; zespół cech wyrazu kojarzących się wtórnie z jego głównym znaczeniem. Por. prozodia semantyczna w dalszej części tego rozdziału.

Ekwiwalencja formalna

Ekwiwalencja formalna polega w zasadzie na odtworzeniu właściwości tekstu wyjściowego. Uwaga tłumacza skupia się na komunikacie źródłowym, jego treści i formie. Tekst przekładu musi być jak najbardziej zbliżony do tekstu wyjściowego (TW) przez zachowanie w tłumaczeniu analogicznego typu tekstu, podobnej struktury zdaniowej i frazowej (Dąmbska-Prokop 2000: 69). Według Barbary Kielar typowymi przekładami opartymi na zasadzie ekwiwalencji formalnej są teksty tłumaczone dla celów etnolingwistycznych (Kielar 1988: 61).

Ekwiwalencja dynamiczna

Ekwiwalencja dynamiczna[27] (*sense-for-sensetranslation*) dąży natomiast do takiego przekazania tekstu wyjściowego w tekście docelowym, żeby reakcja odbiorcy przekładu była podobna do reakcji odbiorców tekstu oryginalnego (Dąmbska-Prokop 2000: 70). Uwaga tłumacza skupia się na odbiorcy tekstu docelowego – na jego potencjalnej bądź faktycznej reakcji (Nida 1964: 159–166).

> *Tłumacz dążący do ekwiwalencji dynamicznej ma na celu całkowitą naturalność wypowiedzi i stara się konfrontować odbiorcę z zachowaniami mającymi znaczenie w kontekście jego własnej kultury, nie zmusza go natomiast do tego, by rozumiał wzorce kulturowe z kontekstu kultury języka wyjściowego* (Hejwowski 2009: 38).

W przypadku ekwiwalencji dynamicznej tłumacz, koncentrując się na czytelniku przekładu, stara się więc osiągnąć pełną naturalność w brzmieniu wypowiedzi w języku docelowym (JD)[28], a jednocześnie nie oczekuje od swego odbiorcy zrozumienia faktów kulturowych zawartych w języku wyjściowym (JW); jest to szczególnie istotne w sytuacji, kiedy mamy do czynienia z pojęciami JW nieistniejącymi w JD, bądź budzącymi w danych językach odmienne konotacje[29].

---

[27] Pojęcie ekwiwalencji dynamicznej wiązane jest w literaturze przedmiotu również z innymi terminami: ekwiwalencja funkcjonalna, ekwiwalencja pragmatyczna (Koller 1995), ekwiwalencja tekstowa (Baker 1992), ekwiwalencja kognitywna (Tabakowska 1993).

[28] Eugene Nida postulował, aby tekst przekładu stanowił możliwie najbliższy naturalny ekwiwalent komunikatu źródłowego. Jednocześnie zauważał, iż tłumaczenie, które buduje pomost łączący odległe od siebie kultury, nie jest w stanie uniknąć śladów obcego tła kulturowego (Kielar 1988: 60–63).

[29] Uwaga ta jest niezwykle ważna w kontekście niniejszego opracowania, w którym zaprezentowane zostaną m.in. jednostki nie mające odzwierciedlenia pojęciowego w języku docelowym; innymi słowy w języku docelowym nie odnajdujemy takiego pojęcia (w sensie konceptu), jakie wyrażone jest w języku wyjściowym.

Problem ekwiwalencji dynamicznej związany jest też z pojęciem ekwiwalentu funkcjonalnego i ekwiwalentu kulturowego (Newmark 1981). Według Petera Newmarka zastosowanie ekwiwalentu funkcjonalnego to użycie elementu neutralnego dla rożnych kultur, natomiast ekwiwalent kulturowy polega na zastąpieniu wyrażenia kulturowego w języku źródłowym innym wyrażeniem w języku docelowym. To rozróżnienie P. Newmarka mogłoby stanowić uszczegółowienie teorii E. Nidy o ekwiwalencji dynamicznej, choć sam Eugene Nida już wcześniej zdystansował się wobec tego terminu[30].

Terminem *ekwiwalencja dynamiczna* posłużył się również Tomasz Krzeszowski (Krzeszowski 1990: 15–22), który definicję ekwiwalencji w ujęciu szerokim przedstawia w zestawieniu z pojęciem *tertium comparationis*, rozumianym jako przyjęta dla celów porównawczych „platforma odniesienia" – *platform of reference* (Dąmbska-Prokop 2000: 71–72). Owa platforma odniesienia przedstawia podobieństwa dwóch porównywanych obiektów i może ukazywać jakości, które stanowią o ich różnicy[31].

Ekwiwalencja w gramatyce kognitywnej

Nieco inne postrzeganie ekwiwalencji prezentuje gramatyka kognitywna. Znaczenie w tej teorii oparte jest na konceptualizacji i obrazowaniu (Tabakowska 1991, 1995).

[30] Kiedy teoria E. Nidy zyskała popularność oraz stała się obiektem różnych interpretacji, sam E. Nida dokonał ponownego sformułowania tej teorii, w której broni tezy, iż powielenie relacji komunikacyjnej (tekst ↔ odbiorca tekstu oryginalnego / tłumaczenie ↔ odbiorca przekładu) przekład – odbiorca nie może odbywać się kosztem zerwania więzi pomiędzy przekładem a tekstem oryginału (Kielar 1988: 62–63). Eugene Nida zdystansował się też sam wobec terminu ekwiwalencja dynamiczna (*dynamic equivalence*) i przychylał się do określenia *ekwiwalencja funkcjonalna*; przy czym termin ten nie miałby wskazywać na ekwiwalencję pomiędzy funkcją tekstu źródłowego (TW) w kulturze źródłowej a funkcją tekstu w języku docelowym (w języku tłumaczenia – JD) w kulturze docelowej, ale na „funkcję", która może być rozumiana jako własność tekstu.

[31] Krzeszowski postrzega ekwiwalencję i *tertium comparationis* jako dwie strony tej samej monety.

The crucial notion in identifying various kinds of tertia comparationis and determining their character is the concept of equivalence or the relations which provides justifications for why things are chosen for comparison, keeping in mind that only equivalent items across languages are comparable. The various principles motivating equivalence and, eo ipso, contrastive studies will provide grounds for dividing tertia comparationis and, consequently, contrastive studies into various categories, each being connected with a specific kind of equivalence which motivates the comparisons. In other words, equivalence is the principle whereby tertium comparationis is established inasmuch as only such elements are equivalent for which some tertium comparationis can be found, and the extent to which a tertium comparationis can be found for a particular pair of items across languages determines the extent to which these elements are equivalent. Thus, equivalence and tertium comparationis are two sides of the same coin (Krzeszowski 1990: 21).

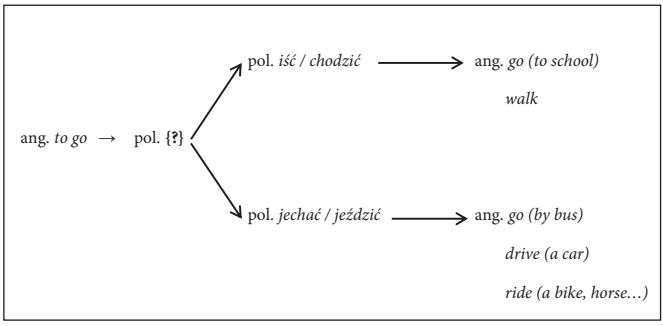
Problem *tertium comparationis* powróci w kolejnym rozdziale.

Kopia nigdy nie będzie wierna do końca – po pierwsze dlatego, że ten sam subiektywizm w tworzeniu obrazów świata, który pozwala na niezliczone warianty obrazowania, wyklucza jednocześnie identyczność interpretacji. Po drugie zaś dlatego, że techniki dostępne twórcy oryginału (językowe konwencje) mogą nie mieć odnośników w języku kopisty (Tabakowska 1991: 60).

Współgra to z spojrzeniem na przekład, widzianym jako gra *rekonceptualizacji* (Lewandowska-Tomaszczyk 2010A: 9–31, 2010B); tłumacząc wypowiedzi, dokonuje się bowiem ponownej konstrukcji danych pojęć[32]. Dzieje się tak np. w przykładzie przytaczanym przez Lewandowską-Tomaszczyk (Lewandowska-Tomaszczyk 2010B): *I'd like Tom to come earlier.* Język polski nie umożliwia pojawiania się w tłumaczeniu bezokolicznika; wymaga pojawienia się zdania podrzędnego, a w konsekwencji i wskaźników czasu gramatycznego, aspekty, liczby i rodzaju. Oprócz tego mówiący w języku polskim musi „ustalić", czy Tom idzie, czy jedzie. W rezultacie można wskazać szereg potencjalnych ekwiwalentnych zdań; każde z nich dotknięte jednak będzie rekonceptualizacją:

Chciałabym / chciałbym, żeby Tomek przyjechał / przyszedł wcześniej.

Istota rekonceptualizacji związana jest z pojęciem asymetrii semantycznej, której przykład prezentuje w jednej ze swych prac Lewandowska-Tomaszczyk:



Obraz 1. Przykład asymetrii semantycznej – ang. *displacement of senses* (Lewandowska-Tomaszczyk 1996A)

Cykle rekonceptualizacji obserwuje się również w procesie przekładu pomiędzy dwoma bliskimi językami; dotyczy to także języka czeskiego i polskiego

[32] Zjawisko rekonceptualizacji dotyczy również kontekstów monolingwalnych; znaczenie / koncept doznaje zniekształcenia (czy ponownej konceptualizacji) podczas przekazu werbalnego. To zniekształcenie (czy rekonceptualizacja) zachodzi ponownie podczas odbioru przez adresata wypowiedzi. Dochodzi do rekonceptualizacji oryginalnych znaczeń, wyłaniających się jako interpretacje wypowiedzi przez adresata (Lewandowska-Tomaszczyk 2010A; Kaczmarska 2015C).

(Kaczmarska 2015C). Wskazać tu można choćby rekonceptualizację „zubożeniową", jak np. w przypadku czeskiego czasownika *mrzet*. Dwuznaczny (dla polskiego odbiorcy) czasownik czeski podczas przekładu na język polski musi zostać ujednoznaczniony; język polski dysponuje całym zbiorem możliwych odpowiedników czasownika *mrzet*:

mrzet		
przykro	*gniewać*	*rozgoryczyć*
żałować	*zły*	*smutek*
przepraszać	***drażnić***	*ubolewać*
żal	***irytować***	*bolesny*
szkoda	*niezadowolony*	*oburzać się*
martwić	*dręczyć*	*przeprosiny*
pożałować	***mierzić***	*wkurzać*
zmartwić	*przeszkadzać*	*wyrzuty*
przykrość	*rozczarować*	*zawód*

Obraz 2. Polskie odpowiedniki czeskiego czasownika *mrzet*[33]

Struktury ze słowem *przykro* (najczęściej z zaimkiem *mi*) czy czasownik *żałować*, sugerowane zarówno przez słownik, jak i obecne wśród ekwiwalentów korpusowych, są jednak bliższe znaczeniowo czasownikowi *litovat* czy frazie *být líto*. Wobec czasownika *mrzet* są ekwiwalentne tylko w części przypadków[34]. Czasownik *mrzet* oprócz żalu zawiera także nutę złości; z tego powodu bywa tłumaczony na polski również jako *gniewać*, *drażnić*, czy *irytować*. Pomiędzy czasownikami *żałować* i *drażnić* jest jednak w języku polskim semantyczna przepaść.

Rekonceptualizacja może dotyczyć również przekładu czeskiego czasownika *toužit*[35]. Choć dla rodzimych użytkowników języka czeskiego jest to czasownik w zasadzie jednoznaczny, dla polskiego odbiorcy to wiązka znaczeń trzech polskich czasowników: *pragnąć, tęsknić, marzyć*. Wybór któregokolwiek z nich oznacza z jednej strony konkretyzację, z drugiej zaś – poważne zubożenie treści komunikatu. Również w tym przypadku dopatrzeć się można asymetrii semantycznej:

[33] Na podstawie danych z korpusu InterCorp, generowanych przy użyciu narzędzia Treq.

[34] Synonimia dotyczy przypadków, kiedy mówiący mówi np. o swoim przewinieniu – *Přykro mi.* / *Mrzí mě…* – *Je mi líto*.

[35] Czasownik ten zostanie szczegółowo omówiony w dalszej części tego opracowania.

toužit	*pragnąć*	*chtít* *toužit* *přát* *potřebovat* *prahnout* *dychtit*
	tęsknić	*stýskat* *toužit* *postrádat* *chybět* *tesknit* *scházet*
	marzyć	*snít* *toužit* *přát* *zdát se* *chtít* *těšit se*

Obraz 3. Przykład asymetrii semantycznej: czeski czasownik i jego odpowiedniki[36]

Rekonceptualizacja doprowadza do zubożenia, może nieść też ujednoznacznienie, czy rozszerzenie. Trudno jednak wyobrazić sobie sytuację, że w tłumaczonej na inny język wypowiedzi cykl rekonceptualizacyjny może w ogóle nie zaistnieć.

Ekwiwalencja funkcji

Pojęcie ekwiwalencji bywa jednak w nowszych teoriach częściowo marginalizowane.

> *Oryginał praktycznie nie jest obecny w świadomości odbiorcy przekładu. Czytając Hamleta, odbiorca sądzi, że zaznajamia się z tekstem „Szekspira", a nie tłumacza, którego nazwiska nawet nie zauważa. Można więc założyć, iż z punktu widzenia odbiorcy, ekwiwalencja wobec oryginału jest nieistotna dla funkcjonowania przekładu* (Lewicki 1993: 10).

Traktując przekład jako dzieło zamknięte i skończone, można by się z tym założeniem zgodzić. Ekwiwalencja nie dotyczy jednak tylko całego tekstu jako całości.

[36] Na podstawie danych z korpusu InterCorp.

Mierzymy się z nią, przekładając poszczególne fragmenty, czy zdania, a nawet słowa. Ekwiwalentne mogą być cel, przesłanie, a także funkcja tekstu. Właśnie autorką koncepcji ekwiwalentnej funkcji tekstu źródłowego (TW) i tekstu przekładu (TD) była Juliane House (House 1977, 1997, 2001). Według jej pierwotnej teorii tekst tłumaczenia powinien mieć funkcję ekwiwalentną wobec funkcji tekstu źródłowego (House 1977: 30; Kielar 1988: 65); zasada ta wymagałaby korekty dotyczącej elementów kulturowych, istniejących w komunikacji z udziałem pośrednika językowego. House w swojej teorii wyróżnia dwa typy tłumaczeń: tłumaczenia ewidentne (ang. *overt translation*) i tłumaczenia ukryte (ang. *covert translation*)[37].

Tłumaczenie ewidentne

Tłumaczenie ewidentne (House 1997: 66–69) nie zwraca się wprost do odbiorcy przekładu i nie jest na niego nakierowane; tekst oryginalny (TW) jest silnie osadzony w kulturze języka źródłowego (JW).

> *An 'overt' translation is one in which the addressees of the translation text are quite "overtly" not being directly addressed: thus an 'overt' translation is one which must overtly be a translation not, as it were, a "second original". In an 'overt' translation the source text is tied in a specific manner to the source language community and its culture. The source text is specifically directed at source culture addressees but at the same time points beyond the source language community because it is, independent of its source language origin, also a potential general human interest. Source texts that call for an 'overt' translation have an established worth or status in the source language community and potentially in other communities* (House 1997: 66).

Ze względu na swój charakter (treść i formę) tekst oryginalny może stać się obiektem zainteresowania czy badań innych społeczności językowych i kulturowych (House 1997: 66–69; Kielar 1988: 20)[38]; to zbliża poniekąd ideę tłumaczenia ewidentnego do omawianej wcześniej ekwiwalencji formalnej. Ponieważ w ramach tłumaczenia ewidentnego nie ma możliwości dopasowania funkcji tekstu

[37] Dla użytego przez House terminu *overt translation* (zwłaszcza w zestawieniu z *covert translation*) trafniejszym przekładem byłby przymiotnik *jawny* (tłumaczenie jawne). W tym opracowaniu autorka posługuje się jednak zaproponowanymi przez Barbarę Kielar i przyjętymi w polskojęzycznej literaturze przedmiotu terminami – tłumaczenie ewidentne i tłumaczenie ukryte.

[38] Tłumaczeniu ewidentnemu podlegają teksty uwarunkowane historycznie lub okazjonalnie (np. przemówienia polityczne), teksty o wymiarze ponadczasowym (powstają w określonym czasie i w określonej kulturze, trafiać mogą jednak do szerokiego kręgu odbiorców – np. utwory literackie. Utwory literackie nie odnoszą się zazwyczaj do jednostkowej sytuacji. Odbiorca (zarówno oryginału, jak i przekładu) potrafi odnieść fikcyjne wątki do swojego życia, wszystkie bowiem informacje niesie sam tekst (Kielar 1988: 65).

przekładu do funkcji, jaką spełniał oryginał, J. House dostrzegła konieczność zmiany funkcji tekstu w procesie tłumaczenia. House stwierdza, iż tłumacz powinien dążyć do osiągnięcia tzw. *funkcji drugiego poziomu*, co polegałoby na tym, iż tłumacz tworząc przekład, powinien myśleć nie tylko o odbiorcach przekładu, ale również o porównywalnych potencjalnych odbiorcach w kulturze wyjściowej, którzy nie są pierwotnymi adresatami. Chcąc np. tłumaczyć *Pana Tadeusza*, tłumacz uwzględniłby oczekiwania współczesnego odbiorcy w języku oryginalnym (cechy archaiczne języka z XIX, ówczesna leksyka bezekwiwalentna) i podobne znaczniki umieściłby w przekładzie. Analogicznie mogłoby się zadziać w przypadku elementów dialektalnych[39].

Tłumaczenie ukryte

W przypadku tłumaczenia ukrytego (House 1997: 69–71) tekst przekładu nie nosi cech tekstu oryginalnego – ani kulturowych, ani pragmatycznych; ma charakter i cechy oryginalnego tekstu języka docelowego (JD); nieprzenoszenie kontekstów zawartych w kulturze języka wyjściowego do tekstu przekładu może z kolei przypominać zasady ekwiwalencji dynamicznej.

> *A 'covert' translation is a translation which enjoys the status of an original source text in target culture. The translation is 'covert' because it is not marked pragmatically as a translation text of a source text but may, conceivably, have been created in its own right. A 'covert' translation is thus a translation whose source text is not specifically addressed to a particular source culture audience, i.e., it is not particularly tied to the source language and culture. A source text and its 'covert' translation text are pragmatically of equal concern for source and target language addresses. Both are, as it were, equally directly addressed. A source text and its 'covert' translation have equivalent purposes, they are based on contemporary, equivalent needs of comparable audience in the source and target language communities. In the case of 'covert' translation texts, it is thus both possible and desirable to keep the function of the source text equivalent in the translation text* (House 1997: 69).

Tłumacz musi sprostać oczekiwaniom adresatów języka docelowego, dlatego House postuluje potrzebę zastosowania *filtra kulturowego* pomiędzy tekstem wyjściowym (TW) a przekładem (TD).

> *In a 'covert' translation, the translator has to make allowances for underlying cultural differences by placing what I call a 'cultural filter' between the source text and the*

[39] Tu jednak J. House proponuje dwa rozwiązania: 1) albo pozostawić niezmienioną specyfikę kultury źródłowej jako fakt historyczny i kulturowy pomnik przeszłości, 2) albo jawnie dopasować poszczególne elementy tekstowe do kultury docelowej (Kielar 1988: 66).

translation text. The translator has, as it were, to view the source text through the glasses of a target culture member (House 1997: 70).

Tłumacz powinien spojrzeć na tekst źródłowy przez pryzmat kultury docelowej[40]. Przy wyborze metody tłumaczenia J. House zwraca uwagę na rodzaj tłumaczonego tekstu, treść tekstu, jego formę i powiazania kulturowe[41].

Pięć typów ekwiwalencji Komissarova

Zgodnie z przeświadczeniem, iż ekwiwalencja nie dotyczy wyłącznie poziomu stricte językowego, Komissarov, opierając się na koncepcji komunikacyjnej równowartości[42], wyróżnił pięć typów ekwiwalencji: leksykalny, frazeologiczny, informacyjny, sytuacyjny i związany z celem komunikacji (Komissarov 1973, 1980; Dąmbska-Prokop 2000: 69).

Na najniższym poziomie hierarchii jednostek językowych, na którym występuje ekwiwalentność przekładowa, Komissarov umieszcza ekwiwalencję leksykalną, która zachodzi na poziomie wyrazów.

Kolejnym typem ekwiwalencji pomiędzy tekstem źródłowym (TW) a tekstem docelowym (TD) jest zbieżność frazeologiczna na poziomie wypowiedzi, obejmująca struktury leksykalne i składniowe.

Następnym rodzajem wyróżnionym przez rosyjskiego językoznawcę jest ekwiwalencja polegająca na zachowaniu informacji. Dalszym wyróżnionym przez Komissarova typem jest równoważność sytuacji w tekście źródłowym (TW) i w tekście przekładu (TD). Zarówno oryginał, jak i przekład przedstawiają tę samą sytuację – rzeczywistość, jednakże ze względu na możliwość pojawienia się wyrażeń idiomatycznych oba teksty mogą ją wyrażać innymi słowami.

---

[40] Należy jednak wystrzegać się zbyt daleko idącej kreatywności, bowiem zbyt wielkie zmiany na różnych poziomach tekstu prowadzą do powstania ukrytej adaptacji, a nie do translacji (Kielar 1988: 66).

[41] J. House zwraca uwagę na konieczność wprowadzenia elementu, który początkowo sama wykluczyła – celu tłumaczenia. Faktor ten może bowiem wywrzeć decydujący wpływ na decyzje translatorskie. House rezygnuje z pierwotnego sformułowania zasady *ekwiwalentnej funkcji*; stwierdza, iż tekst przekładu może osiągnąć tę samą funkcję co oryginał tylko w przypadku tłumaczenia ukrytego, o ile odpowiednio zastosuje się filtr kulturowy. Natomiast tłumaczenie ewidentne nie prowadzi do osiągnięcia ścisłej ekwiwalencji pod względem funkcji i konieczne się staje wprowadzenie *funkcji drugiego poziomu* jako podstawy adekwatnego tłumaczenia (House 1977, 1997; Kielar 1988).

[42] Komunikacyjna równowartość oznacza, iż różne postaci komunikatu w akcie komunikacji mogą być uznane za tożsame (Kielar 1988: 74–75). Analiza tych komunikatów pozwala na ustalenie, jakie środki (z których poziomów języka) są podstawą tej równoważności.

Najważniejszym według Komissarova jest jednak typ ekwiwalencji związany z wyrażeniem za pomocą tekstu przekładu analogicznego celu komunikacji, jaki był zawarty w tekście oryginalnym. Na tym poziomie nie można konfrontować warstwy leksykalnej czy składniowej TW i TD. Odbiorca powinien zrozumieć cel wypowiedzi (Komissarov 1980: 59–61) i to ów cel determinuje przekład. Zachowanie ekwiwalencji na tym poziomie jest kluczowe i jakby nakręca spiralę równoważności na niższych poziomach. Trafnym określeniem posłużyła się Justyna Walczak (Walczak 2013: 40), pisząc o tym jako o reakcji łańcuchowej:

> *Koncepcja Komissarowa zakłada swego rodzaju reakcję łańcuchową procesu tłumaczenia. Oznacza to, że zachowanie ekwiwalencji na niższym poziomie implikuje ekwiwalencję wyższych pięter tekstowych. W przeciwnym razie tekst staje się dosłowny, co obniża jego komunikatywność lub nawet uniemożliwia odczytanie sensu komunikacyjnego.*

Ekwiwalencja w kontekście

> *Główna moja teza w odniesieniu do ekwiwalentów w tłumaczeniu jest taka, iż nie istnieją one obiektywnie, ale są tworzone przez tłumacza dla konkretnego przekładanego tekstu. Wobec tego rożnego rodzaju typologie ekwiwalentów są najczęściej po prostu chybione, ponieważ zakładają stałość znaczeniową wyrazów w dwóch językach i, co za tym idzie, stałość relacji między elementami dwóch języków. Poza tym typologie te odnoszą się do jednostek izolowanych, bez kontekstu, a jednostki leksykalne bardzo rzadko występują w izolacji, zwykle znajdujemy je w kontekście, jak nie językowym, to społecznym. W takich typologiach pisze się o „pełnych" bądź „niepełnych" ekwiwalentach* (…) (Piotrowski 2011: 47–48).

Tadeusz Piotrowski przytacza przykład angielskiego wyrazu *dog*, którego ekwiwalentem słownikowym w języku polskim jest między innymi *pies*. Zauważa też, iż tłumacz może jednak wybrać dowolną jednostkę w języku polskim (*psina, mała psina, niewierna suka, wielki brytan, wierny przyjaciel człowieka, kudłacz*)[43], kierując się zarówno cechami, jakie przypisywane są w tekście obiektowi nazwanemu *dog*, jak i cechami wiążącymi się w kulturze z danymi obiektami realnymi, wreszcie kierując się też wybranymi przez siebie strategiami (Piotrowski 2011: 48–49). W swoim tekście Tadeusz Piotrowski wyraża przypuszczenie, iż w przypadku języków bliskich – o podobnej strukturze i osadzonych w podobnej kulturze zestaw tych odpowiedników może być mniejszy niż w przypadku języków odległych. Jednak i w kontakcie pomiędzy blisko spokrewnionymi językami odnaleźć można interesujące przykłady tłumaczeń wyrazów,

[43] Po więcej tego typu przykładów Tadeusz Piotrowski odsyła do książki Krzysztofa Hejwowskiego (Hejwowski 2009) czy Kristen Malmkjær (Malmkjær 2004).

które ekwiwalentami słownikowymi absolutnie nie są, ale, jak się okazuje, bywają ekwiwalentami przekładowymi, np. *herbata* w języku polskim i *káva* w języku czeskim, jak prezentują przykłady z korpusów równoległych:

No dobrze, ale po co było wyrzucać akta i wylewać ***herbatę****? – pytają.* SLC2002pl[44]

Dobře, ale proč jste vyhodil spisy a vylil ***kávu****? – ptají se.* SLK1983cz

Mieszałem w milczeniu ***herbatę****, uważając, że jak najbardziej na miejscu będzie powściągliwość.* SLPZWW1961pl[45]

Mlčky jsem míchal ***kávu*** *v předpokladu, že nejlepší bude projevovat zdrženlivost.* SLDNVV1999cz.

Obydwa powyższe przykłady pochodzą z prozy Stanisława Lema i mogłoby się wydawać, iż problem *herbata – kawa* dotyczy tylko jednego tekstu z jego twórczości; zdania zostały wyekscerpowane jednak z różnych książek, a przekład tworzyli różni tłumacze. Co ciekawe, to samo zjawisko dotyka również innych języków[46]:

Det dröjde inte länge, förrän Pippi hade ***kaffet*** *färdigt. Och bullar hade hon bakat dagen förut.* ALPL1945se[47]

It wasn't long before Pippi had ***tea*** *ready. And she'd baked buns the day before.* ALPL1954en.

Za chvíli měla Pipi přichystanou ***kávu*** *i buchty, které napekla den předtím.* ALPDP2010cz.

Niebawem Pippi nadeszła z ***kawą****. Bułeczki upiekła już poprzedniego dnia.* ALPP1961pl.

---

[44] Przykład z korpusu InterCorp (Rosen i Vavřín 2014; Čermák i Rosen 2012; Bańczyk, Dybalska i Vavřín 2017) – Stanisław Lem, *Cyberiada*, Kraków: Wydawnictwo Literackie, 2002. Czeski przekład: Stanisław Lem, *Kyberiáda*, przeł. František Jungwirth, Praha: Československý spisovatel, 1983.

[45] Przykład z korpusu ParaSol (Waldenfels i Meyer 2006; von Waldenfels 2011; von Waldenfels 2006) – Stanisław Lem, *Pamiętnik znaleziony w wannie*, Kraków: Wydawnictwo Literackie, 1961. Czeski przekład: Stanisław Lem, *Deník nalezený ve vaně*, przeł. Pavel Weigel, Praha: Baronet, 1999.

[46] W korpusie InterCorp pojawiają się też inne przykłady, które ilustrują to ciekawe zjawisko:

Odprowadzałem Neszę do stacji metra i po drodze wstąpiliśmy do kafeterii przy Czternastej Ulicy, żeby coś przegryźć i wypić filiżankę ***herbaty***. IBSMIW1991pl (Issac Beshevis Singer, *Miłość i wygnanie*, przeł. Lech Czyżewski, Wrocław: Wydawnictwo Dolnośląskie, 1991).

Cestou jsme se stavili v bufetu na Čtrnácté ulici trochu se najíst a vypít ***kávu***. IBSLAV1997cz (Issac Beshevis Singer, *Láska a vyhnanství*, Praha: Argo, 1997).

W korpusie nie ma jednak oryginału, dzięki któremu można by stwierdzić, który z tłumaczy dokonał wybrał bardziej „rodzimy" napój. Nieobecność tekstu oryginału w korpusie równoległym to duży problem, uniemożliwiający również inne badania (Chlumská 2017).

[47] Przykład z korpusu InterCorp: Astrid Lindgren, *Pippi Långstrump*, Stockholm: Rabén & Sjögren, 1945. Przekład angielski Astrid Lindgren, *Pippi Longstocking*, przeł. Edna Hurup, Oxford: Oxford University Press, 1954. Przekład czeski: Astrid Lindgren, *Pipi Dlouhá punčocha*, przeł. Dagmar Hartlová, Josef Vohryzek, Praha: Albatros: 2010. Przekład polski: Astrid Lindgren, *Pippi Pończoszanka*, przeł. Irena Szuch-Wyszomirska, Warszawa: Nasza Księgarnia: 1961.

Wybierając wyraz *kawa* czy *herbata*, tłumacz skłania się ku kulturze obszaru języka docelowego; w tym przypadku – co pije się w danym miejscu w takich okolicznościach.

Problem ekwiwalencji i ustalenia tego, co jest ekwiwalentne łączy się bezpośrednio ze zjawiskiem nieprzekładalności, szeroko opisywanym w literaturze przedmiotu.

Nieprzekładalność czy niezwykle trudny przekład?

> *Definicja tłumaczenia tekstu a sformułowanego w języku A na język B polega na sformułowaniu tekstu b w języku B, który to tekst b wywołałby u jego odbiorców skojarzenia takie same lub bardzo zbliżone do tych, które u odbiorców wywoływał tekst a* (Wojtasiewicz 1957: 27).

Dokonując przekładu, nie sposób zawsze osiągnąć taki idealny stan, jaki określa podana definicja. Zdarza się, iż tłumacz boryka się z niemożnością przetłumaczenia, czyli właśnie sformułowania w pewnym języku odpowiednika tekstu sformułowanego uprzednio w innym języku. Według Wojtasiewicza w wielu wypadkach nie można w języku docelowym sformułować tekstu, który by u odbiorców wywoływał takie same (lub bardzo podobne) skojarzenia, jak tekst w ich języku rodzimym. Przyczyny takiego stanu upatruje się w dwóch problemach: 1) język, na jaki zamierzamy dokonać przekładu, nie rozporządza środkami strukturalnymi, istniejącymi w języku oryginału; 2) w języku, na jaki zamierzamy dokonać przekładu, nie można oddać pewnych pojęć dających się wyrazić w języku oryginału. Cechy strukturalne danych języków na ogół nie odgrywały wielkiej roli w badaniach translatologicznych, ponieważ przyjmuje się, iż dotyczą jedynie formalnej strony przekazu informacji, a nie wpływają na treść informacji[48]. Czasem jednak cechy strukturalne języków mają, przynajmniej w niektórych przypadkach, wpływ na treść informacji zawartej w danym tekście, jak choćby w przytaczanym przez Wojtasiewicza przykładzie *Astrid spoke first – Astrid odezwała się pierwsza*. Operacja wygląda na pozór na czysto formalną, ale zmienia, tu – rozszerza, treść (Wojtasiewicz 1957: 31). Przypadki problemowe pojawiają się też m.in. przy przekładach z języków bogatszych w środki strukturalne na języki uboższe w środki strukturalne, przy przekładach z języków uboższych w środki strukturalne na języki bogatsze w środki strukturalne;

[48] Badania dotyczące porównania struktur gramatycznych analizowanych języków pojawiają się w opracowaniach translatorycznych bazujących na danych korpusowych; w 2015 roku w Czechach ukazała się monografia porównująca struktury gramatyczne języka czeskiego, francuskiego, włoskiego, portugalskiego i hiszpańskiego (Čermák i Nádvorníková 2015).

trudne również są przypadki, gdy język oryginału i język przyszłego przekładu dysponują odmiennymi środkami strukturalnymi, przy czym trudno jest ustalić, który z dwóch języków rozporządza środkami bogatszymi (np. rodzajniki w języku angielskim / brak rodzajników w języku polskim, czyli konkret versus ogólnikowość)[49].

O wyjątkowo trudnej sytuacji podczas przekładu można też mówić, kiedy w grupie ludzi posługujących się językiem przekładu nie powstają takie skojarzenia, które dany wyraz czy dany zespół wyrazów budzi w grupie ludzi posługujących się językiem oryginalnym, również kiedy w języku, na jaki przekład jest dokonywany, nie można oddać pewnych pojęć dających się wyrazić w języku oryginału[50].

Sytuacje niemożności oddania w języku przekładu pewnych pojęć wyrażalnych w języku oryginału, a więc niemożność wywołania u odbiorcy identycznych skojarzeń, jakie powstają u odbiorcy oryginału dają się podzielić na grupy (Wojtasiewicz 1957: 65–98), np. przekład terminów technicznych (m.in. nazwy zwierząt – *puma*, nazwy zjawisk meteorologicznych – *mosun*, nazwy miar – *wiorsta*)[51], aluzje[52] (erudycyjne[53], językowe[54] lub połączenie erudycyjnych i językowych), czy gry językowe[55].

Komunikowanie się w różnych językach związane jest z sytuacjami, w których w jednym języku trudno wyrazić treść (myśl) zawartą w języku drugim. Często dotyczy to konkretnego leksemu, który może sprawiać problemy zarówno ze zrozumieniem, jak i przetłumaczeniem. Przez osobę tłumaczącą może to być odczuwane jako brak leksykalny w danym języku (Kaczmarska i Rosen 2014A). Braki leksykalne nie są zjawiskiem dotyczącym wyłącznie języka polskiego. O nieprzekładalnych (lub bardzo trudno przekładalnych) słowach piszą także

[49] Jeśli tłumaczymy z języka bogatszego w środki strukturalne, to otrzymujemy przekład mniej precyzyjny od oryginału: przy czym różnica zależy od tego, w jakim stopniu ścisłość sformułowań oryginału zależy od środków strukturalnych.

[50] Taka sytuacja ściśle dotyczy również części omawianych tutaj czasowników.

[51] Szerzej na ten temat (Wojtasiewicz 1957; Hejwowski 2009).

[52] Wśród aluzji często wyróżnia się całkowicie przekładalne, częściowo przekładalne i nieprzekładalne.

[53] Aluzje erudycyjne często polegają na pewnej symbolice, co może powodować różne reakcje odbiorców; np. brak określonych reakcji u odbiorcy (storczyk w Chinach – mądry i wierny doradca, minister), reakcje inne niż zamierzone w oryginale (kolor biały na Dalekim Wschodzie – kolor żałoby, w Europie – czystość, niewinność, dziewictwo, strój ślubny), reakcje, które mogą się częściowo pokrywać (w dawnych Chinach gołąb symbolizował płochość, dla Europejczyków – pojednanie i pokój, ale też zaloty miłosne) (Wojtasiewicz 1957: 87–88).

[54] Np. posługiwanie się w oryginale dialektami lub gwarami lokalnymi, czy zawodowymi (itp.), celowe popełnianie błędów gramatycznych. używanie w oryginale słów i zwrotów zaczerpniętych z języka innego niż język oryginału, ewentualnie wprowadzenie całych zdań bądź paragrafów z innego języka (Wojtasiewicz 1957: 89–94)

[55] Gry językowe wynikają często z wieloznaczności słów (polisemia, homonimia).

portale internetowe, podając najczęściej przykłady jednostek, których przekład na język angielski jest niemożliwy lub niezwykle trudny [np.: *toska* (rosyjski), *lítost* (czeski), *prozvonit* (czeski), *saudade* (portugalski), *Schadenfreude* (niemiecki)] lub wymieniając słowa, które nie istnieją jako pojedyncze leksemy w innych językach [np. *cafuné* (brazylijska odmiana portugalskiego) – delikatna pieszczona, głaskanie po włosach ukochanej osoby, *utepils* (norweski) – picie piwa na dworze w słoneczny dzień, *iktsuarpok* (inuicki) – przeczucie, że ktoś się zbliża, które sprawia, że człowiek wygląda przez okno] (Wiking 2017: 38–39).

Trudności w przekładzie oraz sytuacje nazywane przez teoretyków jako nieprzekładalność (całkowita, częściowa, obiektywna) są często opisywane w literaturze przedmiotu. Również ta praca odwołuje się do tego problemu, ukazując skomplikowany przekład niektórych jednostek czasownikowych z języka czeskiego na język polski. Warto jednak powtórzyć frazę, której autorem jest Claude Lévi-Strauss, iż *istotą języka jest możność przekładu*, a zatem teksty tworzone w jakimś języku udaje się bardziej lub mniej wiernie przełożyć (Hejwowski 2009: 16).

> *Nieprzekładalność może dotyczyć jedynie pewnych wypadków szczególnych, może nawet bardzo licznych, ale dających się interpretować jako wyjątki od ogólnej zasady przekładalności z jednego języka na drugi* (Wojtasiewicz 1957: 28).

Ekwiwalencja a korpusy

Pisząc o ekwiwalencji, warto wspomnieć, iż w odniesieniu do poświadczeń z danych korpusowych w angielskojęzycznej literaturze przedmiotu pojawia się też termin *corrrespondence*:

> *What we observe in the corpus are correspondences, and we use these as evidence of cross-linguistic similarity or difference or as evidence of features conditioned by the translation process. Analysing the correspondences we may eventually arrive at a clearer notion of what counts as equivalent across languages* (Johansson 2007A).

W tym opracowaniu autorka nie posługuje się tym określeniem.

Wiązanie automatyczne

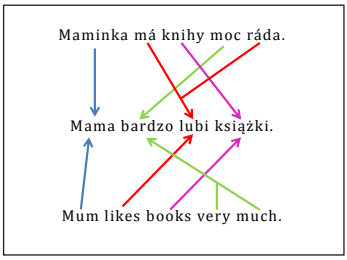
Dalekim od sztuki tłumaczenia, ale znanym i powszechnie wykorzystywanym procesem jest przekład maszynowy (Bojar 2009; 2012; Homola 2009), oparty na identyfikacji relacji translacyjnych pomiędzy słowami i jednostkami

wielowyrazowymi oraz wykorzystujący wiązanie[56] 'słowo do słowa', określane w literaturze przedmiotu jako *word to word alignment* (Och i Ney 2003; Tiedemann 2003; 2004; Kaczmarska, Rosen, Hana i Hladká 2015). Wiązanie (alignment)[57] jako automatyczne wyszukiwanie ekwiwalentów wykorzystywane jest nie tylko w tłumaczeniu maszynowym, ale także w wielojęzycznej leksykografii (Piotrowski 2011).

O wiązaniach i ich technikach pisze w swoich pracach m.in. Jörg Tiedemann, przedstawiając wiązanie słów w korpusie równoległym jako wyzwanie mające na celu identyfikację relacji przekładowych pomiędzy wyrazami a jednostkami wielowyrazowymi (*multi-wordunits*) (Tiedemann 2004: 212). Wiele strategii wiązania opartych jest na połączeniach pomiędzy pojedynczymi słowami, co w uproszczeniu pokazuje poniższy obraz.

Obraz 4. Proste wiązanie pomiędzy językami

Jednak taka równoznaczność – trafność wiązania jeden (wyraz) do jednego – nie jest sytuacją częstą. Dlatego do znalezienia optymalnego wiązania wykorzystywane są strategie uwzględniające występowanie jednostek wielowyrazowych odpowiadających jednemu wyrazowi w innym języku[58], co ponownie w uproszczeniu ukazuje poniższy obraz.



Obraz 5. Wiązania pomiędzy pojedynczymi słowami a frazami wielowyrazowymi

[56] Termin "wiązanie tekstów" (alignment) za: (Lewandowska-Tomaszczyk 2005).

[57] *Word alignment is the task of identifying translational relations between words in parallel corpora with the aim of re-using them in natural language processing. Typical applications that make use of word alignment techniques are machine translation and multi-lingual lexicography* (Tiedemann 2004: 212).

[58] *Word alignment approaches focus on the automatic identification of translation relations in translated texts. Alignments are usually represented as a set of links between words and phrases of source and target language segments* (Tiedemann 2003: 339).

W swoich pracach Tiedemann prezentuje różne algorytmy wiązania wyrazów[59] uwzględniając przy tym metody statystyczne i heurystyczne, a nawet analizę manualną (Tiedemann 2003, 2004). Jak ukazują te studia, badania nad wiązaniem mają już swoje sukcesy, a także są obiecujące pod względem przyszłych zastosowań.

Prozodia semantyczna

W pracach korpusowych dotyczących ekwiwalencji ważną rolę odgrywa pojęcie prozodii semantycznej, czyli relacji pomiędzy daną jednostką a wyrazami najczęściej się z nią pojawiającymi[60]. Victoria Kamasa (Kamasa 2015: 106) przytacza definicję Billa Louwa i pisze o prozodii semantycznej jako o powtarzającej się konsekwentnie aurze znaczeniowej, którą dany wyraz zostaje przepojony przez swoje kolokaty (Louw 1993: 156). Odnosi się też do definicji Johna Sinclaira (Sinclair 2003: 117), który twierdzi, iż prozodia semantyczna to szczególne znaczenie wyrazów, które związane jest raczej z współwystępowaniem słów i jego przyczyną niż z ich znaczeniem słownikowym (Kamasa 2015: 106)[61]. Różne definicje przytacza też Signe Oksefjell Ebeling (Ebeling S.O. 2013); we wspólnej

[59] Tiedemann, zajmując się tłumaczeniem maszynowym, podejmował również temat wiązania w napisach filmowych i telewizyjnych: np. wiązanie oparte na synchronizacji czasu (*time-based alignments*, *time-overlap algorithm*, *time slot alignment*, *sentence alignment with time overlaps*) czy długości segmentów (*length-based sentence alignment*), zdając sobie przy tym sprawę, że nie ma bezpośredniej ekwiwalencji „jeden do jednego" pomiędzy blokami napisów i zdaniami (Tiedemann 2007, 2008; Lison i Tiedemann 2016).

[60] Prozodia semantyczna jest zjawiskiem stosunkowo szeroko opisanym w literaturze (Lewandowska-Tomaszczyk 1996B; Stubbs 2001; Tognini-Bonelli 2001, 2002; Sinclair 2003; Partington 2004; Sinclair 2004; Xiao i McEnry 2006; Hunston 2007; Partington 2007; Bańko 2008; Morley i Partington 2009; Oster 2010; Smith i Nordquist 2012; Ebeling S.O. 2013; Jurko 2017).

[61] V. Kamasa szeroko omawia różne definicje prozodii semantycznej. Sama wskazuje też punkty wspólne dla różnych definicji oraz dodatkowe, relewantne dla tego zjawiska kwestie.

Dla wszystkich przytoczonych definicji wspólne jest przekonanie, że zjawisko prozodii dotyczy relacji między wybranym wyrazem a wyrazami występującymi w jego sąsiedztwie. Drugim wspólnym elementem tych definicji jest skupienie się na nacechowaniu ewaluatywnym, czy też związanym z postawami wobec określonych zjawisk. W naszym przekonaniu istotne jest także sprecyzowanie i zoperacjonalizowanie „sąsiedztwa" poprzez odniesienie do wzorców kolokacyjnych. Samo pojęcie powstało na gruncie językoznawstwa korpusowego i dlatego uznajemy, że powinno być definiowane poprzez pojęcia, których badanie jest możliwe na tym gruncie. Analizy dużych zbiorów tekstów nie pozwalają bezpośrednio orzekać na temat reprezentacji mentalnych, funkcji pełnionych przez wyrazy czy też zmian zachodzących w ich znaczeniu. Umożliwiają one jedynie obserwowanie wzorców współwystępowania o określonej sile i określonym prawdopodobieństwie, które określa się jako kolokacje. Dlatego też uznajemy, że prozodię semantyczną powinno się definiować właśnie przez odniesienie do nacechowania kolokatów danego wyrazu (Kamasa 2015: 107).

publikacji z Jarle Ebelingiem szeroko omawiają również prozodię semantyczną w badaniach kontrastywnych (Ebeling i Ebeling 2013)[62], zwracając uwagę na możliwe różnice w prozodii pomiędzy odpowiednikami w różnych językach; według autorów np. angielska jednostka (*nothing*) *out of the ordinary* i wskazywane jako jej odpowiednik norweskie słowo *uvanlig* mają różną prozodię (Ebeling i Ebeling 2013: 125, 217)[63].

Autorka zdaje sobie sprawę z możliwości wynikających z analizy prozodii semantycznej, jednak w tym opracowaniu do zjawiska tego będzie się odnosić jedynie wyjątkowo.

Ekwiwalenty w tej pracy

Badania przedstawione w tej pracy stanowią próbę ustalenia polskich ekwiwalentów wybranych czeskich czasowników wieloznacznych, odnoszących się do różnych stanów emocjonalnych. Autorka wychodzi od definicji danego leksemu zawartych w słownikach jednojęzycznych i zestawia je z odpowiednikami oferowanymi przez słownik dwujęzyczny. Ekwiwalenty słownikowe poddaje wstępnej weryfikacji, analizując poświadczenia korpusowe, wyekscerpowane z korpusu równoległego; odpowiedniki badanego czasownika znajdujące się w tych poświadczeniach można uznać za ekwiwalenty przekładowe. Zastosowane przez autorkę metody mają na celu nie tylko utworzenie zaktualizowanej listy ekwiwalentów (również słownikowych), ale przede wszystkim ułatwienie doboru odpowiedników – ekwiwalentów przekładowych – ujednoznacznionej jednostki czeskiej i wskazanie czynników, które ten dobór warunkują[64].

[62] Zob. też badania słoweńsko-angielskie (Jurko 2017).

[63] Wiąże się to poniekąd z pojęciem konotacji w przekładzie (Dąmbska-Prokop 2000).

[64] Intuicyjnie można by chcieć nazwać poszukiwane jednostki ekwiwalentami kontekstowymi, jednak nie należy zapominać, iż nazwa ta jest zarezerwowana dla rezultatów tłumaczenia *inteligentnego*, uwzględniającego kontekst współwystępujących słów w celu trafniejszego doboru odpowiedników. Przekład kontekstowy (ang. *Context-Based Machine Translation*) wykorzystuje analizę statystyczną korelacji różnych ekwiwalentów z występującymi w danym „kontekście" wyrazami czy strukturami gramatycznymi w celu identyfikacji ekwiwalentów językowych o jednoznacznej wartości semantycznej w konkretnym kontekście (Kaczmarska, Rosen, Hana i Hladká 2015).

Przekład a studia kontrastywne

W ostatnich latach pomiędzy studiami kontrastywnymi[65] i badaniami translatorskimi doszło do szczególnego zacieśnienia i wzajemnego wzbogacenia; przyczyn można upatrywać między innymi w tym, iż wykorzystują podobne zasoby i posługują się podobnymi metodami. W szczególności analiza tłumaczonych tekstów zajmuje ważne miejsce w obu dziedzinach, nawet jeśli perspektywy i cele epistemologiczne są różne. Poza metodami obserwacyjnymi (opartymi na korpusach) na obu polach pojawia się coraz więcej eksperymentalnych analiz np. statystycznych czy wykorzystujących *eye-tracking* (pol. okulografia, śledzenie ruchów gałek ocznych). Najnowsze badania w obu dziedzinach często ukazują ich wzajemne przenikanie się i, co być może istotniejsze, ich synergię.

Językoznawstwo kontrastywne ma w świecie nauki długą tradycję i w zależności od tradycji naukowych często określane jest jako komparatywne, porównawcze lub konfrontatywne; przy czym warto zaznaczyć, iż dla wielu badaczy określenia te nie są w pełni synonimiczne[66]; zajmujący się pracami komparatywnymi stosują terminy: *badania konfrontatywne* (Korytkowska i Mazurkiewicz-Sułkowska 2014; Korytkowska 1992), *cross-linguistic analysis*/*cross-linguisticstudy* (Lewandowska-Tomaszczyk 1999), *contrastive analysis* (James 1980), *contrastive studies* (Krzeszowski 1990), *comparative descriptive linguistics* (Ellis 1966)[67].

Znaczenie lingwistyki kontrastywnej jest nie do przecenienia, a jej osiągnięcia przyczyniają się do rozwoju innych dziedzin. Nie bez powodu mówi się też o niej, iż opisuje podobieństwa i różnice pomiędzy dwoma językami (lub większą ich liczbą) na poziomie fonologii, gramatyki czy semantyki w celu usprawnienia nauki języków obcych i udoskonalenia tłumaczenia (McArthur 1992;

[65] Mowa tu o lingwistycznych studiach kontrastywnych.

[66] Terminy te odnoszą się do różnych typów badań czy wręcz języków, np. badania porównawcze często utożsamiane są analizami diachronicznymi, a kontrastywne ze względu na terminologię – z językami germańskimi (niem. *konfrontative Linguistik*). W tym opracowaniu określenia *kontrastywne*, *komparatywne* i *konfrontatywne* uważane są za synonimiczne i odnoszące się do badań synchronicznych, natomiast fraza *badania porównawcze* dotyczy badań diachronicznych.

[67] Na niezręczność terminu *contrastive linguistics* zwracał też uwagę Jacek Fisiak (Fisiak 1980).

Skandera i Burleigh 2005)[68]; lingwistyce kontrastywnej przypisuje się też zasługi w budowaniu porozumienia pomiędzy kulturami oraz w tworzeniu słowników wielojęzycznych (Tuong 2016: 21). Jarle Ebeling i Signe Oksefiell Ebeling w swojej książce wskazują na różne cele badań na przestrzeni lat; wcześniejsze studia usiłowały identyfikować obszary w języku, które były wyjątkowo trudne czy niezrozumiałe dla obcokrajowców uczących się danego języka[69]. Stopniowo zaczęły ukazywać się bardziej systematyczne badania dwu (lub większej liczby) języków, co doprowadziło do wykrystalizowania się różnicy pomiędzy teoretycznymi i stosowanymi studiami kontrastywnymi[70]. Ta różnica jest widoczna również w aktualnych badaniach (Ebeling i Ebeling 2013: 13).

Badania kontrastywne na przestrzeni lat definiowane były często jako kilkupoziomowa analiza, złożona z „kroków" lub „etapów"; Lado[71], Di Pietro[72] i Krzeszowski postulowali przeprowadzanie trzystopniowych procedur w konfrontowaniu

[68] (…) *a branch of linguistics that describes similarities and differences among two or more languages at such a level as phonology, grammar, and semantics, especially in order to improve language teaching and translation* (McArthur 1992).

Contrastive linguistics describes similarities and differences between two or more modern languages especially in order to improve language teaching and translation (Skandera i Burleigh 2005: 2).

[69] Studia kontrastywne mają długą tradycję i jako metoda badawcza zaczęły się pojawiać od końca XIX wieku (Jaszczolt 2011).

[70] Szczegółowo o teoretycznych problemach gramatyki kontrastywnej w opracowaniach m.in. Jacka Fisiaka (Fisiak 1980).

[71] Roberto Lado postulował trzystopniową analizę: 1) ustalenie optymalnego opisu struktur (funkcji językowych) analizowanych języków, 2) podsumowanie wszystkich struktur, 3) właściwe porównanie danych języków (Lado, 1957: 67). Głównym celem podejścia Roberta Lado było podkreślenie potencjalnych problemów osób uczących się danego języka jako obcego. Ebelingowie zastanawiają się jednak, na jakiej podstawie Roberto Lado ustala ekwiwalencję poszczególnych fraz, np. ang. *Is he a farmer?* i hiszp. *Es un campesino?* Potrzebne są według nich dowody poza formami wyrażonymi na powierzchni, aby ustalić, czy te zdania w ogóle można porównywać (Ebeling i Ebeling 2013: 14–15).

[72] Gramatyka generatywno-transformacyjna upatrywała w badaniach konfrontatywnych możliwość porównania struktur badanych języków; podobieństwa i różnice struktur syntaktycznych analizowanych języków mogą być zestawiane z analizą obu języków w strukturze głębokiej. Trzystopniowe kontrastywne badanie prowadzone w tym duchu miało obejmować według Roberta Di Pietro następujące elementy: 1) identyfikację różnic pomiędzy strukturami powierzchniowymi dwu języków, 2) stwierdzenie podstawowych uniwersaliów analizowanych struktur powierzchniowych, 3) sformułowanie reguł przechodzenia od struktury głębokiej do powierzchniowej dla każdego języka objętego analizą kontrastywną (Di Pietro 1971: 29–30). Roberto Di Pietro zwraca uwagę na możliwości składniowe analizowanych języków, ale nie ukazuje, jakie okoliczności wpływają na użycie konkretnych struktur. Natomiast różnicę w stosowaniu porównywalnych zasad w analizowanych językach tłumaczy różnicami w stylu pomiędzy tymi językami. Roberto Di Pietro nie nobilitował też specjalnie roli translatoryki w badaniach konfrontatywnych; według niego tłumaczenie mogłoby być jednak odpowiednią techniką do zainicjowania badań kontrastywnych (Di Pietro 1971: 49). Por. (Ebeling i Ebeling 2013: 15–16).

dwóch lub większej liczby języków (Lado 1957; Di Pietro 1971; Krzeszowski 1990).

Tomasz P. Krzeszowski przedstawia trzystopniowy projekt analizy kontrastywnej[73] oraz wskazuje na konieczność znalezienia wspólnej własności porównywanych obiektów – *tertium comparationis*[74]; jeśli porównywane obiekty maja coś wspólnego, w opozycji do podobieństw można ustalić różnice. Celem analizy kontrastywnej jest bowiem często znalezienie elementów, które są ekwiwalentne w dwu językach pod pewnymi względami, przy czym elementy te mogą być tłumaczeniami dla siebie wzajemnie (Krzeszowski 1984). Sytuacja komplikuje się jednak, ponieważ dane obiekty mogą być porównywane z uwzględnieniem różnych cech; w efekcie mogą się okazać pod pewnymi względami podobne, pod innymi względami – różne[75]. W zależności od platformy odniesienia (*tertium comparationis*), którą zostanie zastosowana, te same obiekty mogą się okazać podobne albo różne.

Czeskie i polskie badania konfrontatywne

Uznani na świecie teoretycy badań kontrastywnych, m.in. Jacek Fisiak[76], Tomasz Krzeszowski[77], Jarle Ebeling i Signe Oksefiell Ebeling[78]czy Volker Gast[79] przedstawiają w swoich pracach ważne według nich prace kontrastywne[80], bazujące zarówno na ekscerpowanym ręcznie materiale, jak i danych pochodzących z korpusów; wśród wymienianych prac czeskie i polskie studia są jednak rzadkością[81].

[73] Te stopnie to: 1) opis, 2) zestawienie, 3) analiza właściwa (Krzeszowski T 1990).

[74] (…) *properties which the compared items share but which are outside the scope of comparison itself* (Krzeszowski 1990: 117).

[75] Krzeszowski wyjaśnia to na przykładzie kwadratu i prostokąta. Pod pewnymi względami są podobne – mają cztery ściany i kąty proste; pod innymi względami są inne – jeżeli chodzi o długości boków.

[76] (Fisiak 1980, 1984).

[77] (Krzeszowski 1990).

[78] (Ebeling i Ebeling 2013).

[79] (Gast 2012A, 2012B).

[80] M.in. Roberto Lado (Lado 1957), Roberto Di Pietro (Di Pietro 1971), Carl James (James 1980), Andrew Chesterman (Chesterman 1998), Bengt Altenberg (Altenberg 1999), Sylviane Granger (Altenberg i Granger 2002; Granger 2010).

[81] Najczęściej cytowane są badania angielskojęzyczne i odnoszące się do języka angielskiego.

Czeskie i polskie synchroniczne badania konfrontatywne również mają swą tradycję; zestawiają różne języki, prezentują szeroki wachlarz tematyczny, reprezentują różne nurty studiów kontrastywnych i dotyczą między innymi[82]:

- natury gramatycznej języków: kategorii przypadka semantycznego w języku polskim, bułgarskim i serbsko-chorwackim (Korytkowska 1984), wykładników modalności imperceptywnej w języku polskim i litewskim (Roszko 1993), formacji nominalizowanych w oryginale oraz czeskim i polskim przekładzie *Harrego Pottera* (Kaczmarska 2007), imiesłowu przysłówkowego współczesnego *gérondif* w języku francuskim i jego czeskich ekwiwalentów (Nádvorníková 2010), bezokolicznika jako dopełnienia czasowników w języku czeskim i słoweńskim (Stankovska 2018);
- zagadnień semantycznych: znaczeń angielskiego czasownika *to see* i jego polskiego czasownika *widzieć* (Zawisławska 1998), anglicyzmów w czeskiej i polskiej prasie dla mężczyzn (Kaczmarska 2006A), angielskich ekwiwalentów czeskich przyimków (Klégr, Malá i Šaldová 2012), procesu adaptacji anglicyzmów w języku słowackim (Greň 2015), leksykografii przekładowej polsko-rosyjskiej (Charciarek 2018), walencji i synonimii czasowników w dwujęzycznym, czesko-angielskim kontekście (Urešová, Fučíková, Hajičová i Hajič 2018);
- ekwiwalencji pragmatycznej (funkcjonalnej): nominalizacji w języku polskim i słowackim (Papierz 1982), czeskiego i polskiego przekładu niemieckiego passivum i zdań z „*man*"–Sätze (Rytel-Kuc 1990), związków frazeologicznych w językach słowiańskich (Kaczmarska 2003), obcych nazw geograficznych w języku czeskim i polskim (Siatkowski 2006), realizacji kategorii określoności w języku polskim i słowackim (Papierz 2009), polskich wyrażeń metatekstowych i ich odpowiednikach w języku czeskim i rosyjskim (Charciarek 2010), zapożyczeń niemieckich w języku polskim, czeskim, słowackim i łużyckim (Kaczmarska i Kłos 2012), związków frazeologicznych w języku czeskim, polskim i rosyjskim (Dobrotová i Chlebda 2015), korpusowego badania czeskiego i angielskiego języka mówionego (Čermáková 2018).

W literaturze przedmiotu liczne są też pozycje konfrontujące bezpośrednio język czeski i polski; przedmiotem ich zainteresowania są różne elementy

[82] Poniżej wymienione badania to tylko nieliczne wybrane przykłady opracowań niecytowanych w tej pracy w innych miejscach. Badania konfrontatywne dotyczące języka polskiego i czeskiego stanowią liczną grupę badań – szczególnie w zestawieniach z językiem angielskim, rosyjskim, niemieckim czy językami romańskimi.

gramatyczne i semantyczne[83], np.: *Leksykalne środki wyrażania modalności w języku czeskim i polskim* (Rytel 1982), *Z problematyki konfrontatywnego badania czasownika w językach blisko spokrewnionych (na materiale polskim i czeskim)* (Greń 1992), *Uniwerbizacja w języku czeskim a polskim* (Szczepańska 1994), *Semantyka i składnia czasowników oznaczających akty mowy w języku polskim i czeskim* (Greń 1994), *Multiwerbizacja w języku czeskim i polskim* (Mietła 1998), *Badanie struktury walencyjnej czeskich i polskich predykatów posiadających pozycję Experiencera* (Kaczmarska 2001), *Tykání a jeho zdvořilejší protějšek v češtině a polštině* (Skwarska 2001), *Czeskie czasowniki 'vzkázat' /'vyřídit' i ich polskie ekwiwalenty* (Greń 2003), *Nominalizacje odczasownikowe w języku polskim i czeskim (wybrane zagadnienia)* (Kaczmarska 2003), *Jeszcze o czesko-polskich pułapkach językowych* (Szczepańska 2004), *Netykieta i dyskusje polityczne w czeskim i polskim Internecie* (Dobrotová 2005), *Nominalizacje w czeskiej i polskiej prasie dla rodziców małych dzieci* (Kaczmarska 2006B), *Funkcja stylistyczna konstrukcji z nominalizacjami w języku polskim i czeskim* (Kaczmarska 2008A), *Polskie i czeskie artykuły sponsorowane* (Kaczmarska 2008B), *Relacje temporalne w konstrukcjach z nominalizacjami w języku polskim i czeskim* (Kaczmarska 2009B), *Kilka uwag o tytułach artykułów jako skrytych komunikatach reklamowych (na podstawie polskiej i czeskiej prasy dla mężczyzn)* (Kaczmarska 2009A), *Analiza zdolności konotacyjnych polskich i czeskich predykatów odnoszących się do strachu, złości i wstydu* (Kaczmarska 2010), *O „bezpiecznej" prawdzie w tabloidach polskich i czeskich* (Kaczmarska, Stefaniak i Susfał 2011), *Tendencje rozwojowe we współczesnym polskim i czeskim systemie adresatywnym* (Charciarek 2014), *K teorii mezijazykové interference v lexikálním plánu češtiny a polštiny* (Muryc 2015), *Polsko-czeska homonimia jako problem translatorski* (Muryc 2017), *Symbolika v českých a polských somatických frazémech s lexémy sluch/sluch a ucho a jejich vzájemná ekvivalence* (Tkaczewski 2017A), *Symbolika w czeskich i polskich frazeologizmach somatycznych z leksemami zrak/wzrok i oko oraz ich ekwiwalencja wzajemna* (Tkaczewski 2017B), *O obrazie i stereotypie rzemieślnika w polskich i czeskich mediach* (Kolberová 2018), *Mentalnościowy i językowy „obraz sąsiada". Polsko-czeskie i czesko-polskie stereotypy językowe i aproksymaty* (Tkaczewski 2018).

Część wymienionych prac wprost odwołuje się do zjawiska ekwiwalencji istniejącej na różnych poziomach języka i wypowiedzi, np. odpowiedniości funkcji gramatycznych (Rytel 1982), ekwiwalencji pragmatycznej (Skwarska 2001), czy leksykalnej (Charciarek 2018)[84].

[83] Również w tym przypadku zostały wymienione tylko niektóre prace spośród synchronicznych badań konfrontatywnych polsko-czeskich.

[84] Wymienione zostały tylko przykładowe badania; kategorie te są licznie reprezentowane przez konfrontatywne badania czeskie i polskie.

Wśród tekstów kontrastywnych polsko-czeskich pojawiają szczególnie bliskie prezentowanej pracy, opracowane z wykorzystaniem korpusu InterCorp[85]: *Polské konstrukce s neosobními tvary slovesa na -no/-to a jejich české překladové ekvivalenty – analýza úrovně ekvivalence* (Bańczyk 2010), *Jak se překládají české univerbizáty do polštiny* (Hebal-Jezierska 2010), *How good is parallel corpus data? Contrasting Polish and Czech must and can in InterCorp* (von Waldenfels 2010), *Korespondence kolokací s lexémem ženský/mužský v překladu z češtiny do polštiny* (Zasina 2016), *Překlad a korpus. Textová ekvivalence vybraných českých adjektiv* (Zasina 2016), *Možno stivyužití korpusu InterCorp v česko-polské překladové lexikografii* (Charciarek 2018)[86].

Badania kontrastywne oparte na materiale z korpusu równoległego InterCorp nie wykorzystują jednej uniwersalnej metody; różne analizy (gramatyczne, stylistyczne, słowotwórcze) wymagają zastosowania innych podejść. Cechą wiążącą te badania jest jednak bazowanie na wyrównanych segmentach[87], bez względu na to, który z typów wyszukiwania został wybrany. Segmenty te są tłumaczeniami, a ponieważ nie wszyscy teoretycy badań kontrastywnych dopuszczają korzystanie z tłumaczeń (Krzeszowski 2016), już sam wybór korpusu równoległego jako źródła materiału językowego jest określoną decyzją metodologiczną[88].

[85] Badania kontrastywne i translatologiczne prowadzone są na danych pochodzących z różnych typów korpusów. Sylviane Granger oraz Anthony McEnery i Zhonghua Xiao w swych pracach wskazują na problemy terminologiczne dotyczące korpusów wielojęzycznych: *comparable corpus, parallel corpus, translation corpus* (McEnery i Xiao 2008; Granger 2010). O wykorzystaniu korpusów wielojęzycznych w pracach konfrontatywnych mowa również w innych opracowaniach, m.in. (Johansson 1998, 2007A; Anderman i Rogers 2008; McEnery i Xiao 2008; Gast 2012B).

[86] Do badań konfrontatywnych polsko-czeskich wykorzystujących dane z korpusu InterCorp należą również prace autorki cytowane w tym opracowaniu.

[87] Segmenty te po eksportowaniu do dokumentu edytowalnego można dowolnie tagować, co ułatwia analizę zarówno jakościową, jak i ilościową.

[88] Por. o wykorzystaniu przekładów literackich w pracy nad dwujęzycznym słownikiem walencyjnym (Greń i Rytel-Kuc 1991).

Dane korpusowe

InterCorp

InterCorp[89] (Čermák i Rosen 2012; Kaczmarska i Rosen 2014B; Rosen 2016; Rosen i Vavřín 2014; Rosen, Vavřín i Zasina 2018) to akademicki i niekomercyjny projekt, który powstał w Pradze na Wydziale Filozoficznym Uniwersytetu Karola. Jego celem jest wybudowanie obszernego równoległego korpusu synchronicznego, który zawierałby teksty z jak największej liczby języków[90]. InterCorp jest jednocześnie nazwą tego ciągle rozrastającego się synchronicznego korpusu paralelnego, obejmującego aktualnie 40 języków (stan na marzec 2019). InterCorp jako korpus jest też częścią Czeskiego Korpusu Narodowego[91].

InterCorp zawiera szereg jednobrzmiących tekstów, których „parą" jest zawsze tekst czeski (bądź jako oryginał, bądź jako tłumaczenie). Język czeski jest więc dla InterCorpu językiem kluczowym[92].

Korpus równoległy[93] służy, między innymi, jako źródło danych do badań teoretycznych, prac translatorskich, analiz gramatycznych i leksykograficznych, projektów dotyczących nauki języków obcych, aplikacji komputerowych, czy poszukiwań studenckich[94].

---

[89] Źródło i dostęp: http://www.korpus.cz/intercorp/

[90] W pierwotnym założeniu miały to być wszystkie języki nauczane na Wydziale Filozoficznym Uniwersytetu Karola w Pradze.

[91] www.korpus.cz

[92] Inaczej wygląda sytuacja w przypadku korpusu ParaSol, również wielojęzycznego korpusu równoległego: *W odróżnieniu od znacznie większego korpusu InterCorp w ParaSolu większą wagę przykłada się do reprezentacji i zrównoważenia jak największej liczby języków i żaden z tych języków nie jest podstawowy, tak jak podstawowy jest czeski w korpusie InterCorp, gdzie każdy tekst musi mieć czeską wersję*. – (Łaziński 2014). Por. (Łaziński, Kuratczyk i inni 2012; von Waldenfels 2012).

[93] Obszernie o różnych korpusach równoległych (Gruszczyńska i Leńko-Szymańska 2016).

[94] O wykorzystywaniu korpusów równoległych między innymi również w: (Lewandowska-Tomaszczyk 2005: 43; Johansson 2007A; Ebeling i Ebeling 2013; Kaczmarska i Rosen 2013: 105; Kaczmarska i Rosen 2014B: 207; Čermák i Nádvorníková 2015; Čermáková, Chlumská i Malá 2016; Hebal-Jezierska, Kaczmarska i Rosen 2016; Cosma i inni 2016). Warto zauważyć, iż możliwość

W tworzeniu InterCorpu uczestniczą pracownicy i studenci Wydziału Filozoficznego Uniwersytetu Karola (FF UK), osoby związane z Czeskim Korpusem Narodowym, a także współpracownicy zewnętrzni.

Opis korpusu InterCorp

Korpus InterCorp składa się z dwóch części – trzonu (tzw. *core*) i kolekcji (*collections*). Trzon stanowią przede wszystkim teksty beletrystyczne, ręcznie wiązane[95]. Oprócz nich korpus obejmuje również automatycznie opracowane teksty; są to widoczne w interfejsie wspomniane wcześniej kolekcje. Obecnie w ramach tych kolekcji udostępnione są artykuły publicystyczne, teksty ze stron Project Syndicate oraz Presseurop, akty prawne Unii Europejskiej z korpusu Acquis Communautaire, zapisy posiedzeń parlamentu Europejskiego z lat 2007–2011 (z korpusu Europarl), bazy napisów dialogowych z serweru Open Subtitles, a także *Biblia*. Teksty w kolekcjach są wiązane automatycznie, dlatego w wyszukanych konkordancjach może pojawić się więcej zdań, które sobie nawzajem nie odpowiadają.

Wszystkie teksty w InterCorpie mają zawsze wersję czeską (jako oryginał lub tłumaczenie); jak już wcześniej zostało wspomniane, język czeski jest dla InterCorpu językiem kluczowym (*pivot*), a czeska wersja jest wiązana z jedną lub wieloma wersjami innojęzycznymi.

Najnowszą edycją jest opublikowana w październiku 2018 r. wersja 11 InterCorpu[96]. Całkowita objętość dostępnej części tej wersji korpusu to w przybliżeniu 283 milionów słów w wyrównanych obcojęzycznych tekstach w trzonie i 1225 milionów słów w wyrównanych obcojęzycznych tekstach w kolekcjach (Rosen, Vavřín i Zasina 2018). Nawiązania do wersji 11 są jednak w tym opracowaniu jedynie incydentalne. Badania przedstawione w tej pracy przeprowadzono w oparciu o dane pochodzące z wersji 10 (oraz wcześniejszych)[97].

wykorzystania przekładów literackich w pracach nad dwujęzycznym słownikiem walencyjnym została zauważona dużo wcześniej (Greń i Rytel-Kuc 1991).

[95] Termin „wiązanie tekstów" (alignment) za: (Lewandowska-Tomaszczyk 2005).

[96] Wcześniejsze wersje (począwszy od wersji 6) są jednak ciągle dostępne.

[97] Całkowita objętość dostępnej części wersji 10 InterCorpu (opublikowanej w 2017 roku) to 361 417 611 słów w tekstach trzonu korpusu oraz 1 314 875 602 słów w tzw. kolekcjach (Rosen, Vavřín i Zasina 2017).

Zazwyczaj raz w roku publikowana jest nowa wersja InterCorpu. Z każdą nową wersją rośnie objętość tekstu, mogą pojawiać się nowe języki, zmiana anotacji tekstów czy korekty dotyczące kolekcji[98].

W perspektywie InterCorp ma zawierać również teksty, które nie mają czeskiej wersji; jeszcze jednak nie wiadomo, kiedy to nastąpi. InterCorp straciłby wówczas swoją wyraźną orientację na język czeski (Kaczmarska i Rosen 2014B).

Teksty w tym korpusie mają anotację zewnętrzną i wewnętrzną. Jeżeli chodzi o anotację wewnętrzną czy segmentację tekstu, dla każdego języka może być ona inna[99]. Ta anotacja jest kwestią zasadniczą dla twórców korpusu (Hebal-Jezierska, Kaczmarska i Rosen 2016: 41–65). Dla użytkowników niezwykle istotną podczas wyszukiwania, a przede wszystkim późniejszych badań, jest anotacja zewnętrzna; dzięki wskaźnikom – kryteriom można dowolnie filtrować materiał już na etapie wyszukiwania. Do dyspozycji jest kilkanaście tzw. atrybutów dotyczących samego tekstu, między innymi: język tekstu (doc.lang), identyfikacja tekstu (text.id), przynależność do określonej części korpusu – trzon, kolekcje – (text.group), autor tekstu (text.autor), tytuł tekstu (text.title), miejsce wydania (text.pubplace), rok wydania (text.pubyear), typ tekstu (text.txtype) czy wreszcie ważne z punktu widzenia badań nad przekładem: język oryginału (text.srclang), tłumacz tekstu (text.translator), płeć tłumacza (text.transsex), płeć autora (text.authsex) i informacja kluczowa dla analiz dotyczących przekładu i ekwiwalencji – czy tekst jest oryginałem (text.original)[100]. Zastosowanie pozostałych oznaczeń może mieć znaczenie w zależności od typu przeprowadzanego badania. Relewantna dla analizy może być np. płeć autora lub tłumacza, jeśli w badaniu bierzemy pod uwagę, czy tekst został napisany przez kobietę, czy przez mężczyznę[101] (np. przy porównywaniu tekstów w kryminalistyce). Nieocenione w badaniach językowych jest też oznaczenie języka oryginału (Kaczmarska i Rosen 2014B: 210).

W korpusie InterCorp domyślne ustawienia obejmują wszystkie teksty wybranych języków. Chcąc zawęzić liczbę przeszukiwanych teksów, można każdorazowo na etapie wyszukiwania wprowadzić ograniczenia (o czym wspomniano wyżej) lub utworzyć subkorpus (Kaczmarska i Rosen 2014B: 220). Wytworzenie subkorpusu jest wygodniejszym sposobem; umożliwia to bowiem korzystanie

[98] Niektóre teksty z korpusów Acquis Communautaire i Europarl były częściowo poprawione i wyselekcjonowane, dlatego mogą się różnić od oryginału formą i objętością. Zredukowana została także baza napisów dialogowych (Kaczmarska i Rosen 2014B: 208).

[99] Szerzej na ten temat w: (Kaczmarska i Rosen 2014B: 210–212). Również na stronach internetowych InterCorpu są odesłania do opisów anotacji dla danych języków.

[100] Atrybut „text" pojawił się dopiero w wersji 11; do tej pory występował w tym miejscu atrybut „div" (Kaczmarska i Rosen 2014B: 210).

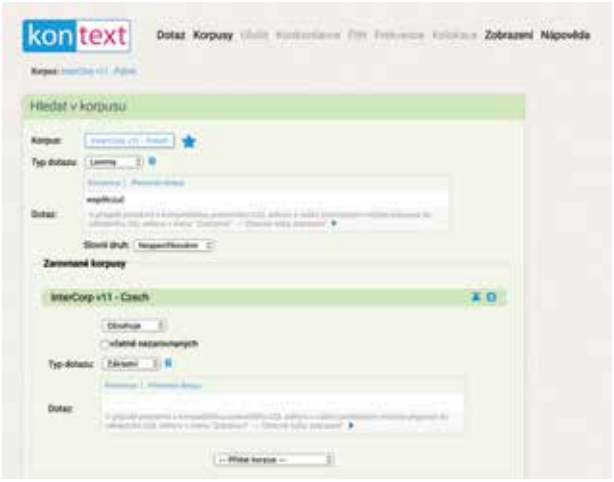
[101] Próby badań tego typu, choć niezwykle trudnych w rzetelnym przeprowadzeniu, były już podejmowane (Leonardi 2007; Zaliwska-Okrutna 2002; Simon 1996).

z dokładnie tego samego zestawu tekstów przy każdym badaniu[102]. Możliwość tworzenia subkorpusów poprzez interfejs KonText w podanym zakresie możliwa jest jednak dopiero od wersji 7 InterCorpu, opublikowanej 19.12.2014 r.

Dostęp do tekstów

Obecnie InterCorp można przeszukiwać, używając dostępnego od stycznia 2014 r. interfejsu KonText; od kwietnia 2015 r. jest to jedyna wyszukiwarka Czeskiego Korpusu Narodowego (Hebal-Jezierska 2014: 29–52; Kaczmarska i Rosen 2014B: 212). Wcześniej wykorzystywane były również inne interfejsy: Park i NoSketch Engine (Kaczmarska i Rosen 2014B: 212–228).

Ikonka KonTextu widoczna jest po otwarciu strony głównej korpusu (www.korpus.cz)[103]. Wykorzystanie wszystkich funkcji tej wyszukiwarki (i tym samym dostęp do tekstów) możliwe jest po darmowej rejestracji (https://www.korpus.cz/signup). Rejestrując się, użytkownik zyskuje dane dostępowe do wszystkich korpusów ČNK; KonText obsługuje bowiem zarówno korpusy jednojęzyczne, jak i równoległe (Hebal-Jezierska 2014: 29–52; Kaczmarska i Rosen 2014B: 212).



Obraz 6. Interfejs wyszukiwarki KonText

[102] Wytworzenie subkorpusu zmniejsza też niebezpieczeństwo błędu użytkownika na etapie ustalania zapytania; przy każdorazowym zaznaczaniu kryteriów zawężających liczbę tekstów łatwo bowiem o pomyłkę.

[103] KonText dostępny jest również bezpośrednio ze strony kontext.korpus.cz

Obraz 7. Interfejs wyszukiwarki KonText po wyszukaniu paralelnych konkordancji

Polsko-czeska i czesko-polska część InterCorpu

Materiał językowy wykorzystany w prezentowanych badaniach pochodzi przede wszystkim z polsko-czeskiej i czesko-polskiej części InterCorpu[104], z trzonu korpusu (*core*). Interfejs KonText umożliwia dodatkowe zawężenie wyszukiwania (kluczowe podczas badań nad przekładem) i tak podczas poszukiwań haseł polskich przeszukiwane mogą być wyłącznie teksty napisane w języku polskim i wiązane z tekstami w języku czeskim. Wytyczona w ten sposób część polsko-czeska zawiera 3 300 158 pozycji (w ramach trzonu korpusu, wersja 10 InterCorpu). Analogicznie poświadczenia czeskie wyszukiwane mogą być tylko w tekstach napisanych oryginalnie w języku czeskim wiązanych z tekstami w języku polskim[105]. Czesko-polska część zawiera 2 878 776 pozycji. Takie rozwiązanie nakłada jednak na użytkownika konieczność ustalania zawężeń za każdym razem podczas wyszukiwania poszczególnych haseł. Inną możliwością, zastosowaną również w większości badań przedstawionych w tym opracowaniu, jest tworzenie subkorpusów, o czym była mowa wcześniej[106].

[104] Bańczyk, Ł., Dybalska, R., Vavřín, M.: *Korpus InterCorp – polština, verze 10 z 1.12.2017*. Ústav České honárodního korpusu FF UK, Praha 2017. Dostęp: http://www.korpus.cz

[105] Objętość przeszukiwanych tekstów możemy zawęzić poprzez doprecyzowanie, które metadane te teksty mają charakteryzować. Dzieje się to w momencie wpisywania zapytania (Kaczmarska i Rosen 2014B: 215).

[106] Pierwsze badania wykonywane były przy użyciu wyszukiwarek Park i NoSkech Engine; po pojawieniu się możliwości tworzenia subkorpusów zostały powtórzone na materiale zgromadzonym w wytworzonym specjalnie do tego celu subkorpusie.

Treq

Treq jest usługą automatycznego generowania słowników przekładowych na podstawie danych zebranych w korpusie równoległym InterCorp, dostępną ze strony Czeskiego Korpusu Narodowego – http://treq.korpus.cz (Vavřín i Rosen 2015; Škrabal i Vavřín 2017).

Pomysł automatycznego generowania listy ekwiwalentów narodził się podczas ustalania polskiego ekwiwalentu czeskiego czasownika *zdát se* w latach 2012–2013 (Kaczmarska 2012A, 2012B). Badanie ukazało, iż InterCorp oferuje kilkadziesiąt różnych odpowiedników tego czasownika; stwierdzenie tego było jednak możliwe dopiero po wyeksportowaniu wyników do dokumentu Excel, następnie manualnej analizie i ręcznym otagowaniu wszystkich poświadczeń, wreszcie posortowaniu, przefiltrowaniu i przeliczeniu wyników. Było to niezwykle praco- i czasochłonne i niemożliwym wydawało się powtórzenie całego procesu dla wszystkich interesujących autorkę czasowników. Jedynym rozwiązaniem, które się nasuwało, było zautomatyzowanie tego procesu. Autorka tego opracowania, wspólnie z dr. inż. Alexandrem Rosenem, podjęła próbę automatycznej ekstrakcji par ekwiwalentów z czesko-polskiego korpusu równoległego, wykorzystując narzędzie GIZA++[107]. Lista par utworzyła swoisty słownik wygenerowany metodą automatyczną[108]. Nie była to jednak metoda idealna, gdyż wymagała zaawansowanych znajomości procedur informatycznych i ponownie zajmowała wiele czasu. Poza tym aktualizacja tegoż słownika wymagałaby powtórzenia całej procedury od początku (na nowych danych). Za idealne rozwiązanie autorka przyjęła możliwość generowania takiej listy wprost z korpusu.

Pierwszy raz autorka wspomniała o tym podczas wykładu gościnnego w Czeskim Korpusie Narodowym 29.04.2014 r. Pomysł zakładał, iż w trakcie wyszukiwania w InterCorpie dwujęzycznych konkordancji na ekranie z wynikami pojawi się okienko – tabelka ze wszystkimi dostępnymi w korpusie ekwiwalentami, wraz z danymi o ich liczbie i udziale procentowym. W dyskusji wziął udział dr Michal

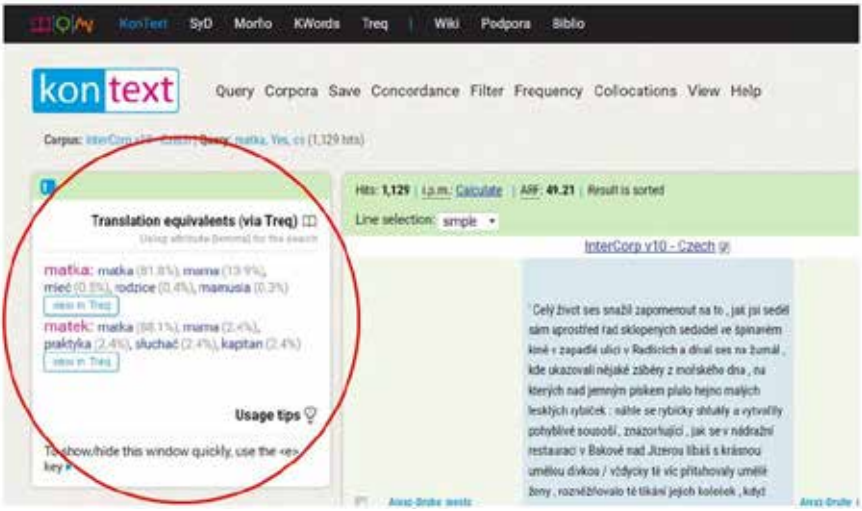
[107] GIZA++ jest programem wolno dostępnym, wytworzonym w celu statystycznego przekładu maszynowego, który zawiera moduł wiązania segmentów na poziomie wyrazu (word-to-word-alignment) – http://www.statmt.org/moses/giza/GIZA++.html (Och i Ney 2003; Kaczmarska i Rosen 2013).

[108] W trakcie ekstrakcji wykorzystane zostały teksty z korpusu InterCorp (wersja 6). Uwzględniona została jedynie beletrystyka (bez rozróżnienia na czeskie, polskie czy obce oryginały) – w tym teksty w języku czeskim: 11,885 milionów słów i teksty w języku polskim 11,860 milionów słów. Wiązanie segmentów (zdań) zostało ograniczone do 1:1; oznacza to, że były wykorzystane tylko segmenty wiązane 1:1, co zwiększyło wiarygodność wiązania na poziomie zdania i wyrazu. Metodą tą wyekstraktowano 8,651 milionów par lematów. Po złączeniu identycznych par (z zachowaniem informacji o ich liczbie) powstało 528 tysięcy haseł dwujęzycznych (Kaczmarska i Rosen 2013: 117–118). Podobną metodę wykorzystywał Karel Jirásek w swej pracy:*Využití paralelního korpusu InterCorp k získávání ekvivalentů pro chorvatsko-český slovník* (Jirásek 2011).

Křen, który zaproponował utworzenie specjalnej aplikacji, która wykorzystywałaby wiązanie *word-to-word*. Dopiero wyniki wyszukiwań poprzez tę aplikację odsyłałyby do poświadczeń korpusowych. Tak narodził się pomysł stworzenia aplikacji nazwanej później Treq (Vavřín i Rosen 2015), (Škrabal i Vavřín 2017).

Obraz 8. Treq – interfejs

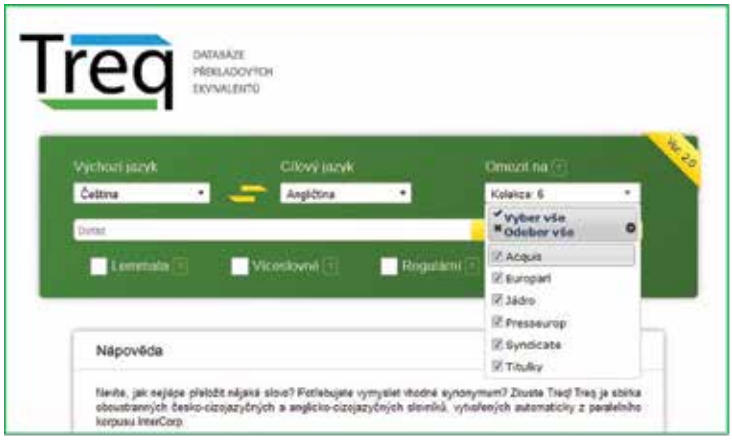
Okienko z proponowanymi ekwiwalentami z Treq, pojawiające się po wyświetleniu wyników wyszukiwania w korpusie InterCorp poprzez KonText, zaczęło się pojawiać od wersji 10 InterCorpu; na ilustracji zaznaczone czerwonym kółkiem:



Obraz 9. Okienko z ekwiwalentami z Treq pojawiające się w interfejsie KonText podczas wyświetlenia wyników wyszukiwania

Obraz 10. Treq – informacja bibliograficzna na stronie internetowej interfejsu

Treq jest narzędziem umożliwiającym generowanie list ekwiwalentów, w których jednym z języków musi być czeski albo angielski, a drugim – którykolwiek dostępny w tekstach InterCorpu. Oznacza to, że możemy wygenerować np. listę serbskich ekwiwalentów czeskiego słowa i listę czeskich ekwiwalentów serbskiego słowa (czy angielsko-japońską i japońsko-angielską), ale język polski (podobnie jak wszystkie inne poza czeskim i angielskim) wystąpi w kombinacji wyłącznie z językiem angielskim i czeskim.



Obraz 11. Treq – moduł wyszukiwania ekwiwalentów w języku czeskim

Obraz 12. Treq – moduł wyszukiwania w ekwiwalentów w języku angielskim

Korzystając z tego serwisu, należy jednak wziąć pod uwagę, iż narzędzie to stosuje metodę automatyczną, a wśród odpowiedników mogą znaleźć się sporadycznie również antonimy czy przypadkowe wyrazy. Nie jest więc to typowy słownik, zawierający wyłącznie trafione ekwiwalenty (Kaczmarska 2016; Kaczmarska i Rosen 2013).

Wykorzystując aplikację Treq, trzeba mieć również świadomość, iż poświadczenia w tym serwisie zliczane są z całego korpusu bez względu na język oryginału. Dużą zaletą jest natomiast automatyczna aktualizacja danych; Treq pobiera listę ekwiwalentów z najnowszej dostępnej wersji InterCorpu.

Studia przypadków

Badania pilotażowe z uwzględnieniem wybranych jednostek przeprowadzane były w latach 2011–2018. Poszczególne opracowania dotyczyły jednego lub kilku etapów opisanego w rozdziale 11 algorytmu i odnosiły się do jednej lub kilku czeskich jednostek leksykalnych. Kolejne rozdziały to przedstawienie tych badań[109]. Chociaż niektóre jednostki analizowane były kilkukrotnie, w różnych okresach czasu i przy pomocy różnych narzędzi, w prezentowanym opracowaniu wszystkie badania dotyczące tej samej jednostki stanowią odrębny rozdział, innymi słowy każda jednostka przedstawiona jest w oddzielnym rozdziale. Podrozdziały (w rozdziałach 6–10) mają jednak zbliżoną strukturę:

- definicja słownikowa,
- słownik dwujęzyczny,
- analiza korpusowa,
- wnioski.

Każdy z rozdziałów zawiera więc definicję słownikową[110], odniesienie do słownika dwujęzycznego[111] i analizę korpusową, przy czym analiza ta wraz z rozwojem metodologii z biegiem czasu ulega z jednej strony rozszerzeniu i wzbogaceniu, z drugiej strony – niektóre elementy analizy zostają zminimalizowane czy usunięte, co prezentują kolejne badania. Nie bez konsekwencji jest też fakt, iż czasowniki te – ze względu na swą specyfikę – często wymagają wdrożenia jakiejś specjalnej techniki. Z tych powodów części dotyczące analizy korpusowej są zbudowane podobnie, ale nie identycznie.

[109] Autorka starała się, by badania w tej części pracy przedstawione były w porządku chronologicznym, w jakim powstawały, i tym samym wskazywały na tok rozwoju metodologii. Nie w każdym przypadku zachowanie tego porządku było jednak możliwe.

[110] W tym opracowaniu wykorzystywany jest najczęściej *Slovník spisovného jazyka českého* (Havránek 1989) oraz jego wersja elektroniczna dostępna na stronie: http://ssjc.ujc.cas.cz/search.php?db=ssjc

[111] W częściach dotyczących słowników znajdą się odniesienia do słowników jednojęzycznych, dwujęzycznych, a także (kilkukrotnie) do słowników walencyjnych.

Ze względu na rozpiętość czasową badań autorka pracowała z różnymi wersjami korpusu InterCorp i w przypadku każdej analizy te dane podaje[112]. Wykorzystywała dostępne interfejsy: Park, NoSketch Engine i KonText (Kaczmarska i Rosen 2014B: 212–228). W trakcie badań wyszukiwane były pojedyncze leksemy (np. *mrzet*) – rodzaj zapytania *Lemma* i jednostki złożone (np. *být lito, mít rád*) – zaawansowany typ zapytania *CQL* lub *Fráze* (Hebal-Jezierska 2014: 62–64; Kaczmarska i Rosen 2014B: 220–225).

W kręgu zainteresowań proponowanego opracowania nie znalazły się derywaty omawianych jednostek; w wyszukiwaniach nie były uwzględniane więc jednostki np. *zatoužit* (*derywat* od *toužit*) czy *utrápit* (*derywat od trápit*); są traktowane jako odrębne jednostki i jako takie wymagają oddzielnej analizy. Poza rozważaniami znalazł się też problem aspektu czeskiego czasownika.

Wybierając do badań trzon InterCorpu (a dokładniej jego czesko-polską część), autorka zdawała sobie sprawę z jego ograniczeń, np.: mogącą się pojawić niską liczbę poświadczeń (Hebal-Jezierska, Kaczmarska i Rosen 2016: 43–62). Problemem przy szeroko zakrojonych badaniach (np. analizą dyskursu) mogłaby być w przypadku InterCorpu jego niereprezentatywność i niezrównoważenie (Górski i Łaziński 2012: 25–36). W przypadku badania tak specyficznej grupy czasowników autorce zależało jednak na zgromadzeniu jak największej liczby poświadczeń z uwzględnieniem kierunku przekładu oraz – w miarę możliwości – na tym, żeby punktem wyjścia były teksty oryginalne

Ograniczenia te znalazły swoje odzwierciedlenie w metodach badawczych. W językoznawstwie korpusowym wyróżnia się pięć głównych typów podejść:

- badania ilustrowane przez korpus (*corpus-illustrated* / *corpus-informed approach*), przy okazji których, jak pisze Łukasz Grabowski, korpus traktuje się jak informatora, rodzimego użytkownika języka, a dane korpusowe służą argumentacji (Grabowski 2015: 28)[113];
- oparte na korpusie (*corpus-based* / *corpus-supported approach*), w których poświadczenia korpusowe służą potwierdzeniu lub zanegowaniu wcześniej postawionej hipotezy (Hebal-Jezierska 2008: 11–12; Grabowski 2015: 28–29);
- badania sterowane korpusem (*corpus-driven approach*), w których z rezultatów badań nad poświadczeniami korpusowymi może być budowana nowa teoria i nie stawia się tu żadnych hipotez przed rozpoczęciem analizy (Hebal-Jezierska 2013: 25–26; Grabowski 2015: 28–29);

[112] W przypadku kilku jednostek konieczne było powtórzenie całości badania na danych z wersji 10 InterCorpu.

[113] Por. (Hebal-Jezierska 2008: 11–12; Górski 2012: 291–300).

- wywoływane czy indukowane korpusem (*corpus-induced approach*), polegają na wykorzystaniu danych korpusowych do prowadzonych na dużą skalę najczęściej automatycznych analiz ilościowych w celu pozyskania informacji dla komercyjnych zastosowań praktycznych (Grabowski 2015: 28–29)[114];
- badania wykorzystujące jednoczesną analizę danych korpusowych oraz pozakorpusowych, np. ankiety, wywiady (Baker 2010: 8; Hebal-Jezierska 2013: 25); ich angielska nazwa – *corpus-assisted analysis* – jest jednak niefortunna i bywa utożsamiana z badaniami z wykorzystaniem korpusu w ogóle.

W przedstawianym opracowaniu nie było możliwe typowe badanie *corpus driven* (McEnery, Xiao i Tono 2006: 9–11; Hebal-Jezierska 2008: 11–12; Grabowski 2012: 35–38), które wyłącznie na podstawie danych korpusowych umożliwiłoby określenie znaczenia danej jednostki i wskazanie ekwiwalentu konkretnego czasownika w danym kontekście (Kaczmarska 2015A; Kaczmarska, Rosen, Hana i Hladká 2015). Nie zostało też wybrane tradycyjne podejście *corpus based* (McEnery, Xiao i Tono: 2006: 9–11; Hebal-Jezierska 2008: 12). W opracowaniu tym zostało zastosowane podejście *mieszane*[115], wykorzystujące korpus w celu:

- potwierdzenia i zilustrowania założenia przyjętego wcześniej na podstawie definicji słownikowych,
- modyfikacji tej hipotezy w wyniku przeprowadzenia dalszego etapu analizy,
- rozszerzenia i konkretyzacji założenia – tu inwentarza potencjalnych ekwiwalentów po zakończeniu badania.

Zastosowana tu metoda łączy więc trzy klasyczne podejścia: *corpus-illustrated*, *corpus based* i *corpus driven*, sporadycznie wykorzystuje również metodę *corpus-assisted analysis*.

Czeskie jednostki poddane analizie w tym opracowaniu zostały wybrane spośród kilkudziesięciu wyekscerpowanych ze słowników leksemów, odnoszących się do emocji, uczuć i przeżyć. Grupa ta jest niezwykle barwna i należą do niej jednostki związane z różnymi stanami psychicznymi. Są wśród nich m.in. jednostki odnoszące się do :

- strachu (*bát se, děsit, hrozit se, lekat se, mít strach, obávat se, polekat, polekat se, strachovat se, trápit, trápit se, vystrašit, vystrašit se, zděsit se, zneklidnit, zneklidnit se, znepokoit, znepokoit se*);

[114] Por. (Lee 2008: 88; Grabowski i Hebal-Jezierska 2016: 68–70).

[115] Szerzej na temat różnych podejść w badaniach korpusowych (również mieszanych) w artykule Łukasza Grabowskiego i Mileny Hebal-Jezierskiej (Grabowski i Hebal-Jezierska 2016).

- złości (*dopálovat*, *dopálovat se*, *dráždit*, *hněvat*, *hněvat se, iritovat, popuzovat*, *rozčilit*, *rozčilit se*, *rozhorlit*, *rozhorlit se*, *rozhořčit*, *rozhořčit se*, *rozlítit*, *rozlítit se*, *rozvzteklit*, *rozvzteklit se*, *rozzuřit*, *rozzuřit se*, *rozmrzet*, *rozmrzet se*, *vztekat*, *vztekat se*, *zlobit*, *zlobit se*, *zuřit*);
- wstydu (*hanbit se*, *ostýchát se*, *stydět se*, *zahanbit*, *zmást*, *zmást se*, *zastydět se*);
- wzruszenia (*dojmout*, *dojmout se*, *rozlítostnit*, *rozlítostnit se*, *slitovat se*, *vzrušit*, *vzrušit se*);
- smutku (*bolet*, *mrzet*, *soucítit*, *stýskat se*, *trápit*, *trápit se*, *trpět*, *zarmoutit*, *zarmoutit se*, *zoufat si*, *želet*);
- negatywnych emocji (*nenávidět*, *nudit*, *nudit se*, *omrzet*, *omrzet se*, *pohrdat*, *štítit se*, *závidět*, *zklamat*, *zklamat se*);
- pozytywnych emocji (*líbitse*, *milovat*, *mítrád*, *nadchnout*, *nadchnout se*, *obšťastnit*, *okouzlit*, *radovat se*, *těšit*, *těšit se*, *uchvátit*, *zbožňovat*);
- zaskoczenia (*divit se*, *překvapit*, *šokovat*);
- pragnienia (*chtít se*, *postrádat*, *potřebovat*, *toužit*);
- innych emocji, przeżyć i uczuć [(*snít*, *zdát se*), (*přivyknout*, *zvyknout si*), (*bláznit*, *blbnout*, *šílet*, *zbláznit se*)].

Leksemy te zostały poddane wstępnej analizie i na jej podstawie wybranych zostało osiem jednostek najbardziej, według autorki, interesujących pod względem przekładu czesko-polskiego[116].

Kolejno zostaną przedstawione jednostki: *zdát se* (Przypadek 1), *trápit* (Przypadek 2), *soucítit* (Przypadek 3), *závidět* i *žárlit* (Przypadek 4) oraz *mít rád* i *milovat* (Przypadek 5). Po omówieniu projektu algorytmu ułatwiającego ustalanie ekwiwalentów (w rozdziale 11) zostanie przedstawiony także czasownik *toužit* (Przypadek 6) w rozdziale 12.

[116] Wybrane jednostki można by równie dobrze nazwać najbardziej kłopotliwymi. Czasowników sprawiających trudności w procesie przekładu czesko-polskiego jest oczywiście więcej; niektóre z nich zostały opisane w odrębnych pracach np. *postrádat* (Kaczmarska i Rosen 2018, 2019) czy *mrzet* (Kaczmarska 2014B; Kaczmarska i Rosen 2018).

Przypadek 1 – *zdát se*

Pierwszą analizowaną jednostką był czasownik *zdát se*[117] (Kaczmarska 2012A, 2012B), a pierwszym etapem – sprawdzenie, jak definiowany jest ten czasownik w źródłach jednojęzycznych (Havránek i inni 1989; Slovník českých synonym a antonym 2007) oraz jakie polskie ekwiwalenty ma ta jednostka w słowniku dwujęzycznym (Siatkowski i Basaj 2002). Kolejnym krokiem była ekscerpcja ręczna czeskiej i polskiej wersji tego samego tekstu[118], następnie wyszukiwanie ekwiwalentów w korpusie równoległym InterCorp.

Definicja słownikowa

Słowniki języka czeskiego odnotowują kilka znaczeń tego czasownika, prezentując przy tym dość szeroką gamę jego synonimów: 1) 'jevit se ve snu', 'mít sen'; 2) 'mít dojem', 'mít pocit', 'domnívat se'; 3) 'budit dojem', 'jevit se', 'připadat', 'vypadat', 'projevovat se', 'ukazovat se'; 4) 'líbit se', 'zamlouvat se' (Havránek i inni 1989).

Słownik dwujęzyczny

Choć jednostka *zdát se* została odnotowana w słownikach języka czeskiego jako czasownik wieloznaczny, popularny słownik dwujęzyczny czesko-polski nie

[117] Angielski odpowiednik tego czasownika – *seem* – był analizowany między innymi przez Stiga Johanssona (Johansson 2007B) w ujęciu kontrastywnym (angielski – norweski).

[118] Ekscerpcja manualna została przeprowadzona na podstawie trzech powieści (wersja czeska: K. Čapek, *Hordubal – Povětroň – Obyčejní život,* Praha, 1958; tłumaczenie na język polski: K. Čapek, *Hordubal – Meteor – Zwyczajne życie,* tłum. Paweł Hulka-Laskowski, Warszawa, 1977). W trakcie analizy ustalono między innymi, iż tłumacz użył tylko trzech czasowników jako ekwiwalentów jednostki *zdát se*: *śnić się*, *zdawać się*, *wydawać się*. Ten etap analizy nie został powtórzony przy żadnym kolejnym badaniu i nie zostanie tu bliżej omówiony; szczegółowe wyniki odnaleźć można innej publikacji (Kaczmarska 2012: 248–249).

odzwierciedla tego stanu i podaje odpowiedniki w języku polskim bez odniesienia się do czeskich znaczeń[119]: *śnić się*, *wydawać się*, *zdawać się*, *podobać się* (Siatkowski i Basaj 2002: 1006–1007). Klasyczny słownik dwujęzyczny (papierowy) nie ma też możliwości zaprezentowania wszystkich ekwiwalentów wraz z kontekstami. Stąd decyzja o ekscerpcji materiału z korpusu równoległego.

Analiza korpusowa

Czasownik *zdát se* był wyszukiwany w wersji 4 korpusu równoległego InterCorp[120]. Analizie poddane zostały wszystkie formy wspomnianego czasownika oraz jego odpowiedniki w języku polskim. Zapytanie zostało zadane w jedynej wówczas dostępnej dla InterCorpu – wyszukiwarce Park.

intercorp_cs (539727 tokens)	intercorp_en (653376 tokens)	intercorp_pl (551869 tokens)
Kwic		
Možná si šel zaběhat a zanedbal obvyklý postup .	show context Maybe he 'd gone jogging and neglected the routine .	show context Niewykluczone , iż wyruszył na trasę biegu , zapomniawszy dopełnić rutynowych czynności .
O půlnoci si šla lehnout .	show context She went to bed at midnight .	show context Położyła się spać o północy .
Měl jediný pokoj ; dveře se daly zamknout zvenčí a okna nešla otevřít .	show context His was the only room with doors which could be locked from the- outside and windows that would n't open .	show context Patricka umieszczono w niewielkiej salce na samym końcu bocznego skrzydła szpitala . Pojedyncze okno nie dało się otworzyć , drzwi były zamykane od zewnątrz .
Pak pozval své nové kamarády z řad novinářů , aby ho doprovodili pěšky k soudu , až půjde podat žalobu .	show context Then he invited his new pals in the press corps to stroll with him down the sidewalk as he went to file it .	show context A po jej zakończeniu wyruszył wraz z całym tłumem dziennikarzy do kancelarii , aby złożyć pozew .
Paulo také věděl o dost podezřelých individuích , kteří se potulovali u nich v ulici a sledovali ho , když šel nakupovat nebo jel do své pracovny na Pontificia Universidade Católica .	show context Paulo was also tired of the shady little men lurking around his street and following him as he walked to the market or drove to his office at the Pontificia Universidade Catolica .	show context Miał także dosyć ciągłego towarzystwa jakichś podejrzanych typków , którzy nieodmiennie go obserwowali , czy to wychodził na zakupy , czy jechał do swojego gabinetu w Pontificia Universidade Católica .
Oblékla si parku a bosa se šla projít po pláži , jednu ruku v kapse a v druhé hrnek kávy .	show context She put on a parka and went for a walk on the beach , barefoot and bare-legged , with one hand in a pocket and the other holding her coffee .	show context Eva włożyła kurtkę i wybrała się na spacer wzdłuż plaży . Trzymając w jednym ręku kubeczek z kawą , przechadzała się powoli , boso , wzdłuż granicy fal .
Export: xls1 , xls2		View: horizontal

Obraz 13. Interfejs Park – fragment wyniku wyszukiwania

[119] Wśród polskich czasowników zaproponowanych przez słownik dwujęzyczny (Siatkowski i Basaj 2002) znajdują się odpowiedniki dwóch znaczeń wymienionych przez czeski słownik (Havránek i inni 1989): 'jevit se ve snu' – *śnić się*, 'líbit se' – *podobać się*. Trudno jednak wskazać odrębne ekwiwalenty czasowników reprezentujących znaczenie 2 i 3: 'mít dojem', 'mít pocit', 'domnívat se', 'budit dojem', 'jevit se', 'připadat', 'vypadat'; w cytowanym słowniku ich odpowiednikami są jednostki *zdawać się* i *wydawać się*.

[120] Korpus InterCorp – wersja 4; szczegółowe informacje dostępne na stronie internetowej: http://wiki.korpus.cz/doku.php/cnk:intercorp:verze4

Wyszukiwanie wskazało 978 poświadczeń czasownika *zdát se* w tłumaczeniu na język polski[121]. Znalezione jednostki zostały podzielone na kilkanaście grup na podstawie znaczenia[122]. W ramach każdej grupy pojawiają się jednostki synonimiczne. Wyekscerpowane segmenty obejmują też przykłady, które nie zawierają ekwiwalentów czasownika *zdát se*; przykłady te ilustrują błędy w segmentacji i wiązaniu tekstów w korpusie oraz błędy lematyzacji[123]. Wśród poświadczeń pojawiają się też rekordy, w których doszło do opuszczenia omawianego czasownika w procesie tłumaczenia[124]. W wyszukiwaniach znaleziono 45 przypadków (~4,6%), w których czasownik *zdát se* został w tłumaczeniu opuszczony; przykłady dotyczyły różnych konfiguracji składniowych, wskazanie więc reguł, jakimi mogli się kierować tłumacze, nie było możliwe[125].

Zdálo se však, že pilné zaznamenávání zpráv dvorských, vojenských i politických bylo příčinou, že na knihy asketické začala sedat jemná dosud vrstva prachu. JDB1993cz

Jednakże pilne notowanie spraw dworskich, wojskowych i politycznych sprawiło, iż dzieła ascetyczne pokryła leciutka warstwa kurzu. JDZ1959pl

Nebe se zdálo překrásné. JDB1993cz
Niebo było przepiękne. JDZ1959pl

[121] Dane na dzień 18 marca 2012. W wersji 7 byłoby tych poświadczeń 1384 (w identycznym typie zapytania), a w wersji 10–1560. Autorka zdecydowała się jednak na przedstawienie badania, które przeprowadziła z wykorzystaniem wyszukiwarki Park ze względu na jej archiwalny już charakter.

[122] Czasowniki *wydawać się* i *zdawać się* są najczęściej pojawiającymi się ekwiwalentami polskimi; stanowią razem prawie 72% wystąpień i są postrzegane jako synonimiczne. Ponieważ badanie dotyczy możliwości przekładu, należy wziąć pod uwagę, w jakich sytuacjach wybierany jest czasownik *wydawać się*, a w jakich – *zdawać się*. Do tego można wykorzystać program Sketch Engine i wygenerowane przez niego profile kolokacyjne Word Sketches (Kaczmarska 2015A: 140–143). W analizie uwzględnić należy wszystkie elementy łączące się z tymi czasownikami: podmiot zdarzenia – komu się *wydaje* / *zdaje*, a także jak ów podmiot zdarzenia jest wyrażony [imię, rzeczownik pospolity, zaimek osobowy (tu ważną rolę może odegrać zaimek pierwszoosobowy *ja*)], jakiego typu jest obiekt zdarzenia, warstwę stylistyczną tekstu oraz jego rodzaj (partia dialogowa, narracja). Wynik takiej analizy mógłby się okazać szczególnie ważny dla tłumaczy, obcokrajowców uczących się języka polskiego oraz dla leksykografów (Kaczmarska 2012: 250).

[123] W 20 przykładach (~2,04%) został odnaleziony homonim lub też fragment polski, mający być ekwiwalentem tekstu czeskiego, był zupełnie innym tekstem. W tabelce przedstawiającej wyniki oznaczono tę grupę jako „błąd korpusu". O problemach podczas wyszukiwania w korpusie także: (Hebal-Jezierska 2008: 22–24).

[124] Na temat opuszczeń w procesie tłumaczenia także: (Hejwowski 2009: 83–104).

[125] Stig Johansson w swoim badaniu również spotkał się z problemem opuszczania analizowanego angielskiego czasownika *seem* w tłumaczeniu na norweski:

What sparked this study was the observation that the English verb seem sometimes seems to disappear without a trace in translations into Norwegian and likewise may be added, seemingly without any motivation, by English translators in rendering Norwegian original texts. This happens too often for it to be rejected as a mere translation error. What seems to be a problem? (Johansson 2007B: 117).

W swojej analizie Johansson poszukiwał prawidłowości zanikania i pojawiania się norweskich ekwiwalentów czasownika *seem* (Johansson 2007B: 117–138).

Panna již byla vedle něho a zdálo se, že by ho mohla vidět, kdyby chtěl. JDB1993cz
Panna stała obok niego i mogła go ujrzeć, gdyby chciał. JDZ1959pl

wydawać się

W 509 przypadkach jako ekwiwalent pojawił się czasownik ***wydawać się*** (stanowi to ~52,05% wystąpień), np.:

Zdála se stvořena pro němý podiv. KCTNA1982cz
*Góra ćwieczków **wydawała się** stworzoną dla niemego podziwu.* KCFA1971pl

G.H. Bondy klečí v jakési úžasné, jasné blaženosti, chtěl by křičet a zpívat, zdá se mu, že slyší šum nesmírných a nesčíslných křídel. KCTNA1982cz
*G.H. Bondy klęczy w stanie jakiegoś straszliwego uczucia błogości, chciałby krzyczeć i śpiewać, **wydaje** mu **się**, że słyszy szum niezmiernych i niezliczonych skrzydeł.* KCFA1971pl

A tu jsem užasl: hra, jíž jsem se bavil, nabyla náhle podivně skutečných rysů; zdálo se mi totiž, že tu ženu, co se nade mnou v zrcadle sklání, znám. MKZ1991cz
*Wpadłem w osłupienie, i zabawa, którą sobie urządziłem, nabrała nagle rysów autentyczności; **wydało** mi **się** bowiem, że tę dziewczynę, która pochyla się nade mną w lustrze, znam.* MKZ1999pl

zdawać się

W 190 poświadczeniach polskim ekwiwalentem jest czasownik ***zdawać się*** (stanowi to ~19,45% wystąpień):

Panna mu děkovala a zdálo se, že též s úsměvem. JDB1993cz
*Panna odpowiedziała, **zdaje się**, również uśmiechem.* JDZ1959pl

Zdá se, že v Berlíně se nevzdávají naděje na smírnou likvidaci celého nedorozumění. KCTNA1982cz
***Zdaje się**, iż w Berlinie nie tracą nadziei co do pokojowego zlikwidowania całego nieporozumienia.* KCFA1971pl

Zdálo se mu, že našel správný tón. MKVNR1997cz
***Zdawało** mu **się**, że przyjął właściwy ton.* MKWP2001pl

Czasowniki *wydawać się* i *zdawać się* skupiają razem ~71,48% wystąpień i są postrzegane jako jednostki synonimiczne. Obydwie jednostki (jako ekwiwalenty) podaje też wspomniany wcześniej słownik czesko-polski (Siatkowski i Basaj 2002) oraz znalazły się w wynikach ekscerpcji manualnej (Kaczmarska 2012: 248–249).

mieć wrażenie

W 50 przypadkach w tekstach przekładu pojawiła się jednostka ***mieć wrażenie*** lub jednostki synonimiczne, np. *wyobrażać sobie*; stanowi to ~5,11% wszystkich wyekscerpowanych wystąpień.

Hluk se valil a zdálo se, že dunění a řinčení tvrdě pobízelo (…) JDB1993cz

*Huk przewalał się ulicami i **miało** się **wrażenie**, jakby dudnienie i brzęk nawoływały twardo* (…) JDZ1959pl

Pak se zdálo, jako by na stránce knihy se zrcadlily rozpaky čtoucích, jako by věty byly nesrozumitelny a temny. JDB1993cz

*Potem **miał wrażenie**, jakby na stronicach książki odbiło się zakłopotanie czytelnika, jakby zdania były nie dość zrozumiałe, nie dość jasne.* JDZ1959pl

Tak se to zdá vám. JSLE1996cz
*Tak to pan **sobie wyobraża**.* JSLE1965pl

Jednostka *mieć wrażenie* jest bliska czasownikom *wydawać się* i *zdawać się*; może być użyta zarówno w wersji osobowej, jak i nieosobowej[126]. Podobnie jak te czasowniki wnosi element niepewności czy dystansu.

wyglądać

Jako ekwiwalent występuje również czasownik ***wyglądać*** i jednostki względem niego synonimiczne (*widać, widzieć*); dzieje się tak w 39 przykładach (~4%), np.:

V ulicích bylo úzko, domy se zdály na spadnutí a lidé jako by se měli v noci vystěhovat i s kočkami. JDB1993cz

*Wąskie uliczki robiły przygnębiające wrażenie, domy **wyglądały**, jakby miały lada chwila runąć, a ludzie - jakby się mieli wyprowadzać w nocy razem z kotami.* JDZ1959pl

Zdá se, že vám dva roky politického školení, soudruzi, nestačily! JSTP1990cz
***Widać**, że dwa lata szkolenia politycznego to dla was za mało, towarzysze!* JSBC2004pl

Jakub poslouchal Škrétovy řeči a nezdálo se mu, že by v nich nalézal mnoho zajímavého. MKVNR1997cz
*Jakub słuchał oracji Slamy, ale nie **widział** w niej nic szczególnie ciekawego.* MKWP2001pl

Jednostka *wyglądać* (czes. *vypadat*) odwołuje się do wzrokowych możliwości postrzegania; bogate znaczenie czasownika *zdát se* jest w tym przypadku

[126] Czeskim odpowiednikiem frazy *mieć wrażenie* jest jednostka *mít dojem*.

spłycone. Czasowniki *widzieć* (czes. *vidět*) i *widać* (czes. *je vidět*) pozornie również wskazują wyłącznie na percepcję wzrokową, można jednak przypuszczać, iż użyte tu zostały w znaczeniu bliskim jednostkom *zdawać sobie sprawę*, czy *uznać*. Słownik języka polskiego odnotowuje bowiem inne znaczenia tych czasowników: *widać* oznacza również 'widocznie', 'zapewne', 'chyba', 'okazuje się' (Szymczak 1995: 647), a czasownik *widzieć* – między innymi 'zdawać sobie sprawę", 'uświadamiać sobie', 'przekonywać się o czymś', 'stwierdzać coś', 'zauważać' czy 'mieć określony pogląd na coś', 'pojmować coś w pewien sposób', 'patrzeć na coś' (Szymczak 1995: 649–650)[127].

śnić się

W 32 przypadkach jako ekwiwalent pojawia się czasownik ***śnić się*** *(przyśnić się, przywidzieć się, sen)*; stanowi to ~3,28% wystąpień. *Śnić się* to jedno z podstawowych znaczeń analizowanego czeskiego czasownika. Wystąpienie tego ekwiwalentu (oraz jego synonimów) można uznać za oczekiwane.

V noci bylo slyšet z kuchyně křik, Růžence se zdálo, že ji honí čerti. LFVPTS2003cz
W nocy było słychać z kuchni krzyk, Rużence ***śniło się****, że ją gonią diabły.* LFWNNS1970pl

Ten šramot se nám musil zdát. LFVPTS2003cz
A ten chrobot to nam ***się*** *chyba* ***przyśnił****.* LFWNNS1970pl

"To se ti nezdálo", informoval mne o přestávce Jaromír. MVVDVC1994cz
– Nie ***przywidziało*** *ci* ***się*** *– poinformował mnie na przerwie Jaromir.* MVWDWC2005pl

Ta druha půle Tomáše, to byla dívka, o které se mu zdálo. MKNLB1985cz
Ta druga połowa Tomasza to była dziewczyna ze ***snu****.* MKNLB1992pl

myśleć

W 21 przypadkach (~2,15% wystąpień) jako ekwiwalenty pojawiły się czasowniki myślenia – ***myśleć*** (*mniemać, podejrzewać, pomyśleć, rozumieć, sądzić, uświadamiać sobie, uważać, uznać, moim zdaniem*).

Zdálo se mu, že ho něco roztrhne. JDB1993cz
Myślał*, że mu serce pęknie.* JDZ1959pl

[127] Znaczenia te odnoszą się do kondycji intelektualnej – szerzej o tym w dalszym podrozdziale.

Ale ta odvaha nebyla tak velká, jak se mu zdálo, protože výraz, který zvolil, byla něžná zdrobnělina nebo poetická perifráze. MKN1993cz

Nie okazał, prawdę mówiąc, takiej odwagi, jak mniemał, bowiem użyte wyrażenie było czułym zdrobnieniem czy też poetycką peryfrazą. MKN2008pl

Zdálo se, že loutnistka se s ním už nikdy nechce vidět, a je jí nepříjemné mu to říci přímo. MKN1993cz

***Podejrzewał**, że lutnistka postanowiła więcej się z nim nie spotykać, lecz nie ma śmiałości mu tego powiedzieć*. MKN2008pl

Franzovi se najednou zdá, že mu jeho polozapomenutý přítel telefonoval na základě jejího tajného pokynu. MKNLB1985cz

*Franz nagle **rozumie**, jego na wpół zapomniany przyjaciel zadzwonił do niego pod wpływem jakiegoś tajemnego polecenia*. MKNLB1992pl

Pak jsem uklidil rychle láhev, srovnal podušky na gauči, a když se mi zdálo, že jsou všechny stopy uklizeny, zabořil jsem se do křesla u okna (…). MKZ1991cz

*Potem szybko uprzątnąłem butelkę, wyrównałem poduszki na tapczanie, a kiedy **uznałem**, że wszystkie ślady zostały zatarte, zagłębiłem się w fotelu koło okna* (…). MKZ1999pl

Zdá se mi, že moc pracuješ. MKN1993cz
***Moim zdaniem** za dużo pracujesz*. MKN2008pl

Użycie wymienionych w tym punkcie jednostek jako ekwiwalentów analizowanego czeskiego czasownika sprawia, iż podkreśleniu ulega aspekt intelektualny zdarzenia. Zawarty w czasowniku *zdát se* element modalny zostaje zachowany[128].

czuć

Czasowniki typu *czuć* (*poczuć, doznać uczucia, mieć uczucie*) pojawiają się 9 razy (~0,92%).

Laura si tiskla každou rukou ke každému prsu jednu kytici a zdálo se jí, že stojí na scéně. MKN1993cz

*Z bukiecikami przyciśniętymi do piersi Laura **czuła się** jak na scenie operowej.* MKN2008pl

Zdálo se mu, že by potřeboval sedět přikován v hladomorně (…). JDB1993cz
***Miał uczucie**, że powinien siedzieć przykuty w wieży* (…). JDZ1959pl

[128] Więcej o komponencie modalnym w dalszej części tego rozdziału.

Znenáhla však míč začal divoce lítat po hřišti a mně se zdálo, jako by mi horečka přestávala vadit. LFVPTS2003cz

Nagle jednak piłka zaczęła latać po boisku jak oszalała i ***poczułem****, że gorączka przestała mi przeszkadzać.* LFWNNS1970pl

Ulehli jsme pak s Vlastou do vysoko nastlané postele a zdálo se mi, že je to sama moudrá nekonečnost lidského pokolení, erá nás vzala do měkké náruče. MKZ1991cz

Położyli śmy się z Vlastą do zasłanego wysoko łoża i ***doznałem uczucia****, że to sama mądra nieskończoność ludzkich pokoleń bierze nas w miękki uścisk swych ramion.* MKZ1999pl

Wybierając te jednostki, tłumacze decydują się na wzmocnienie elementu emocjonalnego znaczenia czasownika *zdát se*.

czuć strach

W poświadczeniach znalazły się też dwie jednostki nawiązujące do poprzedniej grupy związanej z czasownikiem *czuć*, ale odnoszące się do uczucia strachu. Stanowią zaledwie ~0,2% wszystkich wystąpień.

Zahlédl pot na jejím čele a zdálo se mu, že omdlí. JDB1993cz
Dojrzał pot na jej czole i ***zląkł się****, że zemdleje.* JDZ1959pl

Zdálo se, že mundšenk dlouho nejde. DB1993cz
Niepokoił się*, że podczaszy tak długo nie idzie.* JDZ1959pl

podobać się

W korpusie pojawia się również czasownik *podobać się* (oraz formacja *być zadowolonym*); są to 2 przypadki (~0,2%). To kolejne (po czasownikach *wydawać się* / *zdawać się* i *śnić się*) znaczenie analizowanego czeskiego czasownika, które prezentował również słownik (Siatkowski i Basaj 2002: 1006–1007). Wystąpienie tego ekwiwalentu (oraz jego synonimów) można więc traktować jako oczekiwane.

Venca se potil , jak ho Fonda nutil, a nutil ho tak , že si musel doladťovač trombónu postrčit skoro o decimetr, až už mu to dál nešlo, ale Fondovi se to pořád nezdálo. JSZ1964cz
Wacek aż się spocił, tak go Fonda piłował, a piłował go tak, że Wacek musiał stroik puzonu przesunąć prawie o dziesięć centymetrów, aż już dalej nie szło, ale Fonda ciągle nie ***był zadowolony****.* JST1970pl

„Co se ti na tom nezdá?" říká Malá útočně. MVZOL2003cz
– Co ci się w tym nie ***podoba?*** *– mówi Mała agresywnie.* MVZOM2007pl

okazać się

W dwóch przypadkach pojawił się też czasownik wynikać, a w jednym – *okazać się*.
Cesta byla delší, než se na mapce zdálo, a navíc jsme zabloudili. MVPNK2003cz
Droga okazała się dłuższa, niż to ***wynikało*** *z mapy, do tego udało nam się zabłądzić.* MVSNK2007pl

Po letech se náhle zdá, že ta anonymita byla zemi nebezpečná. MKNLB1985cz
Po latach ***okazuje się*** *nagle, że ta chwilowa anonimowość była dla kraju niebezpieczna.* MKNLB1992pl

Czasowniki *wynikać* i *okazywać się* wnoszą element obiektywizmu i nieosobowości. Tłumacze mogli posłużyć się tymi jednostkami tylko w nielicznych przypadkach, w których *zdát se* występuje bez dopełnienia osobowego.

Wykładniki modalności jako ekwiwalenty

W 56 (~5,72%) przypadkach pojawiły się inne odpowiedniki analizowanego *zdát se*, które często zawierają w swoim znaczeniu informację o przekonaniu mówiącego, czy zaistnienie zdarzenia (o którym mowa) jest prawdopodobne, czy też nie. Wiąże się to z modalnością epistemiczną[129]. Większość wyrażeń poniżej wypunktowanych (np. *chyba*, *najwyraźniej*, *pewnie*, *prawdopodobnie*) to typowe jej wykładniki[130].

[129] Nadawca informuje odbiorcę o swojej postawie wobec treści zdarzenia, które może być związane zarówno z przeszłością, jak i teraźniejszością czy przyszłością. Hipotezy mogą opierać się na własnych doświadczeniach nadawcy, przekonaniach, czy różnych źródłach informacji. Zob. hasło *modalność* (Polański 1999: 371). Szerzej na temat modalności (w tym epistemicznej) i jej klasyfikacjach również w ujęciu konfrontatywnym: (Adamec 1973: 141–147; Bauer i Grepl 1970; Boniecka 1976; Roszko 1993; Rytel 1982; Svoboda 1973: 232–241; Wróbel 1991: 260–270).

[130] Modalność epistemiczna może być wyrażana za pomocą:
1) wykładników leksykalnych: a) czasowników modalnych takich jak *musieć*, *mieć* czy innych jak *wiedzieć*, *wierzyć*, *sądzić* itp.; b) połączenia spójki z przymiotnikiem: *jestem pewien, że…*, *jestem przekonany*, *że…* itp.; c) przysłówków: *podobno*, *wątpliwe*, *na pewno*, *może*, *oczywiście*; d) wyrażeń przyimkowo-rzeczownikowych typu przysłówkowego np.: *z pewnością*, *bez wątpienia*; e) nieosobowych konstrukcji bezokolicznikowych np.: *jest możliwe*, *jest pewne*, *jest jasne*, f) zwrotów typu: *istnieje prawdopodobieństwo*, *istnieją wątpliwości*; g) związków frazeologicznych np.: *Prawdę powiedziawszy, bardzo się cieszymy, że przyjechałeś.*; 2) wykładników gramatycznych: a) czasu gramatycznego futurum, b) trybu rozkazującego, c) środków składniowych (szyk zdania), d) nominalizacji, e) uniwerbizacji; 3) wykładników pozajęzykowych: a) tembr głosu, b) akcent zdaniowy. Szerzej na temat środków wyrażania modalności epistemicznej w języku polskim i czeskim: (Rytel 1982; Kaczmarska, Stefaniak i Susfał 2011: 217–229; Tutak 2003).

chyba

Pak se zdálo, že tam v té jizbě žije či aspoň žila někdy žena. JDB1993cz
*A tam **chyba** była jakaś kobieta.* JDZ1959pl

najwyraźniej

Popošel rychle o krok a napjatě očekával, co zvíře udělá, ale nezdálo se, že by krab jeho přítomnost zaznamenal. MVUZ2001cz
Szybko zrobił krok do przodu i z napięciem oczekiwał, jak zachowa się zwierzę, ale krab najwyraźniej nie zauważył jego obecności. MVW2008pl

pewnie (pewien)

Zdálo se nám, že již nepotkáme zlého hada. JDB1993cz
*Byliśmy prawie **pewni**, że już nie spotkamy złego węża.* JDZ1959pl

prawdopodobnie

…zdála se být vyvolenou některého švédského hrdiny. JDB1993cz
*…była **prawdopodobnie** wybranką któregoś z szwedzkich bohaterów.* JDZ1959pl

jakby (11) i *pozornie* (1)

Výkřik se prodlužoval jako blesk po obloze a zdálo se, že stokráte za sebou se odrážel ode všech zdí. JDB1993cz
*Okrzyk przedłużał się jak błyskawica na niebie, **jakby** stokrotnie odbity od wszystkich ścian kościoła.* JDZ1959pl

Zdálo se, že omeškalé čarodějnice a ďáblové tryskem přijeli oběma stranám na pomoc. JDB1993cz
*Tak **jakby** spóźnione czarownice i diabły pędem nadjechały obu stronom na pomoc.* JDZ1959pl

Zdálo se, že lež se vtělila v nenadálou děsnou pravdu a přichází se chechtat. JDB1993cz
*Tak **jakby** kłamstwo wcieliło się niespodziewanie w przerażającą prawdę i przyszło z nich szydzić.* JDZ1959pl

Najednou s ním zase někdo mluvil jako s lékařem a on cítil, jak k němu z dálky promlouvá jeho někdejší život se svou příjemnou pravidelností, s vyšetřováním nemocných, s jejich pohledem plným důvěry, kterému se zdál nevěnovat pozornost, ale který ho ve skutečnosti těšil a po němž prahl. MKNLB1985cz

Naraz ktoś znów z nim rozmawiał jak z lekarzem, a on czuł, jak z oddali przemawia do niego jego dawne życie ze swą przyjemną regularnością, z badaniem chorych, z ich

pełnym zaufania spojrzeniem, na które ***pozornie*** *nie zwracał uwagi, ale które w rzeczywistości sprawiało mu satysfakcję i na które czekał.* MKNLB1992pl

Zabiegiem zastosowanym w tym przypadku przez tłumaczy było przeniesienie znaczenia czasownika *zdát se* na odpowiedni przysłówek. Uniknięto w ten sposób nagromadzenia czasowników w krótkim zdaniu (**wydawał się nie zwracać uwagi*) oraz zmiany stylu.

Wymienione wykładniki modalności mogą być interpretowane również jako wskaźniki ewidencjalności[131], czyli kategorii, która specyfikuje, na czym opiera się przekazywana informacja oraz jakie jest jej źródło[132] (Hirschová 2017; Guentchéva 2018). Milada Hirschová pisze, iż sam czasownik *zdát se* (podobnie jak *jevit se*) może być uznany za wskaźnik ewidencjalności wyrażający wnioskowanie oparte na percepcji; dzięki temu mogłoby się udać wyjaśnić użycie niektórych ekwiwalentów czasownika *zdát se* z domeny percepcji (*wyglądać*) czy poznawczej (*myśleć*)[133].

Zmiana konstrukcji

W 11 przypadkach doszło do zmiany konstrukcji przekładanego zdania:

Tereze se to zdálo nevysvětlitelné. MKNLB1985cz
Teresa nie umiała sobie tego wytłumaczyć. MKNLB1992pl

Zaprezentowany przykład, typowy dla tego podpunktu, dotyczy sytuacji, w której zaistniała konieczność zmiany struktury nie tylko ze względu na omawiany czasownik *zdát se*, ale także na inny komponent zdania – jednostkę *nevysvětlitelné*. W języku czeskim jednostki typu *nevysvětlitelné* są często używane; to przymiotniki wyrażające niemożność wykonania czynności nazwanej przez podstawę,[134] np. *nenapravitelná, nenahraditelné, nesnesitelný* (Kaczmarska 2002: 96). W języku polskim również obserwuje się takie konstrukcje (np. *niewytłumaczalny*,

[131] *Jedną z charakterystycznych funkcji wykładników ewidencjalnych jest sygnalizowanie faktu, że mówiący opiera swoje twierdzenie na wypowiedziach innych osób; formy o takiej funkcji w terminologii angielskiej noszą nazwę 'quotative' albo 'hearsay marker'* (Holvoet 2011:78).

[132] W przeciwieństwie do modalności epistemicznej ewidencjalność reprezentuje domenę poznawczą, oznaczającą źródło informacji (Ivanová 2011: 145–154; Kyseľová i Ivanová 2013: 202–206). Obie kategorie są blisko spokrewnione (Blakemore 1999: 141–145; Kyseľová i Ivanová 2013: 207–210; Wiemer 2018: 85–108).

[133] Na słowacki czasownik *zdať sa* jako możliwy wskaźnik ewidencjalności zwracają też uwagę Miroslava Kyseľová i Martina Ivanová (Kyseľová i Ivanová 2013: 204).

[134] Wykładnikami modalności są sufiksy: *-elný, -elná, -elné*, a wykładnikiem negacji prefiks *ne-*.

niezmywalny), jednak pojawiają się one zdecydowanie rzadziej. Być może dlatego tłumaczka nie zdecydowała się użyć tego słowa w swoim przekładzie[135].

Wnioski

Korpus paralelny wskazał 978 wystąpień różnych form czasownika *zdát se* w tłumaczeniu na język polski. Procentowy udział tych poświadczeń, podzielonych na grupy (wraz z wartościami bezwzględnymi) ukazuje tabela 1.

Analiza danych zebranych na podstawie korpusu równoległego ukazała szereg możliwości rozumienia i przekładu czasownika *zdát se*, który jawi się jako wieloznaczny i często stawiający przed tłumaczem zadanie trudne nie tylko pod względem leksykalnym, ale również stylistycznym. Autor przekładu mierzy się z problemem oddania w tekście wszystkich elementów wiązki semantycznej, którą reprezentuje czasownik *zdát se* (percepcja wzrokowa, aspekt intelektualny, element emocji, komponent obiektywizmu i nieosobowości, rama modalna), a także przed zagadnieniem dopasowania wybranej jednostki do stylu przekładanego tekstu. Często tłumacz wybiera jednostkę, która akcentuje któryś z odcieni znaczeniowych; najczęściej jest to spowodowane kontekstem; wpływ może mieć także struktura predykatowo-argumentowa danej jednostki.

Przeważająca liczba poświadczeń (w sumie ~71,48% wszystkich znalezionych wystąpień) wykorzystuje najbardziej oczywiste ekwiwalenty – *wydawać się* i *zdawać się*; bliska im znaczeniowo, acz odmienna stylistycznie, jednostka *mieć wrażenie* pojawiła się tylko w ~5,11% wystąpień. Autorzy pozostałych tłumaczeń sięgają po jednostki akcentujące rożne odcienie znaczeniowe analizowanego czasownika *zdát se*, np. percepcję wzrokową (*wyglądać*, *widzieć* i *widać*), aspekt intelektualny (*myśleć, mniemać, podejrzewać, pomyśleć, rozumieć, sądzić, uświadamiać sobie, uważać, uznać, moim zdaniem*), element emocji (*czuć, poczuć, doznać uczucia, mieć uczucie*), komponent obiektywizmu i nieosobowości (*wynikać* i *okazywać się*), czy ramę modalną (*chyba*, *najwyraźniej*, *pewnie, prawdopodobnie*). W korpusie pojawiły się również poświadczenia pozostałych znaczeń czasownika *zdát se*: *śnić się* (*przyśnić się, przywidzieć się*) i *podobać się* (*być zadowolonym*).

[135] Zwłaszcza w kontekście fraza ta wydaje się mało naturalna: *Teresie wydawało się to niewytłumaczalne*.

Upłynął miesiąc i inżynier nie zjawił się. Teresa nie umiała sobie tego wytłumaczyć. Zawiedziona tęsknota zeszła na drugi plan i zastąpił ją niepokój: dlaczego nie przyszedł? MKNLB1992pl

Tabela 1. Ekwiwalenty polskie czasownika *zdát se* – rozkład procentowy

Ekwiwalenty w języku polskim	Liczba wystąpień	Udział procentowy
wydawać się	509	**52,05%**
zdawać się	190	**19,43%**
mieć wrażenie	49	**5,11%**
wyobrażać sobie	1	
wyglądać	34	**4%**
widać	4	
widzieć	1	
sen / śnić się	29	**3,28%**
przyśnić się	1	
przywidzieć się	1	
sen	1	
myśleć	5	**2,15%**
mniemać	1	
podejrzewać	1	
pomyśleć	2	
rozumieć	1	
sądzić	4	**2,15%**
uświadamiać sobie	1	
uważać	4	
uznać	1	
moim zdaniem	1	
czuć	3	**0,92%**
poczuć	2	
doznać uczucia	3	
mieć uczucie	1	
niepokoić się	1	**0,20%**
lękać się	1	
podobać się	1	**0,20%**
być zadowolonym	1	
wynikać	2	**0,3%**
okazywać się	1	
inne	45	**5,72%**
jakby	11	
błąd korpusu	20	**2,04%**
pominięte w tłumaczeniu	45	**4,6%**
RAZEM	978	**100%**

Badanie potwierdziło, że na wybór tłumaczenia wpływa kontekst; nie ukazało jednak, czy pojawia się jakiś formalny element decydujący o wyborze ekwiwalentu. Ręczna analiza poświadczeń pozwoliła natomiast na dostrzeżenie, iż InterCorp oferuje całą gamę możliwych ekwiwalentów również w ramach grup reprezentujących określone odcienie semantyczne (jak w przypadku powyżej wymienionych). Wtedy właśnie pojawił się pomysł stworzenia narzędzia, nazwanego później Treq, generującego listę możliwych ekwiwalentów na podstawie korpusu InterCorp[136].

[136] Treq opisywany był w rozdziale 4.

Przypadek 2 – *trápit*

Jednostka *trápit* była wcześniej przedmiotem analizy w dwóch badaniach (Kaczmarska 2014B: 48–49, 55; Kaczmarska 2018)[137]. W pierwszym opracowaniu na podstawie analizy korpusowej zauważono, iż InterCorp oferuje dwadzieścia siedem odpowiedników analizowanego czasownika, które stanowią szereg bliskoznacznych wobec siebie jednostek[138]. Stwierdzono też, iż tłumacz korzystający z takiego zestawu potencjalnych ekwiwalentów ma do wyboru cały wachlarz możliwości w różnych odcieniach stylu, brzmienia czy składni. Późniejsze badanie pogłębia analizę, koncentrując się głównie na wymaganiach walencyjnych czasownika *trápit* i jego polskich najczęstszych ekwiwalentów wskazanych przez słownik dwujęzyczny i korpus równoległy.

Niniejszemu badaniu poddane zostaną struktury semantyczno-składniowe czasownika *trápit* i jego odpowiedników zaproponowanych przez klasyczny słownik dwujęzyczny oraz znalezionych wśród poświadczeń wyszukanych w wersji 10 InterCorpu (Čermák i Rosen 2012; Kaczmarska i Rosen 2014B; Rosen 2016; Rosen, Vavřín i Zasina 2018). Analiza tych jednostek może odpowiedzieć na pytania, czy i ewentualnie w jaki sposób wymagania walencyjne wpływają na przekład oraz czy tłumacze starają się zachować w tekście docelowym identyczną strukturę semantyczno-składniową[139].

W badaniu wykorzystane zostaną słowniki walencyjne obu języków, a także narzędzie korpusowe Treq[140].

[137] Do badań tych autorka odnosi się w tym rozdziale.

[138] Nie wszystkie te jednostki są czasownikami.

[139] Problem ten wiąże się z dodatkowymi pytaniami – czy przekład powiela ramę walencyjną czasownika z tekstu oryginalnego, czy zachowuje wszystkie obiekty, w jakich przypadkach może dojść do zmiany całej struktury, gdyby tego wymagał ekwiwalent w języku docelowym.

[140] Opisana w rozdziale czwartym usługa automatycznego generowania różnych słowników na podstawie danych zebranych w korpusie równoległym InterCorp (Rosen i Vavřín 2015; Škrabal i Vavřín 2017). Dostęp ze strony Czeskiego Korpusu Narodowego http://treq.korpus.cz/

Definicja słownikowa

Słownik języka czeskiego (*Slovník spisovného jazyka českého* – dalej SSJC) definiuje ten czasownik jako:

1) ‘působit někomu tělesnou nebo duševní bolest, trýzeň, obtíž, nesnáz; soužit, sužovat, týrat, trýznit, mučit’; 2) ‘obtěžovat’[141] (Havránek i inni 1989)[142]. Źródło to przedstawia również charakterystykę walencyjną czasownika *trápit*:

trápit **koho** *(rzadziej* **– co***)* **// koho čím / s čím // že**… // **jak**…

Z informacji zawartej w słowniku jednojęzycznym (Havránek i inni, 1989) można wnieść, iż czeski czasownik *trápit* otwiera miejsce dla pozycji o wartości Experiencer[143], który na powierzchni wyrażony jest frazą nominalną w bierniku. Dodatkowo występuje argument o wartości Stimulus[144], który może na powierzchni przybierać kilka postaci: fraza nominalna w mianowniku, fraza nominalna w narzędniku (przyimkowa i bezprzyimkowa), fraza zdaniowa.

Trápí mě kašel.
Trápí ho výčitky svědomí.
Co tě trápí?
Nejvíc mě trápí, že jsi odjela.
Trápí ho, jak to nejlépe udělat.

Jak ukazują przykłady, w konstrukcjach z tym czasownikiem może pojawić się również pośredni sprawca – Factor[145] wyrażony frazą nominalną w mianowniku[146]; pojawienie się jednak tej pozycji powoduje, iż Stimulus musi być wyrażony przyimkową lub bezprzyimkową frazą nominalną w narzędniku:

Alena ho trápila svými rozmary.
Proč maminka to dítě trápí s učením?

---

[141] 1) powodować ból fizyczny lub psychiczny, udrękę, kłopoty, trudności; dręczyć, nękać, torturować, drażnić, męczyć; 2) kłopotać (tłumaczenie własne).

[142] http://ssjc.ujc.cas.cz/search.php?hledej=Hledat&heslo=tr%C3%A1piti&sti=EMPTY&where=hesla&hsubstr=no

[143] W przykładach – Experiencer – frazy podkreślone. Terminy nazywające role semantyczne zaczerpnięte zostały ze słownika walencyjnego *Walenty* (Przepiórkowski, Hajnicz, Andrzejczuk, Patejuk, Woliński 2017; Hajnicz i Nitoń 2017). Por. (Kaczmarska 2001).

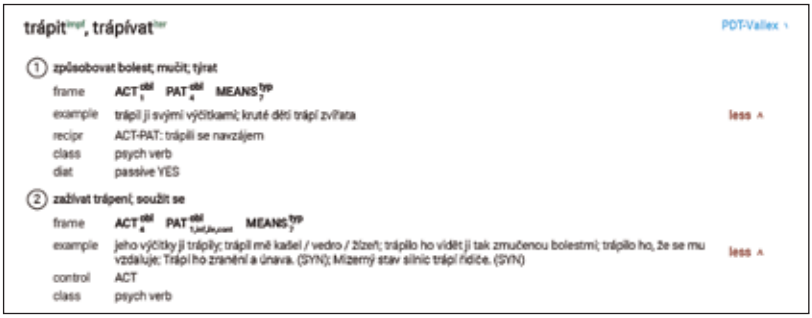
[144] W przykładach – Stimulus – frazy podkreślone faliście.

[145] *Factor – przyczynia się w sposób aktywny, acz pośredni, do zajścia sytuacji, ma na nią wpływ* (Hajnicz i Nitoń 2017: 31). Por. zależności pomiędzy rolami Agens – Experiencer – Źródło (Kaczmarska 2001: 177–179).

[146] W przykładach – Factor – frazy podkreślone podwójnie.

Słowniki walencyjne języka czeskiego

Definicja i charakterystyka składniowa zamieszczona w SSJC (Havránek i inni, 1989) jest zgodna z opisem w słowniku VALLEX[147].



Obraz 14. Czasownik *trápit* w słowniku VALLEX

VALLEX prezentuje jednak dalszą możliwość wyrażenia pozycji Stimulus; w strukturze powierzchniowej może pojawić się na tym miejscu również fraza bezokolicznikowa:

Trápilo <u>ho</u> vidět ji tak zmučenou bolestmi.

Kolejne rozszerzenie możliwości powierzchniowej realizacji pozycji o wartości Stimulus przynosi definicja w słowniku PDT-Vallex[148]:

Obraz 15. Czasownik *trápit* w słowniku PDT-Vallex

[147] Dostęp online: http://ufal.mff.cuni.cz/vallex/3.0/#/lexeme/trapi1/1

[148] PDT-Vallex oprócz skrótowej charakterystyki walencyjnej podaje 33 opisane przykłady z korpusu (na obrazku widoczne tylko nieliczne). Dostęp online: http://lindat.mff.cuni.cz/services/PDT-Vallex/PDT-Vallex.html

Stimulus może być też realizowany jako fraza zdaniowa ze spójnikiem *zda*:

Máme lidi, které trápí, zda bude v noci pršet.

Słownik dwujęzyczny

Czasownik *trápit* w najpopularniejszym tradycyjnym polskim słowniku czesko-polskim (Siatkowski i Basaj 2002) jest przekładany jako: *męczyć, dręczyć, gnębić, trapić*; wymagania walencyjne w tym słowniku nie są podane ani przy czasowniku czeskim, ani przy jego polskich odpowiednikach. Jednostki te mają wymagania walencyjne zbliżone do wymagań semantyczno-składniowych czasownika *trápit*, co zostało ustalone na podstawie czeskich słowników walencyjnych VALLEX i PDT-Vallex oraz polskiego słownika *Walenty*[149].

Polskie ekwiwalenty czeskiego czasownika ***trápit*** InterCorp wersja 10, czeskie oryginały (trzon korpusu)	
ekwiwalent polski	liczba poświadczeń
dręczyć	13
gnębić	12
męczyć	4
dokuczać	3
cierpieć	2
gryźć	2
martwić	2
martwić się	2
nękać	2
trapić	2
cierpienie	1
doprowadzić do wściekłości	1
męczyć się	1
napawać niepokojem	1
niepokoić	1
przejmować się	1
zadręczać	1
zamęczać	1
zmartwienie	1
	53

[149] *Walenty* ukazuje szereg fraz pytajno-zależnych cp(int) łączących się z tymi czasownikami, realizowanych dzięki spójnikom: *co, czemu, czy, czyj, dlaczego, dokąd, gdzie, ile, jak, jaki, kiedy, kto, którędy, który, odkąd, skąd* oraz archaiczne *azali, azaliż, jakoby* (Hajnicz i Nitoń 2017; Przepiórkowski, Hajnicz, Andrzejczuk, Patejuk, Woliński 2017). Dostęp online – http://walenty.ipipan.waw.pl

Słowniki nie informują jednak, jak ramy walencyjne czasownika *trápit* funkcjonują w przekładach na język polski; to zjawisko przybliży analiza korpusowa poświadczeń tego czasownika.

Analiza korpusowa

Materiał do badań wyekscerpowany został z korpusu równoległego InterCorp poprzez wyszukiwarkę KonText. Poświadczenia pochodzą z czesko-polskiej części tego korpusu (Bańczyk, Dybalska i Vavřín 2017), zawężonej do tekstów napisanych oryginalnie w języku czeskim.

Tabela 2. Polskie ekwiwalenty czeskiego czasownika *trápit* w korpusie InterCorp (wersja 10)

Polskie ekwiwalenty czeskiego czasownika ***trápit*** InterCorp wersja 10, czeskie oryginały (trzon korpusu)		
ekwiwalent polski	liczba poświadczeń	
	czasownikowe	nieczasownikowe
dręczyć	13	
gnębić	12	
męczyć	4	
dokuczać	3	
cierpieć	2	
gryźć	2	
martwić	2	
martwić się	2	
nękać	2	
trapić	2	
cierpienie		1
doprowadzić do wściekłości		1
męczyć się	1	
napawać niepokojem		1
niepokoić	1	
przejmować się	1	
zadręczać	1	
zamęczać	1	
zmartwienie		1
	49	**4**

Znalezione zostały zaledwie 53 poświadczenia czasownika *trápit* wyrównanych z odpowiadającymi segmentami w języku polskim[150]. Inwentarz potencjalnych ekwiwalentów wyszukiwanego czasownika jest szerszy od tego, który proponuje klasyczny słownik.

W zestawieniu pojawiają się wskazane przez klasyczny słownik odpowiedniki: *męczyć, dręczyć, gnębić, trapić* (Siatkowski i Basaj 2002): trzy z nich na najwyższych pozycjach, czwarty – *trapić* – uplasowany niżej. Jak wynika z tabeli, reprezentują one niewielką liczbę poświadczeń, uniemożliwiającą przeprowadzenie głębszej analizy gramatycznej czy stylistycznej. Na tej podstawie można jednak wysnuć pewne wnioski co do przekładu wyszukiwanego czasownika na język polski. Dla porównania warto pokazać, ile razy pojawiają się interesujące nas jednostki w całym korpusie i jaka jest ich względna częstość występowania[151] (skrót i.p.m. oznacza liczbę jednostek na milion – ang. *instances per million*)[152].

Tabela 3. Częstość występowania polskich ekwiwalentów w korpusie InterCorp (wersja 10)

wybrane czasowniki → część korpusu / liczba słów ↓	***dręczyć***	***gnębić***	***męczyć***	***trapić***
częstość całkowita **teksty pisane oryginalnie w języku polskim** (CORE) / 23 238 000	65	26	187	12
częstość względna **i.p.m. (CORE)**	**2,8**	**1,12**	**8,05**	**0,52**
częstość całkowita **wszystkie testy w języku polskim** (CORE + kolekcje) / 85 176 000	1673	379	1874	413
częstość względna **i.p.m. (CORE + kolekcje)**	**19,64**	**4,45**	**22.01**	**4,85**

Ręczna analiza wyekscerpowanych poświadczeń umożliwia precyzyjną ich analizę oraz wskazanie konkretnych znaczeń zarówno czeskiego czasownika, jak i jego potencjalnych ekwiwalentów. W dalszym badaniu wykorzystane więc będą również definicje ze słownika języka polskiego .

[150] KonText wskazał 61 poświadczeń. 8 z nich zostało odrzuconych z powodu błędu wyszukiwania; w tym 3 stanowiły imiesłowowe formy wyszukiwanego czasownika. Szerzej o problemach podczas wyszukiwania w korpusie (Hebal-Jezierska 2008: 22–24).

[151] Analizowane czasowniki wykazują dużo wyższą częstość występowania w przekładach na język polski niż w tekstach oryginalnie pisanych w języku polskim. Problem ten nie jest jednak tematem tego opracowania.

[152] Dane liczbowe do tabeli pochodzą z materiałów Czeskiego Korpusu Narodowego – dostęp online: http://wiki.korpus.cz/doku.php/cnk:intercorp:verze10#obsah_korpusu.

trápit – dręczyć

Czasownik *dręczyć* jest najczęściej wybieranym ekwiwalentem *trápit*, co zgadza się z wyliczoną częstością względną. *Słownik języka polskiego* definiuje go jako 'zadawać komuś cierpienie fizyczne lub psychiczne'[153]. Jak ukazują przykłady z korpusu InterCorp, czasownik *dręczyć* jako przekład czeskiej jednostki *trápit* wykorzystywany jest do wyrażenia:

- cierpienia fizycznego spowodowanego przez osobę:

Nerad zbytečně manšaft trápím. JHODV1996cz
Nie lubię dręczyć ludzi niepotrzebnie. JHPDW1957pl

- cierpienia psychicznego zadawanego przez osobę:

Věrnost byla stejně jen taková záminka, s kterou se holky oháněly, aby místo jen jednoho mohly trápit několik chlapů najednou. JSZ1964cz

Wierność to zresztą był tylko pretekst, przy pomocy którego dziewczyny się opędzały, aby zamiast jednego mogły dręczyć kilku chłopców naraz. JST1970pl

A ani tady mu nedala pokoj a ještě nad umírajícím Kareninem ho trápila tajným podezíráním. MKNLB1985cz

I tutaj również nie dała mu spokoju i dręczyła go podejrzeniami nawet nad umierającym Kareninem. MKNLB1992pl

- cierpienia psychicznego powodowanego przez problemy, obawy, sytuację, myśli:

Jediné, co ho v této chvíli trápí je snad jen obava, aby mu marmeláda neukápla na noviny. MVZOL2003cz

Jedyne, co go w tej chwili dręczy, to obawa, żeby marmolada nie skapnęła mu na gazetę. MVZOM2007pl

A teď se mi tedy svěřuje s tím, co ji trápí a tíží: ano, jednala špatně, když se rozhodla, že se nebude se mnou vídat (…). MKZ1991cz

I teraz oto zwierza mi się z tego, co ją dręczy i co jej ciąży: tak, źle postąpiła, kiedy postanowiła się ze mną nie widywać (…). MKZ1999pl

We wszystkich przykładach z czasownikiem *dręczyć* przekład powiela schematy walencyjne struktur czeskich. Nie dzieje się to jednak kosztem poprawności polskich zdań.

[153] Dostęp online: https://sjp.pwn.pl/szukaj/dr%C4%99czy%C4%87.html

trápit – gnębić

Słownik języka polskiego wskazuje na dwa znaczenia czasownika *gnębić*: 1) ‘wywoływać niepokój lub przygnębienie’, 2) ‘prześladować kogoś’. W poświadczeniach wyekscerpowanych z korpusu InterCorp czasownik ten wyraża sytuacje:

- jakieś fakty czy myśli wywołują u kogoś niepokój czy przygnębienie (często zwerbalizowaniu tego, co wywołuje te emocje, służy wyraz „coś”):

Občas mě něco trápí stále, pokrčil jsem rameny, hledě stranou, je to různé a je to všechno dohromady. LFVPTS2003cz

Często mnie coś gnębi - wzruszyłem ramionami, patrząc w bok - są różne przyczyny, i w ogóle wszystko razem. LFWNNS1970pl

Už jednou jsi zbledl, když jsem se ptal, máš - li kamarády… tak se mi zdá, že tě něco trápí… LFVPTS2003cz

Już raz zbladłeś, gdy spytałem, czy masz kolegów … tak mi się zdaje, że coś cię gnębi… LFWNNS1970pl

Bylo však jasné, že se s nacistickým barbarstvím neztotožnil a že ho hluboce trápí… LFOMB1980cz

Było jednak jasne, że Romuald z barbarzyństwem nazistów nie utożsamia się i że ono go gnębi… LFOMB1988pl

- wyrzuty sumienia wywołują niepokój czy przygnębienie:

A navíc mě trápí výčitky, že jsem tu dobrou ženu neoprávněně podezíral ze zlomyslnosti. VRR1944cz

I w dodatku gnębią mnie wyrzuty sumienia, że posądzałem tę dobrą kobietę o złośliwość. VRK1954pl

- ktoś jest dręczony przez inną osobę:

Už jsem ho nechtěl dál trápit. MVPNK2003cz

Nie chciałem już gnębić go dalej. MVSNK2007pl

Sotva přestal katolíky pražské trápit pan hrabě z Thurnu, který jako generální komisař krále švédského pro Království české se stejnou dychtivostí i neumělostí hájil zájmy svého nového pána, nové pronásledování začal pan Vavřinec z Hofkirchu, který následoval způsobů Monsieura de Marradas. JDB1993cz

Ledwie pan hrabia z Thurn, który jako generalny komisarz szwedzkiego króla dla Królestwa Czeskiego z jednakowym zapałem i szczerością bronił interesów swego nowego pana, przestał gnębić katolików praskich, zaczęły się nowe prześladowania, które podjął pan Wawrzyniec Hoffkirch, naśladujący sposoby Monsieur de Marradasa. JDZ1959pl

Również w przypadku tego czasownika rama walencyjna czasownika oryginału jest identyczna ze schematem czasownika w przekładzie. Wśród przykładów nie ma ani jednego przypadku, w którym wybór tego czasownika wydawałby się nieodpowiedni.

trápit – męczyć

W *Słowniku języka polskiego* odnaleźć można cztery znaczenia czasownika *męczyć*[154]: 1) 'sprawiać ból i cierpienie', 2) 'powodować zmęczenie, utratę sił', 3) 'naprzykrzać się komuś czymś lub mieć stale pretensję o coś', 4) 'robić coś z wysiłkiem, zbyt długo'[155]. Przykłady odnalezione w czesko-polskiej części korpusu InterCorp są nośnikiem następujących sytuacji:

- ból i cierpienie sprawia choroba:

Trápí ho velké bolesti a rychle slábne. VRR1944cz

Męczą go silne bóle i bardzo osłabł. VRK1954pl

- ból i cierpienie sprawia człowiek:

(…) *tíhu lesního zvířátka, které očima prosí, aby je člověk netrápil a pustil do jeho lesní svobody, aby je ze své moci nechal odejít.* JSLE1996cz

(…) *wyraz schwytanego zwierzątka leśnego, co błaga oczyma, aby człowiek go nie męczył i zwrócił mu wolność lasu, by pozwolił mu odejść z zasięgu swojej władzy.* JSLE1965pl

- zmęczenie i utratę sił powoduje czyjaś cecha:

Její odpor mne nijak nedojímal, zdál se mi nesmyslný, křivdivý, nespravedlivý; trápil mne, nerozuměl jsem mu. MKZ1991cz

Jej upór zupełnie mnie nie wzruszał, wydawał mi się głupi, krzywdzący, niesprawiedliwy; męczył mnie, nie rozumiałem go. MKZ1999pl

Również w przypadku tego czasownika wymogi walencyjne polskiej jednostki pokrywają się z wymogami czeskiej jednostki. Przy przekładzie i tutaj nie dochodziło do żadnych zmian struktury. Wprawdzie nieliczne przykłady zgromadzone w tym punkcie nie wyczerpują możliwości znaczeń czasownika *męczyć*, nakreślają jednak, które z jego znaczeń jest tożsame ze znaczeniem analizowanego czasownika *trápit*.

trápit – trapić

Ostatnim analizowanym polskim ekwiwalentem jest niezwykle bliski czeskiej jednostce w pisowni i wymowie czasownik *trapić*. Warto zwrócić szczególną

[154] Dostęp online: https://sjp.pwn.pl/szukaj/m%C4%99czy%C4%87.html

[155] Znaczenie 3 i 4 oznaczone są jako potoczne.

uwagę na relację tych czasowników; czes. *trápit* i pol. *trapić* nie są „fałszywymi przyjaciółmi" tłumacza[156], a jednak mimo morfologicznego podobieństwa i zbliżonego brzmienia, a także podobnych wymagań dystrybucyjnych nie są wymieniane jako najbliższe ekwiwalenty (Siatkowski i Basaj 2002; Oliva 1995). Przyczyn tego faktu może być co najmniej kilka – różnica częstotliwości użycia, stylu, odmienny rejestr. Mogłoby się wydawać, iż to doskonały ekwiwalent; badanie ukazuje jednak, iż tłumacze wybierają go bardzo rzadko. *Słownik języka polskiego* prezentuje dwa znaczenia tego czasownika: 1) 'niepokoić, martwić', 2) 'nękać, gnębić'[157]. Jedyne dwa, wyekscerpowane jako ekwiwalenty czeskiego czasownika, przykłady wiążą się z sytuacją, kiedy

- kogoś trapi coś:

A zároveň si ihned namítal, že žádné zdržení jeho odjezdu, ať již o den či o léta, by stejně nenapravilo to, co ho teď trápí (…) MKVNR1997cz

A jednocześnie natychmiast kontrargumentował sam sobie, że żadne odroczenie wyjazdu, czy to o dzień, czy o lata, i tak nie naprawiło by tego, co go teraz ***trapi*** (…) MKWP2001pl

Nejsi, Lakmé, v poslední době nějak smutná, řekl nejistě a vzal kleštičky, kterými předtím zatloukal hřebík do zdi, „nebolí tě něco, netrápí, nemáš nějakou starost (…) LFSM1983cz

Nie jesteś przypadkiem, Lakme, jakaś smutna ostatnio - powiedział niepewnie i wziął obcążki, którymi wcześniej przybijał gwóźdź - nie boli cię coś, nie ***trapi****, nie masz jakichś zmartwień?* (…) LFPZ1979pl

Również w przypadku tego czasownika powielenie schematu walencyjnego czasownika czeskiego odbywa się bez uszczerbku dla stylu i poprawności wypowiedzi w języku polskim. Przyczyn niepopularności tego czasownika jako ekwiwalentu możemy upatrywać jednak w odmiennym rejestrze obu czasowników.

Wartym uwagi jest też fakt, iż czasownik *trapić* pojawia się często w kontekście chorób – zarówno realnych – fizycznych, jak i ujętych metaforycznie w przekładach innych czeskich jednostek na język polski[158]:

[156] Obszernie na temat pozornej ekwiwalencji językowej oraz relacjach językowych czesko-polskich m.in. w pracach Zofii Teresy Orłoś (Orłoś 1980, 2003, 2004), Zofii Tarajło-Lipowskiej (Tarajło-Lipowska 2000), Edwarda Lotko (Lotko 1992, 1997).

[157] Dostęp online: https://sjp.pwn.pl/szukaj/trapi%C4%87.html

[158] Poniższe przykłady pochodzą również z czesko-polskiej części korpusu równoległego InterCorp (Bańczyk, Dybalska i Vavřín 2017). Jako *trapić* zostały przełożone na polski m.in. jednostki: *trápit se, stižený, mít záchvat, týrat*.

Ještě sluší dodati, že plukovník stížený dnou měl ve své kanceláři všechno velice demokraticky zařízeno, když právě neměl záchvatu. JHODV1996cz

A dodać jeszcze trzeba, że pułkownik, dotknięty podagrą, wszystko w swojej kancelarii miał urządzone bardzo demokratycznie i był demokratą, jeśli akurat nie ***trapiła*** *go podagra.* JHPDW1957pl

Říkala to, jako kdyby byla plavec stižený kurdějema, který po tříměsíčním bloudění spatřil nějaký ostrov. JSPS1991cz

Powiedziała to tak, jakby była ***trapionym*** *przez szkorbut żeglarzem, który po trzymiesięcznym błądzeniu po falach oceanu zobaczył jakąś wyspę.* JSFS1999pl

Osud se mezitím opřel do jeho poddaných a týral je nemocemi. MKNLB1985cz

Los tymczasem uwziął się na jego poddanych i ***trapił*** *ich różnymi chorobami.* MKNLB1992pl

trápit a inne możliwości

Potencjalnymi ekwiwalentami znalezionymi w poświadczeniach korpusowych są także inne jednostki, których frekwencja była jednak bardzo niska. Były to m.in.:

- *dokuczać*, *sprawić przykrość*;
- *martwić*, *nękać*, *zamęczać*, *niepokoić*, *torturować*, *doprowadzić do wściekłości*, *gryźć*;
- *męka*, *udręka*, *cierpienie*, *zmartwienie*;
- *przejmować się*[159].

Jednostki *dokuczać*[160] i *sprawiać przykrość* posiadają podobne ramy walencyjne; różnią się one jednak od ramy czeskiego czasownika *trápit*. W przypadku

[159] Jednostki te zostały przedstawione w podpunktach grupujących leksemy o tych samych wymogach walencyjnych.

[160] *Walenty* wskazuje na dwa znaczenia *dokuczać*. *Dokuczać*$_1$ dotyczy procesu niecelowego – impulsem (Stimulus) jest tu dolegliwość, bolesność, sytuacja, forma psychiczna czy doznanie. *Dokuczać*$_2$ to czynność celowa – sprawcą (Initiator – podkreślenie linią wykropkowaną) jest człowiek; w tym przypadku *Walenty* definiuje przeżywającego jako Theme (rola podlegająca jakiemuś procesowi czy działaniu lub znajdująca się w jakimś stanie; często podlegająca zmianie w wyniku zajścia sytuacji – podkreślony linią przerywaną). Powód dokuczania *Walenty* nazywa Condition (w przykładach podkreślenie – kropka i kreska); jest to człon, który określa okoliczności, bez zaistnienia których nie doszłoby do zajścia danego zdarzenia, informuje o przyczynie lub warunkach zaistnienia sytuacji (Hajnicz i Nitoń 2017: 30–31), np.:

Dziadek Krystiana, pan Antoni, ma nadzieję, że inne dzieci nie będą dokuczać wnukowi z powodu koloru skóry.

Przyznaje tylko, że sporadycznie w podstawówce koledzy mu dokuczali, że dobre oceny z kartkówki zawdzięcza ojcu.

obu polskich jednostek pojawia się pozycja Experiencera, który na powierzchni wyrażony jest frazą nominalną w celowniku oraz pozycja o wartości Stimulus wyrażona na powierzchni jako fraza nominalna w mianowniku.

Co prawda ostatnio dokuczał mi trochę nadwerężony mięsień, ale z każdym dniem czuję się lepiej…[161]

Dokucza mi jedynie senność…

Kloss stwierdził ze zdziwieniem, że jej reakcja sprawiła mu przykrość.

W przypadku jednostek *dokuczać* i *sprawiać przykrość* (jako ekwiwalentów czeskiego czasownika *trápit*) walencja w przekładzie nie byłaby zachowana:

Neudělala nic, co by vás mělo ***trápit***. MVPNK2003cz

Nie zrobiła niczego, co mogłoby sprawić panu ***przykrość***. MVSNK2007pl

Jednostki *martwić, nękać, zamęczać, niepokoić, torturować, doprowadzić do wściekłości, gryźć* otwierają miejsce dla pozycji o wartości Experiencer, która jest realizowana na powierzchni poprzez frazę nominalną w bierniku. W ramie walencyjnej pojawia się też Stimulus, który może – podobnie jak w przypadku czeskiego czasownika – przybierać na powierzchni postać frazy nominalnej w mianowniku lub frazy zdaniowej[162].

Nas martwiło, że nie podał nazwy ani dokładniejszego położenia swej wyspy.

Jego żywe wspomnienie i mroczna symbolika nękały mnie i niepokoiły.

Najmniejsze niepowodzenia doprowadzały go do wściekłości.

Przy czym jednak cały czas gryzie ją sumienie.

Film oglądałem więcej niż 20 lat temu, ale teraz gryzie mnie, jak się on nazywał.

W konstrukcjach z tymi czasownikami (oprócz jednostki *gryźć*) może pojawić się również Factor wyrażony frazą nominalną w mianowniku; pojawienie się tej pozycji powoduje, iż Stimulus zaczyna być wyrażany frazą nominalną w narzędniku:

Maroko torturowało ją różami, pomarańczami i kadzidłem.

---

[161] Wszystkie polskie przykłady, niebędące przekładami z języka czeskiego, pochodzą ze słownika walencyjnego *Walenty* lub Narodowego Korpusu Języka Polskiego (Przepiórkowski, Bańko, Górski i Lewandowska-Tomaszczyk 2012; Pęzik 2012).

[162] *Walenty* odnotowuje również drugie znaczenie czasownika *nękać* łączącego się z członami o wartości Initiator, Theme i Instrument [człon wspomagający akcję, będący w niej używanym (ale nie zużywanym) przez Initiatora; zazwyczaj jest to obiekt konkretny, materialny, abstrakcyjne Instrumenty mogą się pojawić jedynie w wypadku czasowników dotyczących komunikacji i myślenia] (Hajnicz i Nitoń, 2017: 31).

Ją nękała mnie prośbami o kontakt z trenerem.

Użyte w przekładzie jednostki *martwić, nękać, zamęczać, niepokoić, torturować, doprowadzić do wściekłości* powieliłyby ramę walencyjną czasownika *trápit*; czasownik *gryźć* tylko częściowo.

Pojawienie się rzeczowników *męka, udręka, cierpienie, zmartwienie* jako odpowiedników czasownika *trápit* automatycznie zmienia strukturę wypowiedzi:

*Svěřejí se mně, co jich **trápí**, a může bejt , že jim poradím, poněvadž voják, kterýho vodějí eskortou, ten má vždy větší zkušenosti než ti, co ho hlídají.* JHODV1996cz

*Niech mi się pan zwierzy ze swej **udręki**, a być może, iż zdołam poradzić, ponieważ żołnierz, którego prowadzą pod eskortą, miewa zawsze daleko większe doświadczenie niż ci, co go pilnują.* JHPDW1957pl

*Oba manželé byli zajedno v tom, že ho nesměji nechat zbytečně **trápit**.* MKNLB1985cz

*Obydwoje małżonkowie byli zgodni, że nie mogą mu pozwolić na zbyteczne **cierpienie**.* MKNLB1992pl

Czasownik *przejmować się* w roli ekwiwalentu czeskiej jednostki także nie powiela ramy walencyjnej *trápit*; otwiera bowiem miejsce dla pozycji o wartości Experiencer wyrażonej na powierzchni jako fraza nominalna w mianowniku:

*Tebe to tak **trápí**, Ireno! – říkám a miluju ji.* JSPILD1992cz

*Tak **się** tym **przejmujesz**, Ireno? – pytam i w tej chwili ją kocham.* JSPILD2008pl

Jeszcze więcej odpowiedników udaje się wyszukać dzięki wykorzystaniu narzędzia Treq (Rosen i Vavřín 2015; Škrabal i Vavřín 2017). Tabela 4 przedstawia kolejne możliwości tłumaczenia oraz liczbę poświadczeń tych jednostek jako ekwiwalentów czasownika *trápit*.

Automatyczna ekscerpcja, na której opiera się Treq, nie wychodzi jednak od tekstów oryginalnie pisanych po czesku, ale jako podstawę bierze wszystkie teksty w tym języku (również tłumaczenia) zgromadzone w korpusie. Stąd wysokie liczby poświadczeń. Taki punkt wyjścia jest szczególnie przydatny w sytuacji, kiedy podczas analizy tekstów oryginalnych nie odnajduje się żadne poświadczenie[163]. Wówczas pomocą służą przekłady z trzecich języków. To wielka zaleta Trequ[164].

[163] Przykładem jest np. czasownik *hrotit*, który w subkorpusie czesko-polskim InterCorpu (wyłącznie czeskie oryginały) nie ma ani jednego poświadczenia; szczegółową analizę tego czasownika przeprowadziła w swojej pracy Karolina Kasprzak (Kasprzak 2017)

[164] Te same poświadczenia uzyskać można wpisując zapytanie do KonTextu, ale Treq oferuje natychmiastowe generowanie listy ekwiwalentów z dostępem do poświadczeń.

Tabela 4. Najczęstsze polskie ekwiwalenty czasownika trápit na podstawie serwisu Treq

Najczęstsze polskie ekwiwalenty czasownika *trápit* na podstawie serwisu Treq	
138	*dręczyć*
112	*martwić*
110	*męczyć*
55	*cierpieć*
46	*przejmować*
38	*gnębić*
22	*dokuczać*
20	*trapić*
18	*niepokoić*
16	*zmartwić*
13	*nękać*
12	*torturować*
10	*boleć*
7	*zadręczać*
7	*doskwierać*
7	*kłopot*
7	*cierpienie*

Wnioski

Próba syntaktyczno-semantycznej analizy korpusowej czasownika *trápit* i jego potencjalnych polskich ekwiwalentów okazała się tylko częściowo możliwa. Nadrzędnym problemem jest zbyt mała liczba poświadczeń korpusowych, co jest podczas tego typu badań częstym problemem (Hebal-Jezierska, Kaczmarska i Rosen 2016: 55–57). Z tego powodu nie było możliwe ustalenie, czy pojawienie się lub brak którejś z pozycji walencyjnej lub pewna kombinacja argumentów wpływa na wybór ekwiwalentu. Dokonano jednak próby „rozbicia" jednostek na znaczenia i powiązania ich z ekwiwalentami. Poniższa tabelka przedstawia znaczenia reprezentowane przez analizowane czasowniki (czes. *trápit* i pol. *męczyć, dręczyć, gnębić, trapić*). Znaczek √ oznacza reprezentację znaczenia przez dany czasownik polski, a ● – które znaczenia pokrywa czeska jednostka.

	dręczyć	***gnębić***	***męczyć***	***trapić***
zadawać komuś cierpienie fizyczne / sprawiać ból	√ ●		√ ●	
zadawać komuś cierpienie psychiczne	√ ●	√ ●		√ ●
wywoływać niepokój lub przygnębienie, niepokoić, martwić	√ ●	√ ●	√ ●	√ ●
prześladować kogoś, nękać, gnębić	√ ●	√ ●		
powodować zmęczenie, utratę sił			√	
naprzykrzać się komuś czymś lub mieć stale pretensję o coś			√ ●	
robić coś z wysiłkiem, zbyt długo			√	

Powyższa tabelka, będąca efektem manualnej analizy wszystkich poświadczeń korpusowych, przynosi pewne wskazówki dotyczące możliwości wyboru określonych ekwiwalentów w danych kontekstach.

1) Jeżeli czeska jednostka wyraża zadawanie cierpienia czy bólu fizycznego, wówczas odpowiednimi ekwiwalentami wydają się czasowniki *dręczyć* i *męczyć*.
2) Kiedy *trápit* ma oznaczać zadawanie cierpienia psychicznego, warto rozważyć użycie w polskim przekładzie czasowników *dręczyć*, *gnębić*, *trapić*.
3) Jeśli czasownik czeski niesie informację o prześladowaniu czy nękaniu kogoś, korpus sugeruje zastosowanie czasowników *dręczyć* i *gnębić*.
4) Gdy *trápit* oznacza naprzykrzanie się komuś czymś lub stałe pretensję do kogoś o coś, warto użyć w przekładzie czasownik *męczyć*.
5) Natomiast w sytuacji, kiedy czasownik czeski dotyczy wywoływania niepokoju, przygnębienia lub zmartwienia, prawdopodobnie wszystkie rozważane polskie czasowniki mogą wystąpić w roli ekwiwalentu (*dręczyć*, *gnębić*, *męczyć*, *trapić*).

Wynikające z analizy preferencje (i ograniczenia) semantyczne mogłyby być potwierdzone badaniami na większej liczbie danych[165]. Gdyby przypuszczenia te okazały się prawdziwe, warto byłoby je uwzględnić w opracowaniach leksykograficznych[166].

Automatyczna ekscerpcja poświadczeń i ręczna ich analiza pozwoliły również na wyciągnięcie pewnych wniosków dotyczących wymagań walencyjnych czeskiego czasownika i jego polskich odpowiedników. Ekwiwalenty proponowane przez klasyczny słownik czesko-polski (*męczyć, dręczyć, gnębić, trapić*) organizują przekład danej wypowiedzi w sposób absolutnie niekolizyjny, jeżeli chodzi o schemat walencyjny. Ta sama zasada dotyczy również niektórych odpowiedników wyekscerpowanych z korpusu InterCorp (*martwić, nękać, zamęczać, niepokoić, cierpieć, torturować, doprowadzić do wściekłości,* a także *gryźć* – z ograniczeniami). W tych przypadkach walencja w przekładzie zostaje zachowana. Pozostałe wyszukane w korpusie odpowiedniki (*dokuczać, sprawić przykrość, męka, udręka, cierpienie, zmartwienie, przejmować się*) posiadają odmienną od czeskiego czasownika ramę walencyjną, dlatego użycie ich jako ekwiwalentów *trápit* powoduje, iż walencja pozostaje w przekładzie stracona i można założyć, że również dlatego poświadczenia z tymi jednostkami są sporadyczne.

[165] Należy pamiętać, iż analizy korpusowe nie są badaniami przedstawiającymi stan całego języka i mają do dyspozycji wycinek danych językowych zawartych w korpusie.

[166] Ważna jest też informacja o dodatkowych znaczeniach czasownika *męczyć*, które nie pokrywają się ze znaczeniami czeskiego leksemu: powodować zmęczenie, utratę sił i robić coś z wysiłkiem, zbyt długo. Znaczenia te nie reprezentują jednak typowych odniesień do emocji, uczuć czy przeżyć.

Przypadek 3 – *soucítit* i *soustrast*

W latach 2016–2017 badany był czasownik *soucítit* (Kaczmarska 2017B). W trakcie analizy ustalono, iż w kontekście języka polskiego konieczna jest równoległa analiza tego innej jednostki będącej językowym wykładnikiem emocji związanych ze współczuciem – rzeczownika *soustrast*. W tym rozdziale analizie zostaną poddane struktury semantyczno-składniowe tych jednostek oraz ich polskich odpowiedników, wytypowanych na podstawie znaczeń słownikowych, narzędzia korpusowego Treq i ekscerpcji z korpusu równoległego InterCorp. Zostanie też podjęta próba ustalenia polskich ekwiwalentów obu czeskich jednostek na podstawie poświadczeń korpusowych.

Definicja słownikowa

soucítit

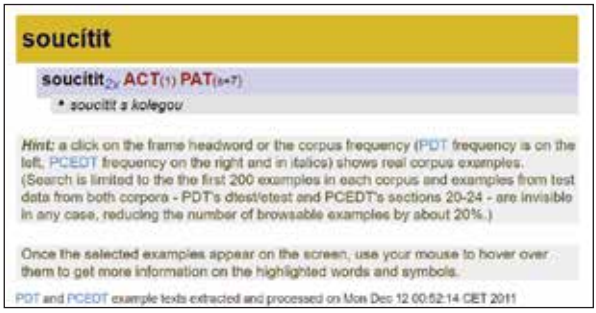
W słowniku literackiego języka czeskiego – SSJC (Havránek i inni 1989) odnotowane jest tylko jedno znaczenie czasownika *soucítit* wraz z jego strukturami składniowymi[167]: 'mít, projevovat soucit (s kým, čím), porozumění pro něčí bol; mít stejné, společné city; sympatizovať. Definicja i charakterystyka składniowa zamieszczona w SSJC jest dość skromna. Nie można jednak skontrolować, czy jest zgodna z opisem w słowniku walencyjnym języka czeskiego VALLEX, ponieważ nie ma w nim hasła *soucítit*[168]. Natomiast PDT-Vallex[169] podaje jeszcze uboższą definicję tego czasownika:

soucítit – kto – s kým

[167] Dostęp 1.12.2018:
http://ssjc.ujc.cas.cz/search.php?hledej=Hledat&heslo=souc%C3%ADtiti&sti=EMPTY&where=hesla&hsubstr=no

[168] Dostęp online: http://ufal.mff.cuni.cz/vallex/3.0/

[169] Dostęp online: http://lindat.mff.cuni.cz/services/PDT-Vallex/PDT-Vallex.html



Obraz 16. Hasło słownika walencyjnego PDT-Vallex[170]

soustrast

W SSJC pojawia się definicja rzeczownika *soustrast*: ‘soucit s trpícím (zejména s pozůstalým), účast na strasti jiného, útrpnost s ní, lítost nad ní; projevy soustrasti; cítit soustrast s pozůstalým; vyslovit, projevit (někomu) soustrasť (Havránek i inni 1989). W definicji pojawiają się też dodatkowe przykłady, które ułatwiają zrozumienie tej jednostki zwłaszcza cudzoziemcom: *přijměte mou upřímnou*, *hlubokou soustrast*; *poslat*, *napsat dopis obsahující vyjádření soustrasti, kondolence*[171].

Słownik dwujęzyczny

soucítit

Tradycyjny słownik czesko-polski (Siatkowski i Basaj 2002: 729) podaje jedno znaczenie tego wyrazu, prezentując wymogi walencyjne obu jednostek – zarówno oryginału, jak i ekwiwalentu: *soucítit* (*s kým, čím*) – *współczuć* (*komu, w związku z czym*). Odpowiednik zaproponowany przez ten słownik nie odpowiada pod względem wymagań składniowych czeskiej jednostce.

soustrast

Słownik czesko-polski (Siatkowski i Basaj 2002: 732) podaje kilka ekwiwalentów tej jednostki: *współczucie*, *kondolencje*, *wyrazy współczucia* oraz

[170] *soucítit* ACT(1) PAT(s+7); *soucítit s kolegou*

[171] Pol. proszę przyjąć moje szczere / głębokie wyrazy współczucia, przesłać, napisać list zawierający wyrazy współczucia, kondolencje.

przykładowe kolokacje: *projevit* (*komu*) *soustrast – wyrazić* (*komu*) *współczucie*, *złożyć* (*komu*) *kondolencje*; *projevy soustrasti – wyrazy współczucia*; *poslat (komu) soustrast – przesłać (komu) kondolencje*.

Analiza korpusowa

Definicje ze słowników oraz z listy odpowiedników wygenerowanej z InterCorpu przybliżają do poznania znaczenia tego czasownika; szerszy obraz oraz ujednoznacznienie znaczeń można uzyskać dzięki przykładom wyekscerpowanym z InterCorpu[172].

soucítit

Analizę rozpoczyna się od automatycznego wygenerowania listy ekwiwalentów badanych czasowników[173] za pomocą narzędzia Treq.

Tabela 5. Najczęstsze polskie ekwiwalenty czasownika *soucítit* na podstawie serwisu Treq

Najczęstsze polskie ekwiwalenty czasownika ***soucítit*** na podstawie serwisu Treq	
31	*współczuć*
1	*podzielać*
1	*serdeczność*
1	*współczuć*
1	*żal*
1	*przejęcie*
1	*sympatia*

Kilka z pojawiających się w tabeli ekwiwalentów nie odpowiada znaczeniu słowa *soucítit*: mogły się tam znaleźć przez przypadek – w wyniku błędu wyrównywania (np. *przejęcie*) lub są częścią zwrotu werbo-nominalnego odpowiadającego badanemu czasownikowi (np. *podzielać*). Wśród proponowanych

[172] Przykłady ekscerpowane były z czesko-polskiej części InterCorpu (wersja 9), z jądra (beletrystyka), oryginalny język źródła – czeski. Na temat wyszukiwania przykładów z subkorpusów – (Kaczmarska i Rosen 2014B).

[173] Poprzednie badania poprzedzone były samodzielnym generowaniem słownika (Kaczmarska i Rosen 2013; Och i Ney 2003; Skoumalová 2008; Jirásek 2011).

ekwiwalentów znalazły się również jednostki bliskoznaczne wobec czasownika *soucítit* (np. *serdeczność, sympatia, żal*).

Wyekscerpowane automatycznie ekwiwalenty są porównywane z poświadczeniami wyekscerpowanymi poprzez wyszukiwarkę KonText[174].

*Do jisté míry jsem **s ním** soucítil, ale právě jen do jisté míry*. MVPNK2003cz
*Do pewnego stopnia **mu** współczułem, ale tylko do pewnego stopnia*. MVSNK2007pl

*„Ta ženská mě zničí," svěřovala se Pamela Maxovi, jenž **s ní** tentokrát docela soucítil.* MVUZ2001cz
*Ta kobieta mnie zniszczy. – Pamela żaliła się Maksowi, który tym razem nawet **jej** współczuł*. MVW2008pl

*Richard Ellman v eseji o Beckettovi píše, že Samuel soucítil i **s humry*** (…) BHANM1992cz
*Richard Ellman w eseju o Becketcie pisze, że Samuel współczuł także **homarom*** (…) BHANM2005pl

*Když teď řeknu, že **s ním** soucítím, „přemítá nahlas Oliver," bude to přestože je to z mé strany bezmála upřímné ještě větší morální hnůj, než kdybych případně tvrdil, že je mi to šumafuk*. MVRPZ2001cz
*Jeśli powiem, że **mu** współczuję – zastanawia się Oliver – będzie to, choć mówię to prawie szczerze, jeszcze większy moralny gnój, niż gdyby m ewentualnie twierdził, że mam to w nosie*. MVPDK2005pl

Wyszukane czeskie poświadczenia wraz z wiązanymi polskim segmentami zostały wyeksportowane do dokumentu Excel[175], dzięki czemu można było materiał dowolnie filtrować i sortować, a po dodaniu kolumn – także tagować. KonText wskazał jednak zaledwie 5 poświadczeń leksemu *soucítit*[176]. Przy każdej konkordancji został oznaczony wybrany przez tłumacza polski odpowiednik. Po otagowaniu okazało się, iż we wszystkich przypadkach w polskim tłumaczeniu występuje czasownik *współczuć*. Ponieważ liczba poświadczeń okazała się zaskakująco

[174] Wyszukiwany czasownik i jego ekwiwalent w wiązanym segmencie zaznaczone są podkreśleniem. Frazy pogrubione to obiekt, któremu się współczuje – *soucítit* **s kým, čím** – *współczuć* **komu**. W podanych przykładach nie pojawiły się argumenty wyrażające powód współczucia.

[175] Wyszukiwarka KonText umożliwia eksport konkordancji do różnego typu plików (Kaczmarska i Rosen 2014B: 216).

[176] Jest to bardzo mała liczba poświadczeń. Na jej podstawie nie jest możliwe przeprowadzenie żadnej głębszej analizy gramatycznej czy stylistycznej. Można jednak wysnuć pewne wnioski co do znaczenia danego słowa i jego tłumaczenia na język polski. Jak twierdzi *Slovník afixů užívaných v češtině*, mała liczba poświadczeń może wynikać z faktu, iż czasowniki czeskie z prefiksem *sou-* są uznawane, w przeciwieństwie do podobnie zbudowanych rzeczowników, za rzadkie i archaiczne (Šimandl 2016); dostęp online: http://www.slovnikafixu.cz/heslar/sou-

mała, zostało przeprowadzone badanie, jak przekładany jest na język czeski leksem *współczuć*[177]. W tym celu jako pierwszy wykorzystany został serwis Treq.

Tabela 6. Czeskie odpowiedniki czasownika *współczuć* na podstawie Treq

Najczęstsze czeskie ekwiwalenty czasownika ***współczuć*** na podstawie serwisu Treq	
307	*soucit*
56	*líto*
53	*soucitně*
51	*litovat*
42	*soucitný*
39	*soustrast*
32	*účast*
31	*soucítit*
26	*lítost*
25	*sympatie*
18	*cítit*
13	*politovat*
11	*pochopení*

Wyniki uzyskane dzięki narzędziu *Treq* ukazują, iż czasownik *soucítit* nie jest ani jedynym możliwym, ani najczęstszym ekwiwalentem polskiego czasownika *współczuć*[178]. Liczba poświadczeń z wyrazem *soucit* jako odpowiednikiem polskiego leksemu *współczuć*, każe baczniej przyjrzeć się temu rzeczownikowi (w dalszej części tegoż rozdziału). Należy jednak wziąć pod uwagę fakt, iż poświadczenia w serwisie Treq zliczane są z całej polsko-czeskiej części korpusu bez względu na język oryginału.

Natomiast wyekscerpowane poświadczenia za pomocą KonTextu to zaledwie 28 trafień czasownika *współczuć* (ekscerpcja z tekstów pisanych w oryginale po polsku).

[177] Według tradycyjnego słownika polsko-czeskiego (Oliva 1995) *współczuć* oznacza *mít soucit* (*s kým*), *vyjadřovat soucit / soustrast* (*komu*), *cítit* (*s kým*).

[178] Niska frekwencja czasownika soucítit skłania do sprawdzenia, jaka jest frekwencja tego czasownika w korpusie języka czeskiego. W największym korpusie syn v4 (jest to połączenie wszystkich istniejących korpusów z serii syn w 2016 roku) obejmującym 4 349 023 692 pozycji wyszukano tylko 2897 poświadczeń czasownika *soucítit*. Dla porównania w tym samym korpusie znaleziono 198 750 poświadczeń czasownika *milovat*. W obu przypadkach wyszukiwano jedynie z tekstów pisanych oryginalnie po czesku.

Tabela 7. Czeskie ekwiwalenty polskiej jednostki *współczuć* na podstawie analizy manualnej danych uzyskanych za pomocą ekscerpcji automatycznej przy użyciu wyszukiwarki KonText

Czeskie ekwiwalenty jednostki ***współczuć***	
cítit s někým	13
soucítit	4
litovat někoho	4
mít soucit s někým	3
být někoho líto	2
politovat	2

Na szczególną uwagę zasługuje zwrot *cítit s někým*, który wypada bezkonkurencyjnie na tle innych proponowanych ekwiwalentów.

Rozumiem go mimo to i współczuję ***mu****, ale nie mogę nic uczynić, aby ulżyć jego doli.* WTOPB1975pl
Přesto ho chápu a cítím ***s ním****, ale nemohu udělat nic, abych mu v jeho trápení pomohl.* WTOPB1981cz

Współczuję ***ci****, że musisz śpiewać w takich miejscach.* ASMP1992pl
Cítím ***s tebou****, že musíš zpívat na takových místech.* ASMO1993cz

Madame Titine współczuła i podzielała nadzieję. MKGO1961pl[179]
Madame Titina cítila ***s ní*** *a sdílela její naději.* MKOH1971cz

Większość przykładów prezentuje typowy dla tej ekwiwalencji układ:

współczuć komu – cítit s někým.

[179] W polskim oryginale nie jest wyrażony obiekt, któremu współczujemy. Można o nim wnioskować na podstawie szerszego kontekstu. Czeski tłumacz decyduje się skonkretyzować obiekt współczucia:

pl. *"On" był jej separowanym mężem, który żył na farmie ze służącą. Od dziesięciu już lat Cottonowie procesowali się o różne sumy i posiadłości, i przez tyleż lat "Dom" stał zamknięty. Jednak gdy Monsieur Cotton miał się zjawić, by obrzucić żonę nową porcją obelg, nacierała oliwą i goździkowym olejkiem starożytne skrzynie i monumentalne szafy z bukietami polnych kwiatów wymalowanymi na drzwiach, ściągała pokrowce z atłasowych kanap, czyściła miedź w nieśmiertelnej nadziei pojednania. Madame Titine* ***współczuła*** *i podzielała nadzieję.*

cz. *"On" byl její rozloučený manžel, který žil na farmě se služkou. Už deset let se Cottonovi soudili o různé peněžité částky a nemovitosti, a po celá ta léta byl "Dům" zavřený. Avšak pokaždé, když se pan Cotton měl objevit, aby svou manželku zahrnul novou sprškou nadávek, natírala olivovým olejem a přelešťovala hřebíčkovým olejíčkem starožitné truhly a monumentální skříně s dveřmi pomalovanými kyticemi polního kvítí , stahovala ochranné povlaky z atlasových pohovek a leštila měděné předměty v naději na smír. Madame Titina* ***cítila s*** *ní a sdílela její naději.*

W poświadczeniach tych pojawia się jednak jeden ciekawy przykład, który pokazuje odmienną ramę walencyjną czasownika *współczuć*, zgadzającą się tym razem z ramą czeskiego odpowiednika. Dla współczesnego języka polskiego rama ta jest jednak obca:

Proszę mi wierzyć, że bardzo współczuję z panem. WRZO1957pl[180]
Věřte mi, prosím vás, že ***s vámi*** *upřímně cítím*. WRZZ1967cz

W przypadku wyników tej ekscerpcji warto też zwrócić uwagę na przykłady z czeskim ekwiwalentem *být někoho líto*:

Tak bardzo ***ci*** *współczuję, załóż stowarzyszenie, niech inne takie wariatki też bez języków walczą przeciwko mnie alfabetem Morse'a…* DMWPR2002pl
Je mi ***tě*** *moc líto, založ spolek, aby i další takový pošahaný jako ty bez jazyku proti mně bojovaly morseovkou…* DMCAB2004cz

Oczywiście współczuł ***jej****, ale drażniła go swoją natarczywą nerwowością, niepokojem, nagłymi wybuchami.* SCHH1997pl
Samozřejmě mu ***jí*** *bylo líto, ale provokovala ho svou agresivní nervozitou, neklidem, záchvaty vzteku.* SCHH2005cz

Na pierwszy rzut oka dochodzi tu do swoistego przesunięcia znaczeniowego, w istocie jednak czeski tłumacz wybiera tłumaczenie jednego ze znaczeń czasownika *współczuć*[181]:

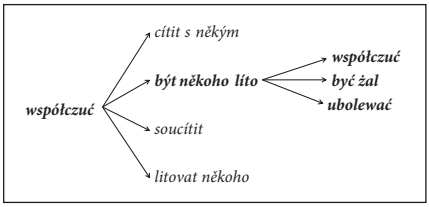
- współczuć,
- odczuwać żal i smutek z powodu złego powodzenia innych ludzi,
- ubolewać nad kimś,
- litować się,
- dawniej: czuć to samo co inni.

W przesunięciu tym można się dopatrywać swoistej rekonceptualizacji (Lewandowska-Tomaszczyk 2010; Kaczmarska 2015C), o której mowa była już

[180] Władysław Reymont, *Ziemia obiecana*, Kraków 1957 (publikowana 1897–1898 i 1899).

[181] Definicja pochodzi z Internetowy Słownik Języka Polskiego – http://sjp.pl/wsp%C3%B3%C5%82czu%C4%87. Definicja oferowana przez Słownik Języka Polskiego PWN (SJP PWN) jest uboższa: 'solidaryzować się uczuciowo z kimś cierpiącym' – http://sjp.pwn.pl/szukaj/wsp%C3%B3%C5%82czu%C4%87.html. SJP PWN podaje jako synonimy jednostki takie jak: *żałować, litować się, użalać się, łączyć się w bólu, zmiłować się*.

w rozdziale 2. W kontekście przekładu obserwuje się bowiem cykle rekonceptualizacyjne wywołane asymetrią semantyczną[182].



Obraz 17. Schemat rekonceptualizacyjny jednostki *współczuć* w odniesieniu do ekwiwalentu *být někoho líto*

Podobne schematy można utworzyć dla niemal każdej jednostki będącej wykładnikiem emocji czy uczuć. Warto przy tym jednak zwrócić uwagę na fakt, iż konstrukcje te mogą mieć różne ramy walencyjne (choćby: *współczuć komu* + + celownik, *być żal* komu + celownik – kogo / czego + dopełniacz, *ubolewać nad* + narzędnik). Ma to olbrzymie znaczenie przy wyborze ekwiwalentu w procesie tłumaczenia; bywa, iż tłumacz, chcąc zachować identyczną konstrukcję składniową (np. w celu utrzymania tempa wypowiedzi), zmuszony jest do użycia ekwiwalentu modyfikującego lekko znaczenie.

soucit

Konieczność analizy tego leksemu pojawiła się podczas automatycznej ekscerpcji ekwiwalentów czasownika *współczuć* przy użyciu narzędzia Treq. W czesko-polskiej części korpusu InterCorp znaleziono 77 poświadczeń rzeczownika *soucit*.

Tabela 8. Polskie ekwiwalenty czeskiej jednostki *soucit* na podstawie analizy manualnej danych uzyskanych za pomocą ekscerpcji automatycznej przy użyciu wyszukiwarki KonText

Polskie ekwiwalenty jednostki ***soucit***	
współczucie	63
współczuć	4
litość	7
wrażliwość	1
pominięcie	1
błąd anotacji	1

[182] Barbara Lewandowska-Tomaszczyk (Lewandowska-Tomaszczyk 2010: 12) ukazuje to na przykładzie angielskiego czasownika *to go* mogącego oznaczać *iść* albo *jechać*. Czasownik *iść* z kolei można rozumieć jako *to go* (*to school*) i *to walk*. *Jechać* może oznaczać *to go* (*by bus*), *to drive* (*a car*), *to ride* (*a bike, horse*).

Znakomita większość poświadczeń posiada polski odpowiednik *współczucie*. Warto się więc przyjrzeć, jakie czasowniki wiążą się z czeską jednostką *soucit*. W zgromadzonym materiale znalazły się 4 frazy werbo-nominalne z komponentem *soucit*.

- *vzbuzovat soucit*

W języku polskim istnieje analogiczna fraza – *wzbudzać współczucie*. W obu językach obiekt, któremu współczujemy występuje w pozycji podmiotu, jakby konstrukcja dopuszczała aktywne wzbudzanie współczucia. Nie musi być natomiast wyrażony przeżywający to współczucie[183]. Właśnie taką sytuację prezentuje poniższy przykład:

Willi říkal, že žebráci alkoholici vzbuzují spíš odpor než soucit, a to já nesmím. LFSM1983cz

Willi mówił, że żebracy alkoholicy wzbudzają raczej wstręt niż współczucie, a tego mi nie wolno. LFPZ1979pl

- *mít soucit s někým*

Fraza polska, będąca odpowiednikiem analizowanej czeskiej jednostki, ma te same komponenty *mieć* + *współczucie*. W polskim przykładzie pojawia się obiekt, któremu współczujemy *mieć współczucie* **dla ludzi** (dla + obiekt w dopełniaczu). Taka pozycja nie pojawia się przy czasowniku *współczuć*. Wymagania czeskiej jednostki *mít soucit* są zgodne z wymaganiami czasownika *soucítit*.

*Sousedka se pohoršovala: nemá soucit **s lidmi**, ale pláče nad psem*. MKN1993cz

*Sąsiadka była oburzona: ta dziewczyna nie ma żadnego współczucia **dla ludzi**, lecz opłakuje swego psa*. MKN2008pl

- *cítit soucit*

Czeska i polska jednostka są zgodne pod względem komponentów. W obu językach, w przypadku zastosowania negacji jest możliwe uniknięcie wyrażenia obiektu współczucia:

Tomáš v sobě necítil žádný soucit. MKNLB1985cz

Tomasz nie czuł w sobie żadnego współczucia. MKNLB1992pl

- *pocítit soucit*

[183] Experiencer może być jednak wyrażony – *vzbuzovat soucit v někom* (wzbudzać współczucie w kim).

Czeska jednostka odnosi się do momentu wystąpienia uczucia współczucia, naznaczając niejako jego początek[184]:

Znovu jí ho přišlo líto, pocítila k němu soucit starší sestry a v té chvíli učinila něco, o čem ještě před vteřinou neměla tušení. MKN1993cz

Znów wzbudził w niej litość, litość starszej siostry; i zrobiła wówczas rzecz, która chwilę wcześniej nie przyszłaby jej do głowy. MKN2008pl

Cztery opisane czeskie struktury oferują wybór różnych możliwości składniowych, a w konsekwencji również zmiany semantyczne. Na te zmiany musi reagować tłumacz; język polski, co ukazują przykłady, dysponuje podobnymi konstrukcjami.

soustrast

Serwis Treq dostarcza szeroką gamę ekwiwalentów:

Tabela 9. Najczęstsze polskie ekwiwalenty jednostki ***soustrast*** na podstawie serwisu Treq

Najczęstsze polskie ekwiwalenty jednostki ***soustrast*** na podstawie serwisu Treq	
232	*kondolencja*
192	*współczuć*
30	*przykro*
7	*strata*
5	*ubolewać*
4	*sympatia*
3	*współczucie*
3	*żal*
3	*solidarność*
2	*litość*

Przeważająca większość przykładów wiąże się z sytuacją składania kondolencji (również w przypadku: *wyrazy współczucia, wyrazy ubolewania, łączyć się*

[184] Polski przekład wykorzystuje konstrukcję *wzbudzić litość*, co jednak zmienia znaczenie całej frazy. Struktura *poczuć współczucie* jest jednak możliwa w języku polskim, choć jest rzeczywiście rzadka:

pl. *Ogólnie uważa się jednak, że dwa główne procesy zachodzące w naszym mózgu muszą być obecne, żebyśmy w ogóle byli w stanie poczuć współczucie wobec innej osoby*.

– z portalu: http://www.psychologia-spoleczna.pl/aktualnosci/2006-neuroempatia.html

w żalu); dotyczy to niemal 100% przykładów, kiedy w języku czeskim pojawiają się frazy *vyjádřit soustrast* **i** *upřímnou soustrast*.

*…moc zdvořilá a dokonalá dáma, s lesknoucíma se očima **jim** vyjádřila soustrast.* JUCZE1993cz

*…prawdziwa grande dame, złożyła **im** kondolencje z błyszczącymi od łez oczyma.* JUCZE2008pl

Ještě jednou – přijměte mou nejupřímnější soustrast. HCTNN2003cz
Ponownie przyjmij moje najszczersze wyrazy współczucia. HCJS2008pl

*Rád bych **jí** vyjádřil naši soustrast.* EPcz
Łączymy wyrazy ubolewania. EPpl

*Závěrem posílám soustrast **nigerijským občanům**…* EPcz
*Wreszcie, szczególnie dzisiaj, łączę się w żalu **z mieszkańcami Nigerii**…* EPpl

Ekwiwalenty te warto porównać z poświadczeniami z korpusu InterCorp. Za pomocą wyszukiwarki KonText ekscerpowane są segmenty, które zostają otagowane, a następnie posortowane pod względem polskich ekwiwalentów. KonText wskazał jednak zaledwie 7 poświadczeń leksemu *soustrast*[185]. Tak mała liczba przykładów nie może być podstawą analizy.

Tabela 10. Polskie ekwiwalenty czeskiej jednostki *soustrast* na podstawie analizy manualnej danych uzyskanych za pomocą ekscerpcji automatycznej przy użyciu wyszukiwarki KonText

Polskie ekwiwalenty jednostki ***soustrast***	
współczucie	5
współczuć	1
współczująco	1

Upřímnou soustrast, soudružko, tiskne Erninu ruku. JGD1990cz
Szczere wyrazy współczucia, towarzyszko – ściska jej rękę. JGK2003pl

Rozszerzenie materiału, który KonText przeszukuje, o teksty z Parlamentu Europejskiego i ponowna ręczna analiza poświadczeń potwierdzają tezę, którą można wysnuć podczas analizy słownikowej. Jeżeli jednostce *soustrast* towarzyszy określenie *upřímná* lub czasownik *vyjádřit*, wówczas odnosi nas to do ekwiwalentu *kondolencje* / *składać kondolencje* / *wyrazy współczucia*.

[185] Wyszukiwanie przeprowadzane było wyłącznie na tekstach oryginalnie czeskich, ręcznie anotowanych (trzon – czes. *jádro*).

Parlament vyjadřuje soustrast ***rodinám*** *obětí a solidaritu zraněným.* EPcz
Parlament składa kondolencje ***rodzinom*** *ofiar i wyraża solidarność z rannymi.* EPpl

Chci se rovněž připojit k vyjádření soustrasti. EPcz
Pragnę również złożyć wyrazy współczucia. EPpl

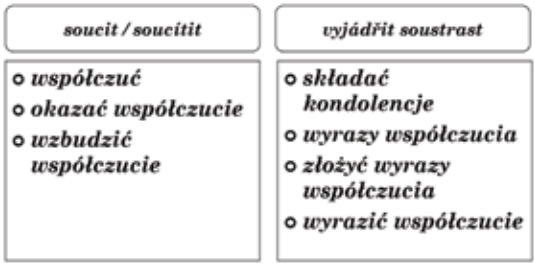
Chtěl bych proto vyjádřit soustrast ***těm****, kteří ztratili člena rodiny.* EPcz
Chciałbym więc wyrazić współczucie ***dla tych****, którzy stracili członków swoich rodzin.* EPpl

Fraza *upřímnou soustrast* może być używana także w znaczeniu przenośnym, kiedy wyrażamy żal (często podszyty ironią) np. w związku z przegraną ulubionej drużyny futbolowej, zgubieniem portfela czy niepłatnymi nadgodzinami.

Doufám, že dostaneš **přesčasy**. *Ne? Upřímnou* **soustrast.**
Mam nadzieje, że płacą ci za **nadgodziny**. *Nie? Moje kondolencje.*

Wnioski

Ręczna analiza wszystkich wyekscerpowanych automatycznie poświadczeń badanych jednostek pozwoliła dostrzec pewną regułę dotyczącą wyboru ekwiwalentów. Prezentuje ją poniższy obraz:



Obraz 18. Rozkład ekwiwalentów czeskich badanych jednostek

Jeżeli pojawia się czeska fraza *vyjádřit soustrast*, jej polskim odpowiednikiem będzie jednostka wyrażająca werbalne przekazanie kondolencji czy wyrazów współczucia (dotyczy to również kondolencji w formie pisanej): *składać kondolencje, składać / przekazywać wyrazy współczucia, wyrazić współczucie*[186]. Jeżeli

[186] Jednostki tego typu – *vyjádřit soustrast / składać kondolencje, składać / przekazywać wyrazy współczucia, wyrazić współczucie* nie należą jednak do klasy czasowników odnoszących się do emocji, uczuć czy przeżyć zdefiniowanej w tej pracy; są predykatami wiążącymi się z komunikacją i aktami mowy, por. (Greń 1994).

natomiast w tekście czeskim pojawią się jednostki *soucítit* czy *soucit*, wówczas polskimi odpowiednikami będą jednostki odnoszące się do uczuć (*współczuć, wzbudzić współczucie*) czy reakcji fizycznych (*okazać współczucie*), ale nie do ustnej czy pisemnej formy kondolencji.

Perspektywy

Próba syntaktyczno-semantycznej analizy korpusowej jednostek *soucítit* i *soustrast* okazała się tylko częściowo możliwa. Problemem jest mała liczba poświadczeń korpusowych analizowanych jednostek.

Dzięki analizie otoczenia jednostek *soucítit* / *soucit* i *soustrast* udało się ustalić regułę pojawiania się polskich ekwiwalentów opisaną we wnioskach powyżej. Nie powiodło się natomiast ustalenie, czy występowanie (lub brak) któregoś z argumentów (obiekt współczucia czy powód współczucia) wpływa na wybór ekwiwalentu.

Tylko częściowo udana analiza determinuje jednak do poszukiwania innych rozwiązań. Pomocne mogłoby się okazać wykorzystanie profili kolokacyjnych Word Sketches[187] w oparciu o materiał z korpusu równoległego InterCorp[188]. Narzędzie Sketch Engine automatycznie ukazałoby, z jakimi obiektami (i w jaki sposób pod względem składniowym) łączą się omawiane jednostki; przy czym obiekty pogrupowane byłyby w tzw. klastry, np. w jednej grupie „człowiek" znalazłyby się obiekty *matka, student, sąsiadka*[189]. Badania czesko-polskie w oparciu o dane z korpusu InterCorp z wykorzystaniem Sketch Engine wciąż jednak można traktować tylko jako eksperyment[190]. Wizja zastosowania tego narzędzia i wygenerowanych przez niego profili Word Sketches zmusza do wysunięcia dwóch postulatów: powiększenia części czesko-polskiej korpusu InterCorp oraz uruchomienia możliwości wykorzystania tego Sketch Engine dla języka czeskiego w ramach InterCorpu.

[187] "A word sketch is a one-page summary of the word's grammatical and collocational behaviour. It shows the word's collocates categorised by grammatical relations such as words that serve as an object of the verb, words that serve as a subject of the verb, words that modify the word etc". – https://www.sketchengine.eu/user-guide/user-manual/word-sketch/

[188] Metoda ta bliska jest poniekąd analizie wymogów walencyjnych.

[189] Pełne wykorzystanie tego narzędzia to jednak kwestia przyszłości, ponieważ o ile można je zastosować na materiale polskim z korpusu InterCorp, o tyle materiał czeski może być analizowany tylko poprzez podkorpusy jednojęzyczne języka czeskiego; to sprawia, że wyniki są nieporównywalne. Problem ten powróci w dalszych częściach tego opracowania.

[190] Wygenerowane tabele profili kolokacyjnych w Załączniku 1.

Przypadek 4 – *závidět* i *žárlit*

Wyrażające negatywne emocje czasowniki *závidět* i *žárlit* były przedmiotem przeprowadzanych w latach 2014–2016 badań, mających na celu ustalenie ekwiwalentów poprzez analizę struktury semantyczno-składniowej tych jednostek oraz ich polskich potencjalnych odpowiedników wskazanych przez słowniki i korpus. Analizowane czeskie czasowniki nie są synonimami, jednak obydwa odnoszą się w języku polskim do pojęcia zazdrości.

Definicja słownikowa

závidět

Znaczenie czasownika *závidět* oraz informację o jego strukturach składniowych przedstawia słownik literackiego języka czeskiego SSJC (Havránek i inni 1989)[191]. Czasownik *záviděti* definiowany jest następująco: ‘pociťovat závist k někomu, nepřát (komu co, koho)’[192]. Słownik ten podaje również przykłady użycia (*záviděti boháčům*, *šťastným lidem*; *lidé si navzájem závidí*; *záviděti mladé dívce krásu, ženicha*; *mnozí mu té slávy záviděli*; *tu funkci ti nezávidím, nechtěl bych ji*[193]) oraz schemat struktur składniowych, w których ten czasownik się pojawia (komu; komu co, koho, †čeho[194]).

Definicja i charakterystyka składniowa zamieszczona w SSJC (Havránek i inni 1989) jest zgodna z opisem w słowniku walencyjnym języka czeskiego VALLEX[195]:

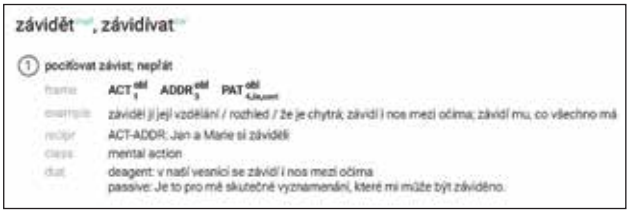
[191] Dostęp 20.01.2019: http://ssjc.ujc.cas.cz/search.php?hledej=Hledat&heslo=z%C3%A1vid%C4%9Bti&sti=EMPTY&where=hesla&hsubstr=no

[192] Pol. czuć zawiść wobec kogoś, nie życzyć komu, co / kogo (biernik) – tłumaczenie własne autorki.

[193] Pol. zazdrościć bogaczom, szczęśliwym ludziom; ludzie sobie nawzajem zazdroszczą; młodej dziewczynie urody, narzeczonego; wielu mu tej sławy zazdrości; tej funkcji ci nie zazdroszczę, nie chciałbym jej – tłumaczenie własne autorki.

[194] komu? komu co / kogo (biernik), czego (przestarzała struktura dopełniaczowa)

[195] Dostęp online: http://ufal.mff.cuni.cz/vallex/2.6.3/data/html/generated/alphabet/index.html



Obraz 19. Hasło *záviděť* w słowniku walencyjnym VALLEX[196]

žárlit

Drugą jednostką związaną z pojęciem zazdrości jest w języku czeskim wspomniany już czasownik *žárlit*. Słownik języka czeskiego przedstawia następujące definicje tej jednostki[197]: 1) ‘pociťovat nelibost, bolest, zlost vůči někomu milovanému, protože má rád někoho jiného’; 2) ‘pociťovat nelibost, bolest, zlost vůči někomu, koho má ráda milovaná osoba’; 3) ‘pociťovat řevnivost, nevraživost na někoho pro jeho úspěchy’; 4) ‘nevražit, řevnit, nepřát, záviděť.[198] Słownik podaje również skrótową informację, w jakich frazach pojawia się analizowany czasownik: *žárlit* na koho, na co; †s kým[199].

Z tymi definicjami zgodny jest opis w słowniku VALLEX:

Obraz 20. Hasło *žárlit* w słowniku walencyjnym VALLEX[200]

[196] Pol. czuć zawiść, nie życzyć; zazdrościł jej wykształcenia / perspektywy / że jest mądra; zazdrości bardzo (dosł. zazdrości i nosa między oczami); zazdrości mu wszystkiego, co on ma; Jan i Maria zazdrościli sobie nawzajem – tłumaczenie własne autorki.

[197] Na podstawie definicji (dostęp online 15.03.2019): http://ssjc.ujc.cas.cz/search.php?hledej=Hledat&heslo=%C5%BE%C3%A1rliti&sti=EMPTY&where=hesla&hsubstr=no

[198] Pol. 1) czuć antypatię, ból, wrogość wobec kogoś ukochanego, ponieważ lubi kogoś innego; 2) czuć antypatię, ból, wrogość wobec kogoś, kogo lubi ukochana osoba; 3) ‘czuć zazdrość, zawiść wobec kogoś ze względu na jego sukcesy’; 4) ‘czuć zawiść, rywalizować, nie życzyć komuś dobrze, zazdrościć’ – tłumaczenie własne autorki.

[199] na kogo, na co; z kim (przestarzała struktura narzędnikowa).

[200] Pol. czuć antypatię / ból / wściekłość / złość z powodu czegoś / do kogoś / zazdrościć; Davídek był zazdrosny o swoją młodszą siostrzyczkę; był zazdrosny o jego sukcesy; nie jestem zazdrosna o jego nową przyjaciółkę; był strasznie zazdrosny, że nie wygrał zamiast ciebie; był zazdrosny o to, jak mu idzie; byli o siebie zazdrośni – tłumaczenie własne autorki.

Słownik dwujęzyczny

Analiza możliwości rozumienia i przekładu czasownika *závidět*

Tradycyjny słownik czesko-polski (Siatkowski i Basaj 2002: 999) podaje jedno znaczenie tego wyrazu: *zazdrościć komu czego*[201]. Odpowiednik zaproponowany przez słownik tradycyjny odpowiada pod względem wymagań składniowych czeskiej jednostce. Większą różnorodność ekwiwalentów przynosi korpusowe narzędzie Treq.

Tabela 11. Najczęstsze polskie ekwiwalenty czasownika *závidět* na podstawie serwisu Treq

Najczęstsze polskie ekwiwalenty czasownika ***závidět*** na podstawie serwisu **Treq**	
zazdrościć	**188**
zazdrość	**26**
pozazdrościć	**16**
zazdrosny	**7**
zawiść	**4**
darzyć	**1**
straszliwie	**1**
współzawodniczyć	**1**
zwyknąć	**1**

Kilka z pojawiających się w tabeli odpowiedników nie odpowiada znaczeniu słowa *závidět*: mogły się tam znaleźć przez przypadek – w wyniku błędu wyrównywania (*zwyknąć*, *straszliwie*) lub są częścią zwrotu werbo-nominalnego odpowiadającego badanemu czasownikowi (*darzyć*). Wśród proponowanych ekwiwalentów znalazł się również czasownik *współzawodniczyć*, który można interpretować jako luźny synonim jednostki *závidět* w jednym z jej znaczeń.

Analiza możliwości rozumienia i przekładu czasownika *žárlit*

Słownik czesko-polski (Siatkowski i Basaj 2002: 1049) podaje dwa znaczenia tego czasownika: *być zazdrosnym (o kogo, o co)* i *zazdrościć (komu)*. Serwis Treq dostarcza szerszą gamę ekwiwalentów.

[201] Ten polski czasownik jest też jednak jednym z ekwiwalentów czeskiej jednostki *žárlit* (Siatkowski i Basaj 2002: 1049). W trakcie analizy należy więc też ustalić, które znaczenia czasownika *zazdrościć* pokrywa *závidět*, a które *žárlit*.

Tabela 12. Najczęstsze polskie ekwiwalenty czasownika *žárlit* na podstawie serwisu Treq.

Najczęstsze polskie ekwiwalenty czasownika ***žárlit*** na podstawie serwisu **Treq**	
zazdrosny	**141**
zazdrość	**25**
zazdrościć	**14**
być	**2**
osiłek	**1**
darzyć	**1**
owszem	**1**
rywalka	**1**
zawiść	**1**

Podobnie jak w sytuacji poprzedniego czasownika, również tutaj wśród oferowanych ekwiwalentów znajdują się jednostki przypadkowo związane (*być*, *osiłek, darzyć, owszem*). Słowa *rywalka* i *zawiść* mogą być traktowane jako elementy struktur synonimicznych wobec pojęcia ‘zazdrość’.

Definicje z obu słowników oraz z listy odpowiedników wygenerowanej z InterCorpu przybliżają nas do identyfikacji i rozróżnienia znaczeń tych czasowników. Pełny obraz można uzyskać dzięki poświadczeniom wyekscerpowanym z InterCorpu[202].

Analiza korpusowa

Analiza korpusowa czasownika *závidět*

Materiał do badań został wyekscerpowany z korpusu równoległego InterCorp poprzez wyszukiwarkę KonText[203].

*Tu jejich zjevnou pohodu a bezstarostnost **jim** upřímně záviděl.* MVUZ2001cz
*Szczerze **im** zazdrościł pogody ducha i beztroski.* MVW2008pl

[202] Przykłady ekscerpowane były z czesko-polskiej części InterCorpu (wersja 8), z jądra (beletrystyka), oryginalny język źródła – czeski.

[203] Tekst podkreślony linią pofalowaną to przedmiot zazdrości – *zazdrościć* kogo czego (w przypadku czasownika *žárlit* będzie to fraza przyimkowa *na* + obiekt, podobnie jak *być zazdrosny o* + obiekt), natomiast tekst podkreślony podwójną linią oznacza powód zazdrości wyrażony frazą zdaniową. Frazy pogrubione to obiekt, któremu się zazdrości – *zazdrościć* **komu**. Pojedyncze podkreślenie linią prostą dotyczy analizowanych czasowników i ich ekwiwalentów.

Agnes záviděla ***Paulovi****, že žije, aniž si musí stále uvědomovat, že má tělo*. MKN1993cz
Agnes zazdrościła ***Paulowi****, że może żyć bez ciągłej świadomości własnego ciała*. MKN2008pl

KonText wskazał 50 poświadczeń leksemu *závidět*[204]. Przy każdej konkordancji zostało zaznaczone, jaki odpowiednik polski został wybrany. Po otagowaniu okazało się, iż w 90% przypadków w polskim tłumaczeniu występuje czasownik *zazdrościć* (ewentualnie wariant *pozazdrościć*).

Tabela 13. Polskie odpowiedniki czasownika *závidět* na podstawie materiału z czesko-polskiej części korpusu InterCorp

Polskie odpowiedniki czasownika ***závidět***	**50**
zazdrościć	45
pozazdrościć	3
być zazdrosny	1
inne	1

Jedyne poświadczenie z ekwiwalentem *zazdrosny*, przedstawia przykład z elipsą.

(…) *spokojen, že* ***mu*** *nemá co závidět…* VPVSP2001cz
(…) *zadowolony, że nie potrzebuje być zazdrosny…* VPTSZ1968pl

Można więc uznać, iż optymalnym ekwiwalentem tej jednostki jest czasownik *zazdrościć*. Ten jednak jest wieloznaczny. *Słownik języka polskiego* przedstawia definicję tej jednostki: 1) 'odczuwać zazdrość'[205]; 2) 'odczuwać żal z powodu tego, że komuś dobrze się powodzi, że ktoś coś ma, pragnąc tego dla siebie' (Szymczak 1995: 916).

Czasownik *závidět* realizuje jedno ze znaczeń jednostki *zazdrościć* ('odczuwać żal z powodu tego, że komuś dobrze się powodzi, że ktoś coś ma, pragnąc tego dla siebie'). Pokrywa też jedno ze znaczeń rzeczownika *zazdrość*: 1) 'uczucie przykrości, żalu spowodowane czyimś powodzeniem, szczęściem, stanem posiadania itp. i chęć posiadania tego samego'; 2) 'uczucie niepokoju co do wierności osoby

[204] Jest to bardzo mała liczba poświadczeń. Na jej podstawie nie jest możliwe przeprowadzenie żadnej głębszej analizy gramatycznej czy stylistycznej. Można jednak wysnuć pewne wnioski co do znaczenia danego słowa i jego tłumaczenia na język polski.

[205] Zazdrość – 1) uczucie przykrości, żalu spowodowane czyimś powodzeniem, szczęściem, stanem posiadania itp. i chęć posiadania tego samego; 2) uczucie niepokoju co do wierności osoby kochanej, podejrzliwość i dążenie do wyłączności w tym zakresie, chęć przeciwdziałania ewentualnemu naruszeniu tej wyłączności. (Szymczak 1995: 916)

kochanej, podejrzliwość i dążenie do wyłączności w tym zakresie, chęć przeciwdziałania ewentualnemu naruszeniu tej wyłączności' (Szymczak 1995: 916).

Odczytując znaczenie czasownika *zazdrościć* (zgodnie z wcześniej przytoczoną definicją) jako 'odczuwać zazdrość', można owo znaczenie rozszerzyć do postaci: 'odczuwać uczucie przykrości, żalu spowodowane czyimś powodzeniem, szczęściem, stanem posiadania itp. i chęć posiadania tego samego'.

Tak jsem tu postával dál, doufal jsem nevěda v co a záviděl jsem ***Vilému Habovi****, jak lehce našel východisko z nouze*. VRR1944cz

Stałem więc nadal, ufając w szczęśliwy zbieg okoliczności i zazdroszcząc ***Wilhelmowi Habie****, że tak łatwo znalazł wyjście z trudnej sytuacji*. VRK1954pl

Zgadza się to również z definicją słownikową czasownika *závidět*.

Analiza korpusowa czasownika *žárlit*

Za pomocą wyszukiwarki KonText wyekscerpowano poświadczenia, które następnie przeanalizowano i posortowano pod względem polskich ekwiwalentów. Przykłady z InterCorpu i w tym przypadku są ilustracją zgromadzonych definicji słownikowych:

Navíc jsem žárlil na Alici, že mě nechala jen tak rozbaleného (…) JGD1990cz

Ponadto byłem zazdrosny o Alicję, bo zostawiła mnie odwiniętego z pieluszek (…) JGK2003pl

KonText wskazał 59 poświadczeń leksemu *žárlit*. Podobnie jak w poprzednim badaniu również w tym przypadku zaznaczone zostały odpowiedniki polskie. Po otagowaniu okazało się, iż w niemal 78% przypadków w polskim tłumaczeniu występuje fraza *być zazdrosny*, wprowadzająca odmienną strukturę składniową.

Tabela 14. Polskie odpowiedniki czasownika *žárlit* na podstawie materiału z czesko-polskiej części korpusu InterCorp.

Polskie odpowiedniki czasownika ***žárlit***	**59**
być zazdrosny	46
zazdrościć	7
zazdrość	3
poczuć zazdrość	1
zawiść	1
błąd	1

W wiązanych segmentach polskich występuje również czasownik *zazdrościć*. W przykładach tych jednak nie występuje najczęściej typowy obiekt zazdrości (wyrażony poprzez frazę nominalną w dopełniaczu), ale pewnego rodzaju powód zazdrości i jest wyrażony poprzez frazę zdaniową:

Povídám, jako vždycky, von na mě žárlí, že jsem mladší než von. BHPHS1994cz
Powiadam, jak zawsze, on ***mi*** *zazdrości, że jestem młodszy niż on.* BHZGS1993pl

Czasownik *zazdrościć* (jako ekwiwalent *žárlit*) pojawia się też konstrukcjach eliptycznych:

Právě proto, že už nechce žárlit, bere vážně a bez podezření jeho tvrzení! MKVNR1997cz
Właśnie dlatego, że nie chce już zazdrościć, przyjmuje jego słowa poważnie i bez podejrzeń! MKWP2001pl

Również w przypadku czasownika *žárlit* można wskazać trafny ekwiwalent w języku polskim; to fraza *być zazdrosnym*[206].

Zjawiskiem związanym z ekwiwalencją jednostek *žárlit* i *być zazdrosnym* jest jednak trudność w identyfikacji obiektu / powodu w strukturze zdania. Problem ten został też poruszony w poście do poradni językowej[207]:

Moj dylemat dotyczy wyrażania zazdrości… Jestem mężatką. Wyobraźmy sobie, ze wokół mojego męża „kręci" się jakaś atrakcyjna kobieta. Mój mąż zwraca na nią uwagę, podchodzi, rozmawia… Ja, będąc oczywiście bardzo wściekłą, mówię: „Nie rób tego nigdy więcej, bo jestem…" No właśnie, jak się mówi: „Jestem o ciebie zazdrosna" czy „Jestem o nią zazdrosna"? Te dwa wyrażenia słyszy się często. Problem w tym, że odnoszą się one do tej samej sytuacji, takiej jak na przykład ta przeze mnie Panu przedstawiona[208].

Problem ten występuje zarówno w języku polskim, jak i czeskim, co paradoksalnie może czynić ekwiwalencję między nimi jeszcze silniejszą. W niektórych przykładach na właściwe znaczenia naprowadza nas kontekst zdania:

Máma byla Pažoutovi po celou dobu věrná, ale on na ni přesto neustále žárlil (…) MVRPZ2001cz

---

[206] Zazdrosny – 1) pragnący tego, co ma ktoś inny, odczuwający żal, że komuś powodzi się lepiej niż jemu; 2) bojący się o swoje dobro, podejrzliwie strzegący swego; zwłaszcza: podejrzliwy wobec współmałżonka, osoby kochanej (Szymczak 1995).

[207] https://sjp.pwn.pl/poradnia/haslo/Jestem-zazdrosna;887.html – dostęp online 15.03.2015 (pisownia oryginalna).

[208] Odpowiedź (fragment): *Istotnie, jest tu pewna językowa nielogiczność, ponieważ w tych dwóch użyciach przymiotnik zazdrosny może być stosowany. W praktyce, gdy mówię, że jestem zazdrosny o żonę, odnoszę to do sytuacji „w ogóle", gdy zaś mówię (np. do żony), że jestem zazdrosny o jej kolegę, odnoszę to do jej relacji z konkretnym człowiekiem. Zazwyczaj wiemy, o co chodzi - ale rzeczywiście, tu nasz język nie sprawdza się za dobrze* (Jerzy Bralczyk, Uniwersytet Warszawski).

Mama była Pažoutowi przez cały czas wierna, ale on mimo to był o nią nieustannie zazdrosny (…) MVPDK2005pl

Když se se Zuzanou brali, žárlil na každého muže, který ji oslovil. MVUZ2001cz

Gdy się pobierali, był zazdrosny o każdego mężczyznę, który się do niej odezwał. MVW2008pl

W niektórych zdaniach jednak znaczenie pozostaje zamazane i bez szerszego kontekstu nie sposób poprawnie zrozumieć sytuację:

Zdálo se mi, že na paní učitelku žárlí a přišlo mi jí líto. IDHB1998cz

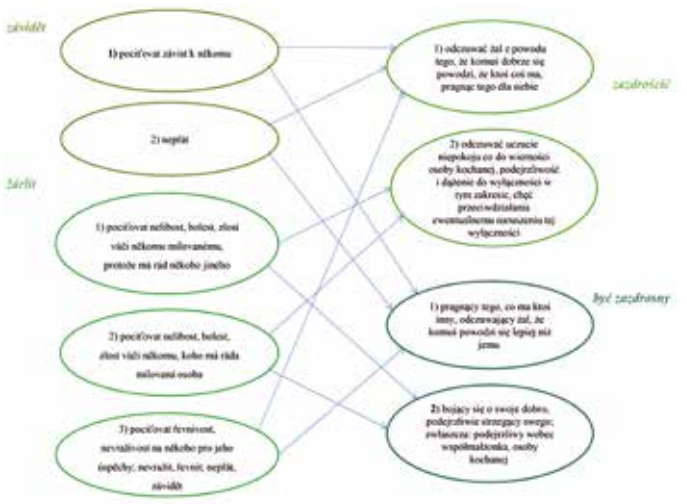
Wydawało mi się, że jest o panią nauczycielkę zazdrosna, i było mi jej trochę żal. IDB2003pl

Wnioski

Analiza jednostek *závidět* i *žárlit* oraz ich polskich odpowiedników wyekscerpowanych z InterCorpu pozwala na stworzenie sieci znaczeń[209]. Schemat ten odzwierciedla tylko sposób rozumienia i przekładu czeskich jednostek na język polski (dlatego strzałki skierowane są tylko w jedną stronę). Wzięte pod uwagę zostały tylko cztery analizowane jednostki; autorka zdaje sobie jednak sprawę z tego, iż na kompleksowej mapie znaczeń powinny się znaleźć również takie potencjalne ekwiwalenty jak np. *zawiść*. Równoważności znaczeń zostały ustalone na podstawie manualnej analizy poświadczeń korpusowych[210].

[209] Sieć powstała na podstawie definicji ze słowników jednojęzycznych: dostępnego online języka czeskiego –http://ssjc.ujc.cas.cz/search.php?hledej=Hledat&heslo=z%C3%A1vid%C4%9Bti&sti=EMPTY&where=hesla&hsubstr=no http://ssjc.ujc.cas.cz/search.php?hledej=Hledat&heslo=%C5%BE%C3%A1rliti&sti=EMPTY&where=hesla&hsubstr=no (dla czasowników *závidět* i *žárlit*) oraz w słowniku tradycyjnym (Szymczak 1995) dla jednostek *zazdrościć* i *być zazdrosnym*.

[210] Czeskie oryginały w przekładzie na język polski; w sumie 109 przykładów w języku czeskim i tyleż w języku polskim.



Obraz 21. Sieć powiązań znaczeń czeskich jednostek *závidět* i *žárlit* oraz polskich – *zazdrościć* i *być zazdrosny*

Podczas analizy manualnej szczególną uwagę zwróciły schematy *zazdrosny o* + O (obiekt), pojawiające się jako odpowiednik tylko jednej frazy: *žárlit na* + O. Ze znalezionych w korpusie 59 poświadczeń czasownika *žárlit*[211] 45 przekładanych jest jako *być zazdrosny*, przy czym tylko 12 realizowanych było poprzez pełny schemat *być zazdrosny o* + O; wśród nich 11 dotyczyło zazdrości, której obiektem była istota ludzka[212], a tylko w jednym przypadku obiekt był inny:

Tereza přijala Karenina takového, jaký byl, nechtěla ho měnit ke svému obrazu, souhlasila předem s jeho psím světem, nechtěla mu ho brát, nežárlila na jeho tajné spady. MKNLB1985cz

Teresa przyjęła Karenina takiego, jakim był, nie chciała go zmieniać na swoje podobieństwo, z góry godziła się na jego psi świat, nie chciała mu go odbierać, nie była zazdrosna o jego tajemne ścieżki. MKNLB1992pl

Analizę można by wzbogacić badaniem wykorzystującym narzędzie Sketch Engine; ukazałoby to, z jakimi typami obiektów (i w jaki sposób pod względem składniowym) łączą się omawiane jednostki. Podobnie jak w przypadku jednostek omawianych w poprzednim rozdziale możliwe jest tu tylko badanie eksperymentalne; wygenerowane tabele profili kolokacyjnych znajdują się w Załączniku 2.

[211] Teksty oryginalnie czeskie w przekładzie na język polski.

[212] 7 razy obiektem zazdrości była kobieta (*přiznal se jí konečně doma, že na ni žárlil / przyznał się wreszcie w domu, że był o nią zazdrosny*), 4 razy mężczyzna (*nemusím na jejího muže žárlit / nie muszę być zazdrosny o jej męża*).

Przypadek 5 – *mít rád* i *milovat*

Kolejną analizowaną jednostką była fraza *mít rád*[213]. W trakcie tego badania podjęto próbę odkodowania jej znaczenia przez pryzmat języka polskiego[214]. Pełne zrozumienie frazy (treści i wartości), które wyraża oryginał, jest fundamentalne dla poszukiwania ekwiwalentu w języku docelowym; stanowi też pierwszy z etapów przekładu[215]. Definicja tej jednostki (przytoczona w całości poniżej) ukazuje, iż *mít rád* to dla osoby polskojęzycznej kombinacja znaczeń 'kochać' i 'lubić' i jako taka reprezentuje pojęcie budzące niezrozumienie. Analizę frazy *mít rád* komplikuje dodatkowo fakt, iż dla wyrażenia znaczenia 'kochać' język czeski posiada jeszcze inny czasownik – *milovat*. Z tego powodu warto analizować jednocześnie obydwie jednostki.

Definicja słownikowa

mít rád

W słowniku języka czeskiego (Havránek 1989) fraza *mít rád*[216] jest definiowana jako 1) 'pociťovat k někomu náklonnost, lásku'[217]; 2) 'milovat'[218]; 3) 'mít v oblibě'[219].

[213] Na temat *mít rád* jako elementu nie w pełni przetłumaczalnego na język polski także w innych pracach (Kaczmarska 2014A, 2015A, 2015B; Kaczmarska i Rosen 2014A).

[214] W rozdziale tym próbowano wykorzystać elementy gramatyki kognitywnej (Langacker 1987, 1991, 2008; Mikołajczuk 1997, 1999; Taylor 2002).

[215] W procesie tłumaczenia wyróżnia się na ogół trzy etapy: analiza, transfer i restrukturyzacja (Nida i Taber 1974: 200).

[216] http://ssjc.ujc.cas.cz/search.php?hledej=Hledat&heslo=r%C3%A1d&sti=EMPTY&where=hesla&hsubstr=no

[217] Znaczenie to oznaczone jest też jako $milovat_2$ i $milovat_1$.

[218] Znaczenie to oznaczone jest też jako $milovat_3$.

[219] Pol. 1) 'czuć do kogoś sympatię miłość'; 2) 'kochać'; 3) 'lubić' – tłumaczenie własne autorki.

milovat

SSJC przedstawia dwa znaczenia czasownika *milovat*[220]: 1) 'pociťovat k někomu mileneckou lásku'; 2) 'pociťovat neerotickou lásku, mít rád'; 3) 'mít v oblibě, dbát o někoho, o něco'.[221] Autorzy słownika w drugim znaczeniu umieszczają wartość *mít rád*; pojawianie się obu jednostek (*mít rád* i *milovat*) we wzajemnych definicjach wskazuje na fakt, iż w niektórych kontekstach są to rzeczywiście synonimy.

Słownik dwujęzyczny

mít rád

Słownik czesko-polski (Siatkowski i Basaj 2002) przedstawia polskie ekwiwalenty: *kochać*, *lubić*, *przepadać*[222]. Odpowiedniki te odnoszą się jednak do zupełnie różnych uczuć. Jak już zostało wspomniane, z punktu widzenia osoby polskojęzycznej połączenie znaczeń 'kochać' i 'lubić' w ramach jednego wyrażenia jest pojęciem nieznanym.

milovat

W słowniku czesko-polskim (Siatkowski i Basaj 2002) leksem *milovat* ma trzy ekwiwalenty: *kochać*, *miłować*, *mieć zamiłowanie.* Podobnie jak w przypadku frazy *mít rád* również i ekwiwalenty czasownika *milovat* proponowane przez słownik czesko-polski odpowiadają znaczeniom wskazanym przez słownik języka czeskiego.

Analiza korpusowa

mít rád

Analizę korpusową rozpoczyna automatyczna ekscerpcja ekwiwalentów za pomocą serwisu Treq[223]; korpus paralelny oferuje szerszy niż słownik wachlarz

[220] http://ssjc.ujc.cas.cz/search.php?hledej=Hledat&heslo=milovati&sti=EMPTY&where=hesla&hsubstr=no

[221] Pol. 1) 'czuć do kogoś miłość erotyczną'; 2) 'czuć miłość platoniczną'; 3) 'lubić, dbać o kogoś, o coś' – tłumaczenie własne autorki.

[222] Czasowniki *lubić* i *przepadać za czymś* można traktować w zasadzie jako synonimy, różniące się nieco intensywnością i stylem.

[223] Tabela została wygenerowana w oparciu o dane z wersji 11 InterCorpu.

odpowiedników frazy *mít rád*. W poniższej tabelce przedstawione zostały polskie ekwiwalenty tej frazy wraz z liczbą poświadczeń.

Tabela 15. Najczęstsze polskie ekwiwalenty czasownika *mít rád* na podstawie serwisu Treq

Najczęstsze polskie ekwiwalenty czasownika ***mít rád*** na podstawie serwisu **Treq**	
1351	*lubić*
574	*kochać*
24	*polubić*
16	*uwielbiać*
9	*pokochać*
11	*przepadać*
4	*sympatia*
4	*preferować*
4	*chcieć*
3	*zależeć*
3	*cenić*
2	*wielbić*

Wszystkie odnalezione odpowiedniki należą w zasadzie do dwóch pól semantycznych odnoszących się do pojęcia 'kochać' i 'lubić'.[224] W wielu kontekstach jest możliwe wskazanie, o które znaczenie chodzi; na przykład w poniższych przykładach jako ekwiwalent można zaakceptować w zasadzie zarówno *lubić, przepadać*, jak i *kochać*:

Katolíci mají rádi mrtvoly. JDB1993cz
Katolicy kochają trupy. JDZ1959pl

On má rád milosrdné lidi. JDB1993cz
On lubi miłosiernych ludzi. JDZ1959pl

Pan Míla má rád fyziku, elektriku, stroje (…) LFSM1983cz
Pan Miłek lubi fizykę, elektryczność, maszyny (…) LFPZ1979pl

[224] Wśród odpowiedników mogą znaleźć się również jednostki związane błędnie i ekwiwalenty innych znaczeń poszczególnych komponentów tej jednostki, np. *mít raději, mít radši* oznacza w języku polskim –*preferować* i ten czasownik również znalazł się w wygenerowanym słowniku. W porównaniu jednak z poprzednią edycją Trequ aktualna generuje dużo dokładniejsze listy, por. tabelki z *mít rád* i *milovat* (Kaczmarska 2015B).

Istnieją jednak przykłady, które sprawiają problemy. Chodzi o sytuacje, w których tłumacz jest zmuszony do użycia jednego z czasowników – *lubić* albo *kochać*. Pozbawia tym samym frazę oryginalną jej wieloznaczności i może doprowadzić do zmiany intencji tekstu źródłowego. Np.:

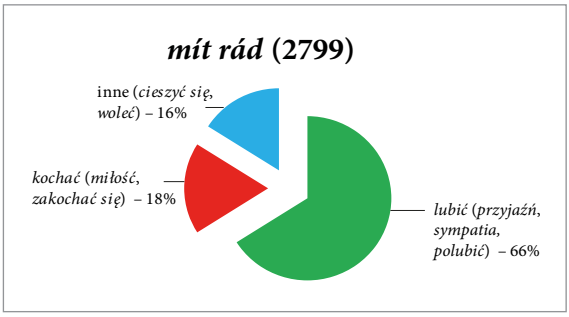
(…) *takže si vysvětlovala jeho zamlklost tím, že ji přestal mít rád.* MKN1993cz
(…) *jego milczenie interpretowała zatem jako dowód na to, że już jej nie kocha.* MKN2008pl

Já se neskrývám tím, že tě mám rada. MKN1993cz
Ja nie ukrywam miłości do ciebie. MKN2008pl

Ale pak se zdálo, že mě má rád. JDB1993cz
Ale potem wydawało mi się, że mnie kocha. JDZ1959pl

Niebezpieczeństwo zmiany znaczenia jest duże, dlatego warto na początek przeprowadzić badanie statystyczne, które wskaże, jak często analizowana fraza *mít rád* poświadczona jest w korpusie równoległym z ekwiwalentem *lubić* (oraz jego synonimami), a jak często występuje z przekładem *kochać* (wraz z synonimami).

Badanie to przeprowadzono wyłącznie na podstawie czeskich oryginałów tekstów beletrystycznych tłumaczonych na język polski. Wyekscerpowano 2799 paralelnych segmentów z jednostką ***mít rád***. Jak wynika z analizy, czasownik *lubić* (i jego synonimy) występuje ponad trzykrotnie częściej (~66%) niż czasownik *kochać* (wraz z synonimami) jako odpowiednik czeskiej frazy *mít rád* (~18%). Wynik ten prezentuje poniższy wykres:



Obraz 22. Rozkład liczbowy i procentowy, ukazujący, jak tłumaczona jest na język polski czeska jednostka *mít rád*

Rezultat badania statystycznego nie może jednak przesądzić o wyborze tłumacza; stojący wobec przekładu jednostki *mít rád* na język polski ma do wyboru czasowniki *kochać* i *lubić* (oraz jednostki wobec nich bliskoznaczne). Fraza

mít rád reprezentuje znaczenie z pogranicza pól semantycznych i jako takie jest dla języka polskiego pojęciem obcym (Kaczmarska 2014A, 2015B; Kaczmarska i Rosen 2014A). Nic więc dziwnego, że w przekładach pojawiają się oba odpowiedniki:

Víte, pane Dvořák, já ***mám rád*** *rodinu a dělám pro ni všechno.* LFSM1983cz
Wie pan, panie Dworzak, jak ***kocham*** *rodzinę i że robię dla niej wszystko.* LFPZ1979pl

Mám *tě strašně* ***rád****, řekl.* KVNR1997cz
Strasznie cię ***kocham*** *– rzekł.* MKWP2001pl

Tak tě ***mám ráda****.* VRR1944cz
Takim cię ***lubię****.* VRK1954pl

Já ***mám*** *podzim* ***ráda****.* IDHB1998cz
Ja ***lubię*** *jesień.* IDB2003pl

Jeżeli dane pojęcie nie istnieje w języku docelowym, próba odtworzenia go w tym języku zawsze jest związana ze stratą części znaczenia lub jego znaczną modyfikacją[225]. Tłumaczenie czeskiej jednostki na język polski wiąże się z wyborem któregoś z ekwiwalentów, a tym samym z wyborem jednego z odcieni znaczeniowych danego czasownika i ze stratą pozostałych. Pozbawiając przekład specyficznej kombinacji znaczeń oryginału, czyni się go niedokładnym. Ponieważ dosłowne tłumaczenie jest niemożliwe, można próbować ustalić, czy istnieją jakieś czynniki formalno-leksykalne warunkujące dobór najlepszego w danej sytuacji ekwiwalentu, a tym samym minimalizujące problem niedosłowności (Kaczmarska i Rosen 2016). W tym konkretnym przypadku warto też przyjrzeć się bliżej jednostce *milovat*, konkurującej z frazą *mít rád* w wyrażaniu znaczenia 'kochać', a także przeanalizować najczęstsze polskie ekwiwalenty tego czasownika.

milovat

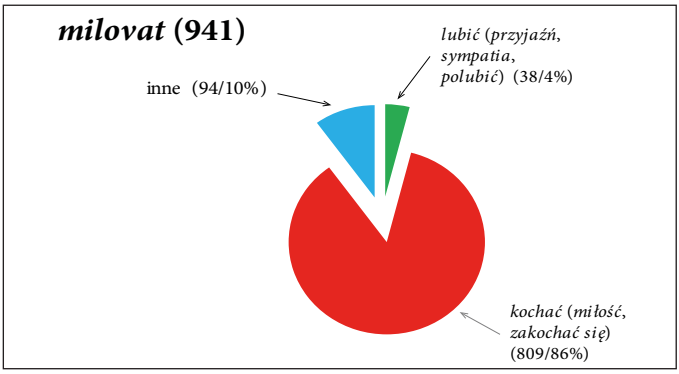
Również w przypadku tej jednostki na podstawie serwisu Treq ustalono, jakie jej ekwiwalenty pojawią się w korpusie InterCorp.

[225] Problem straty części znaczenia lub jego modyfikacji dotyczy także innych jednostek. W języku polskim trudno wskazać trafny odpowiednik również np. dla czeskiego czasownika *toužit*; jego ekwiwalentami w języku polskim są: *tęsknić, pragnąć marzyć* (Siatkowski i Basaj 2002). Czasownik *toužit* omawiany jest w rozdziale 12.

Tabela 16. Najczęstsze polskie ekwiwalenty czasownika *milovat* na podstawie serwisu Treq

Najczęstsze polskie ekwiwalenty czasownika ***milovat*** na podstawie serwisu **Treq**	
3075	*kochać*
142	*uwielbiać*
127	*lubić*
88	*miłość*
87	*pokochać*
53	*miłować*
43	*zakochany*
36	*kochany*
16	*ukochany*
16	*przepadać*
5	*ubóstwiać*

Lista ekwiwalentów wygenerowana dzięki Treq ukazuje, iż najczęstszym ekwiwalentem czeskiej jednostki *milovat* jest czasownik *kochać* (oraz derywaty od niego tworzone). Przewagę czasownika *kochać* jako ekwiwalentu potwierdza również wykres ilustrujący badanie przeprowadzone na 941 poświadczeniach z InterCorpu[226]:



Obraz 23. Rozkład liczbowy i procentowy, ukazujący, jak tłumaczona jest na język polski czeska jednostka *milovat*

[226] Ponownie, podobnie jak w przypadku frazy *mít rád*, analizowane były oryginalne czeskie teksty beletrystyczne, tłumaczone na język polski, z wersji 11 InterCorpu.

Otrzymane wyniki niejako ilustrują przewidywane dążenie rodzimych użytkowników języka polskiego do odwzorowania w języku obcym schematu z własnego języka[227]. W tej sytuacji można by automatycznie przypisać jednostce *milovat* znaczenie 'kochać', a frazie *mít rád* – znaczenie 'lubić'. Warto jednak sprawdzić również znaczenia obu polskich jednostek.

kochać* i *lubić

Słownik języka polskiego PWN[228] przedstawia następującą definicję czasownika *kochać*:

1) 'darzyć kogoś uczuciem miłości';
2) 'bardzo lubić kogoś';
3) 'żywić do osoby płci odmiennej gorące uczucie połączone z pożądaniem'.

W jednym ze znaczeń pojawia się druga analizowana polska jednostka – *lubić*, której definicję słownik języka polskiego przedstawia następującą[229]:

1) 'czuć do kogoś sympatię';
2) 'znajdować w czymś przyjemność';
3) 'o roślinach, zwierzętach, rzeczach: wymagać, potrzebować czegoś'.

Wymienione znaczenia jednostki *lubić* nie zawierają elementu miłości, warto więc rozszerzyć analizę o sprawdzenie, jak polskie czasowniki *kochać i lubić* są przekładane na język czeski. Prezentują to poniższe tabelki wygenerowane na podstawie Treq:

[227] W języku polskim 'lubić' i 'kochać' oznaczają zupełnie inne uczucia. Rożni je przede wszystkim inna jakość, a nie intensywność przeżywania.

[228] Wykorzystana została wersja elektroniczna, dostępna na stronie: http://sjp.pwn.pl (dostęp 29.04.2018).

[229] http://sjp.pwn.pl/szukaj/lubi%C4%87.html (dostęp 30.04.2018)

Tabela 17. Najczęstsze czeskie ekwiwalenty czasownika ***lubić*** na podstawie serwisu Treq

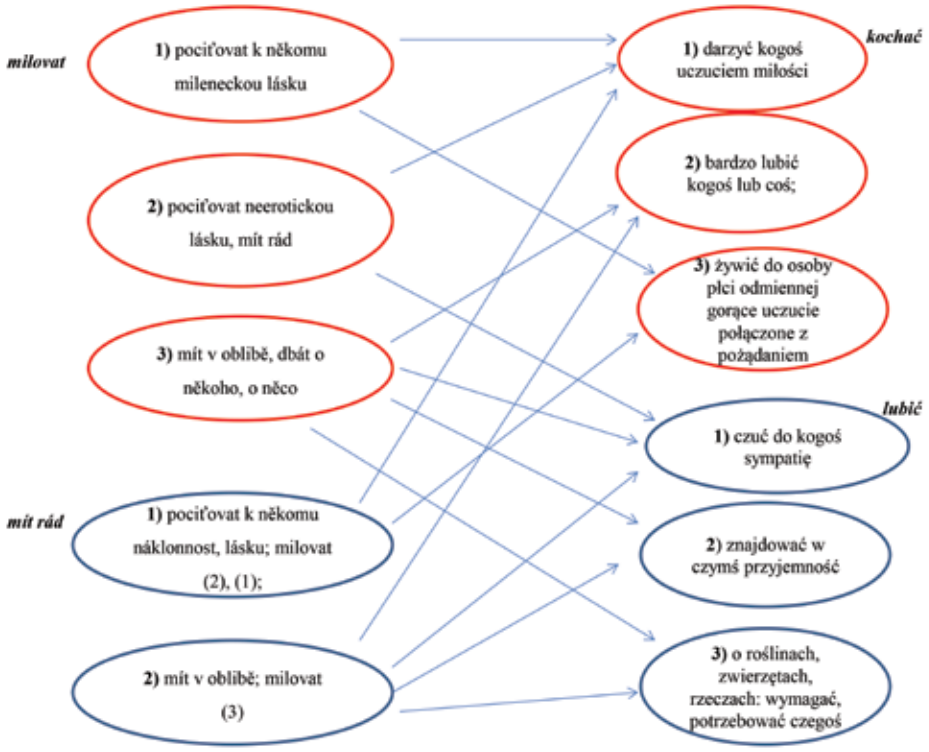
Najczęstsze czeskie ekwiwalenty czasownika ***lubić*** na podstawie serwisu **Treq**	
3116	*rád*
315	*líbit*
127	*milovat*
62	*oblíbený*
60	*snášet*
48	*obliba*
44	*bavit*
32	*potrpět*
31	*oblíbit*
28	*zamlouvat*
24	*láska*
20	*libovat*
18	*těšit*
18	*sympatický*
11	*chuť*
7	*vadit*
6	*zbožňovat*
5	*záležet*

Tabela 18. Najczęstsze czeskie ekwiwalenty czasownika *kochać* na podstawie serwisu Treq

Najczęstsze czeskie ekwiwalenty czasownika ***kochać*** na podstawie serwisu **Treq**	
2779	*milovat*
840	*rád*
181	*miláček*
130	*láska*
72	*drahoušek*
71	*milování*
60	*milující*
54	*zamilovaný*
46	*zamilovat*
37	*pomilovat*
30	*drahý*
27	*zlato*
21	*milovaný*
17	*zlatíčko*
11	*milánek*
11	*zbožňovat*
11	*duše*
8	*milý*
7	*líbit*

Przykłady z korpusu ukazują, iż jednostki czeskie i polskie posiadają po kilka znaczeń, przy czym mogą odnosić się do uczuć zarówno związanych z miłością, jak i niezwiązanych z nią (Głąbska 2014: 46–47). Poniższa mapa pokazuje przecinanie się znaczeń w przypadku tych czterech jednostek[230]:

[230] Mapa powstała na podstawie definicji ze słowników jednojęzycznych dostępnych online: http://ssjc.ujc.cas.cz/search.php?hledej=Hledat&heslo=r%C3%A1d&sti=EMPTY&where=hesla&hsubstr=no; http://ssjc.ujc.cas.cz/search.php?hledej=Hledat&heslo=milovati&sti=EMPTY&where=hesla&hsubstr=no; http://sjp.pwn.pl/szukaj/lubi%C4%87.html; http://sjp.pwn.pl/szukaj/kocha-%C4%87.html. Odzwierciedla tylko sposób rozumienia i przekładu czeskich jednostek na język polski (dlatego strzałki skierowane są tylko w jedną stronę). Wzięte pod uwagę zostały tylko cztery opisywane jednostki; autorka zdaje sobie jednak sprawę z tego, iż na kompleksowej mapie znaczeń powinny się znaleźć również takie potencjalne ekwiwalenty jak np. *uwielbiać*, a także derywaty. Jednostki *kochać* i *lubić* traktowane są tu jako reprezentanci dla grup ekwiwalentów mogących mieć cechę *miłość* lub *sympatia*. Równoważności znaczeń zostały ustalone na podstawie manualnej analizy poświadczeń korpusowych (czeskie oryginały w przekładzie na język polski).



Obraz 24. Sieć powiązań znaczeń czeskich jednostek *mít rád* i *milovat* oraz polskich – *kochać* i *lubić*

Dalsze badania potwierdzają złożony charakter tej ekwiwalencji. Przeprowadzono również badania eksperymentalne.

Eksperymenty

Ankieta

Na podstawie dotychczasowych wyników badań stwierdzić można, iż więcej argumentów przemawia za tłumaczeniem czeskiej jednostki *mít rád* jako *lubić* niż jako *kochać*. Ponieważ czasownik ten jest bardzo częsty i bardzo problematyczny, wśród rodzimych użytkowników języka czeskiego została przeprowadzona ankieta mająca ustalić, czy problem rozumienia fraz *mít rád* i *milovat* dotyczy

w jakimś stopniu również ich[231]. Autorka pytała w ankiecie, 1) czy *mít rád* i *milovat* to jednostki synonimiczne, 2) czy istnieją pomiędzy nimi jakieś różnice znaczeniowe, 3) z jakimi obiektami łączą się te jednostki. Wyniki ankiety przedstawione są w poniższej tabeli:[232]

Tabela 19. Obiekty łączące się z jednostkami *milovat* i *mít rád* według respondentów

	milovat	***mít rád***
obiekty	osoba, jedzenie, muzyka, aktywność, przyroda, zwierzęta	osoba, jedzenie, muzyka, aktywność, piwo, przyroda, rodzice, dziewczyna, życie, zwierzęta
definicja	najwyższy stopień miłości; 'mít rád' ale intensywniej; być w głębokiej relacji; czuć 'love'[20]; coś więcej niż 'mít rád'; bardzo silne pozytywne emocje;	lubić kogoś czy coś; pozytywne emocje;

Wszyscy respondenci odpowiedzieli, iż pomiędzy analizowanymi dwoma jednostkami są różnice semantyczne. Na uwagę zasługuje jednak fakt, iż nie potrafili wskazać jakościowej różnicy pomiędzy tymi dwoma jednostkami; w 100% zgodzili się jednak, iż *milovat* to po prostu „coś więcej" niż *mít rád*[233]. Respondenci mieli jednak problem z dywersyfikacją obiektów łączących się z tymi jednostkami[234].

Czeskie fora internetowe

Niejednoznaczność odpowiedzi ankietowych wzbudza podejrzenie, iż *mít rád* może być sprawiać problemy również rodzimym użytkownikom języka czeskiego. Na czeskich forach internetowych odnaleziono dużą liczbę postów, których

[231] Ankieta przeprowadzona była trzykrotnie (w grudniu 2013, w lutym 2014 i w styczniu 2015 roku); przebadanych zostało w sumie 150 rodzimych użytkowników języka czeskiego w wieku od 19 do 60 lat. Byli to studenci i wykładowcy Technicznego Uniwersytetu w Libercu i Uniwersytetu Karola w Pradze.

[232] Element 'love' – oryginalnie z czeskiej ankiety.

[233] Być może czynnikiem warunkującym ustalenie adekwatnego ekwiwalentu w odniesieniu do *mít rád* / *milovat* jest również płeć osoby wypowiadającej te słowa. Jest to aspekt, który także warto zbadać, jednakże nie jest przedmiotem analizy w tym opracowaniu.

[234] Obiekty w większości powtarzają się w obu kolumnach tabelki, co oznacza, że według respondentów jednostki te mogą się łączyć z takimi samymi obiektami.

autorzy rozważają znaczenie *mít rád* jako wyznania[235]; nie mają natomiast wątpliwości co do znaczenia czasownika *milovat* (*miluju tě*) użytej jako deklaracja[236]. Większość z postów potwierdza przeświadczenie respondentów o tym, że *mít rád* wyraża uczucie mniej intensywne niż *milovat*[237]:

Ovšem mít rád – to člověk může mít knihu, kamaráda, psa... v tom není nic erotického[238].

Vidím to přesně jako Arnika. Pro mě byly hranice teda vždy jasný. Zamilovaná jsem byla ze začátku do současného přítele. Už jsme spolu několik let, ale pořád ho miluju. Ráda mám třeba ex, se kterým jsme se rozešli už před 5 lety, ale v dobrém. Takže asi takhle – zamilovanost zezačátku, miluju někoho potom, co prvotní zamilovanost přešla. A ráda mám kamarády, blízké atd[239].

Ono mít ráda můžu i rajskou, ale milovat... je prostě něco jiného[240].

Tylko w jednym przypadku autorka wpisu utożsamiła czasownik *milovat* z wyrażaniem miłości fizycznej, a jednostkę *mít rád* – z uczuciem głębokiej miłości, wsparciem, zrozumieniem:

Miluji tě – má jistý sexuální náboj. Milenci po setkání odhazují oblečení, cesta vede směrem k ložnici. Je v tom touha, láska, zamilovanost a chtíč. Pro dnešek, zítřek, rok, snad dva. Méně citu a porozumění. Mám tě rád – je v tom všechno: cit, porozumění, láska, podpora. Že se jeden na druhého může spolehnout, budou spolu, až jim bude ouvej. Nebudou nikdy sami. Je to jako v němém filmu, kdy není třeba slov, protože

[235] Należy tu podkreślić, iż wszystkie wpisy rozpoczynające dyskusję na ten temat pochodziły od kobiet. Jeśli w dyskusji uczestniczył mężczyzna, występował w roli osoby komentującej wcześniejsze wpisy.

[236] http://diskuse.doktorka.cz/mit-rad-zamilovat-se-milovat/, http://www.poradte.cz/spolecnost/21684-milovat-nebo-mit-rad.html, http://janajerabkova.blog.idnes.cz/c/194377/Milovat-nebo-mit-rad.html - - dostęp 25.08.2014 Por. (Kaczmarska 2014B, 2015B).

[237] Polskie wersje wypowiedzi z Internetu – tłumaczenie autorki. Jako odpowiednik frazy *mít rád* został tu wybrany czasownik *lubić*. Zwłaszcza ostatni z przykładów ilustruje trudność przekładu tej czeskiej frazy na język polski.

[238] 'Owszem lubić – człowiek może książkę, kolegę, psa… nie ma w tym nic erotycznego' – http://diskuse.doktorka.cz/mit-rad-zamilovat-se-milovat/archiv/0/; dostęp 25.08.2014

[239] 'Widzę to dokładnie tak jak Arnika. Dla mnie granica była zawsze jasna. Zakochana byłam na początku w obecnym partnerze. Jesteśmy razem już kilka lat, ale ciągle go kocham. Lubię na przykład ex, z którym rozstaliśmy się już przed 5 laty, ale w dobrej atmosferze. Także chyba jest to tak – zakochanie na początku, kocham kogoś potem, kiedy pierwsze zakochanie mija. A lubię przyjaciół, bliskich, itd.' – http://diskuse.doktorka.cz/mit-rad-zamilovat-se-milovat/; (dostęp 25.08.2014).

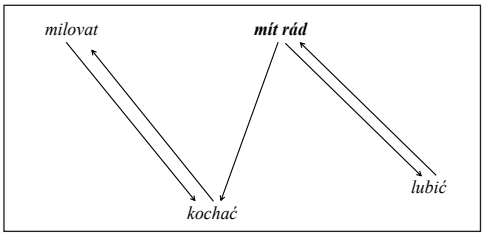
[240] 'Lubić to mogę i zupę pomidorową, ale kochać… to jest po prostu coś innego' – http://www.poradte.cz/spolecnost/21684-milovat-nebo-mit-rad.html; (dostęp 25.08.2014).

hovoří oči, činy. V nich se zobrazuje láska, něha, starost. Mám tě rád už není jen o slovech. Je to o životě[241].

Również rezultaty tej części analizy nie były wiążące. W tej sytuacji wydaje się, iż wybór ekwiwalentu zależy od szerszego kontekstu. Jednakże jak wynika z przeglądu forów internetowych również rodzimi użytkownicy języka czeskiego dyskutują, co dokładnie znaczy *mít rád* (zwłaszcza jako deklaracja). W rezultacie należy stwierdzić, iż nie w każdym przypadku problem z tłumaczeniem *mít rád* może być rozwiązany przez szerszy kontekst.

Wnioski

Czasownik *milovat* nie sprawia trudności w przekładzie; w języku polskim ma swój doskonały ekwiwalent, którym jest czasownik *kochać*.



Obraz 25. Ekwiwalencja czasowników: *mít rád, milovat / lubić, kochać*

W języku polskim nie istnieje natomiast jednostka, która oddawałaby emocje z pogranicza miłości i sympatii i jednocześnie zawierałaby oba te uczucia, ponieważ nie istnieje w tym języku takie pojęcie (koncept). W wielu przypadkach nie można ustalić, jaki jest adekwatny odpowiednik jednostki. Konfrontując język czeski i polski, doświadcza się w tym przypadku niezrozumienia. Również ankiety i poszukiwania w forach okazały się pomocne tylko w małym stopniu. W procesie przekładu frazy *mít rád* to często tłumacz decyduje, czy w tekście docelowym chodzi o miłość, czy jedynie o sympatię.

[241] 'Kocham cię – ma w sobie pewien ładunek seksualny. Kochankowie, spotykając się, zrzucają ubrania, droga prowadzi w kierunku sypialni. Jest w tym tęsknota, miłość, zakochanie i pożądanie. Na dziś, jutro, rok, może dwa. Lubię cię – jest w tym wszystko – uczucie, zrozumienie, miłość, wsparcie. Ze jeden na drugim może polegać, będą razem na dobre i na złe. Nie będą nigdy sami. Jest to jak w niemym filmie, kiedy nie trzeba słów, ponieważ mówią oczy, czyny. W nich wyraża się miłość, czułość, troska. Lubię cię to nie tylko słowa. Tu chodzi o życie.' – http://janajerabkova.blog.idnes.cz/c/194377/Milovat-nebo-mit-rad.html; (dostęp 25.08.2014).

Perspektywy

Kwestia znaczenia jednostki *mít rád* i jej adekwatnego tłumaczenia na język polski nie została w tym rozdziale jednoznacznie rozwiązana. Poza analizą korpusową istnieją jednak inne metody, które warto wykorzystać w dalszych badaniach. Te można zacząć od próby implementacji metody odszukiwania powtarzalnych struktur, w których analizowana jednostka się pojawia. Mowa o Pattern Grammar (Francis, Hunston i Manning 1996; Hunston i Francis 2000; Ebeling i Ebeling 2013), dzięki której udało się np. automatycznie odróżnić dwa znaczenia innej czeskiej jednostki *být líto*: *żal*, *być przykro* (Kaczmarska 2015A: 139–140; Kaczmarska 2017A: 187–188). Wartościowe są też metody gramatyki kognitywnej, zwłaszcza, iż znane są prace dotyczące właśnie miłości (Głąbska 2014; Bierwiaczonek 2002). We wcześniejszych badaniach czyniono też próby wykorzystania obrazów kolokacyjnych Word Sketches, ale rozróżnienie znaczenia na podstawie łączliwości z obiektami w przypadku tych czasowników się nie powiodła (Kaczmarska 2015B).

Bez względu jednak na to, jak skomplikowane metody nie byłyby użyte, w niektórych przypadkach przyjdzie się zmierzyć z niemożnością oddania pewnych treści w języku docelowym. Kwestia niejasności znaczenia fraz *mám tě rád / mám tě ráda*, może przywołać na myśl pytanie: *Czy to jest przyjaźń, czy to jest kochanie?*[242]

[242] Adam Mickiewicz, *Niepewność*, Poezye Adama Mickiewicza, tom 1, Kraków 1899.

Projekt algorytmu ułatwiającego ustalanie ekwiwalentów

Przeprowadzone w poprzednich rozdziałach studia przypadków wykorzystują różne sposoby wyszukiwania ekwiwalentów na podstawie danych z korpusu równoległego InterCorp. Metody te można połączyć i zbudować algorytm ułatwiający ustalanie ekwiwalentów dla analizowanych w tym opracowaniu czasowników[243]. W rożnych ośrodkach na świecie prowadzi się prace nad algorytmami ulepszającymi tłumaczenie maszynowe; prace te w większości opierają się na olbrzymich danych uzyskanych z korpusów (również korpusów równoległych) i bazują na różnych metodach lingwistycznych i matematycznych (Tian, Wong i Chao 2010; Han i inni 2013; Tian, Wong, Chao i Oliveira 2014). Opracowywane algorytmy wykorzystują też różnicowanie znaczenia jednostek wieloznacznych – *Word Sense Disambiguation* (Młodzki, Kopeć i Przepiórkowski 2012; Kędzia, Piasecki, Kocoń i Indyka-Piasecka 2014). Wspomniane algorytmy dotyczą najczęściej języka angielskiego (w różnych konfiguracjach z innymi językami), języka chińskiego, rzadziej innych języków (Tufiş, Koeva, Erjavec, Gavrilidou i Krstev 2008; Sharp, Zock, Carl i Jakobsen 2011; Wu i inni 2016, 2018; Zennaki, Semmar i Besacier 2019)[244].

[243] Należy tu wspomnieć, iż prymarnie algorytm ten był pomyślany jako eksperyment poznawczy i propozycja procedury ustalania ekwiwalentów, mogąca służyć zarówno lingwistom, jak i tłumaczom, a nie jako program komputerowy.

[244] Niektóre systemy przekładu maszynowego (MT) osiągają dobre rezultaty, łącząc kilka metod, np. system *Chimera* (https://ufal.mff.cuni.cz/chimera), który w przekładzie z języka angielskiego na czeski pokonał w roku 2015 Google Translate, zawiera analizę tektogramatyczną, stochastyczny model przekładu oraz automatyczny system postedycji. W roku 2018 praski zespół swój sukces powtórzył, wykorzystując system oparty na sieciach neuronowych (Bojar i inni 2018; Popel i Bojar 2018).

W tym opracowaniu implementowano metody zaczerpnięte z różnych teorii językoznawczych[245] i sprawdzano zależność znaczenia od otoczenia oraz od charakterystyki składniowej (Levin 1993), a tym samym, na podstawie schematów składniowych, ustalano najbliższy ekwiwalent, odnosząc się również do struktury składniowej i otoczenia potencjalnego odpowiednika (Kaczmarska 2015A; Kaczmarska 2017A)[246]. Podobne ujęcie przedstawiał Kalisz (Kalisz 1981), który ekwiwalencję postrzegał jako odzwierciedlenie stopnia podobieństwa pod względem składniowym, semantycznym, pragmatycznym.

Niniejszy algorytm koncentruje się na próbie ustalania polskich ekwiwalentów dla czeskich jednostek.

Obraz 26. Schemat proponowanego algorytmu

[245] *Zwykle propozycje teoretyczne sformułowane są w ramach jakiegoś nurtu badań lingwistycznych, co jest oczywiście naturalne – wcześniejsze prace z tego samego nurtu dostarczają rusztowania, na którym łatwiej oprzeć własne propozycje. Zbyt często jednak badacze zamykają się w jednym paradygmacie, nie próbując sięgać do prac związanych z innymi szkołami lingwistycznymi, choć przecież zestawienie propozycji teoretycznych z różnych paradygmatów często pozwala lepiej zrozumieć ich ograniczenia i ich mocne strony, co w rezultacie zwykle prowadzi do dalszego rozwoju teorii* (Przepiórkowski 2017: 12).

[246] W ostatniej fazie przeprowadzono również eksperyment w oparciu o metody stochastyczne.

Cały algorytm składa się z kilku etapów analizy, przy czym w założeniu czasownik poddawany takiemu badaniu nie musi przechodzić przez wszystkie etapy; optymalny ekwiwalent może być odnaleziony na każdym poziomie analizy[247].

Krok pierwszy – ekwiwalenty słownikowe

To najbardziej oczywisty odcinek algorytmu; na drodze poszukiwania ekwiwalentów niemal zawsze pojawia się słownik. W tym opracowaniu wykorzystywane są słowniki jednojęzyczne (Havránek i inni 1989; Filipec, Daneš, Machač i Mejstřík 1994; Szymczak 1995) i dwujęzyczne (Oliva, 1994, 1995; Siatkowski i Basaj, 2002). Na podstawie definicji zawartych w tych słownikach ustalone zostają pierwsze ekwiwalenty; pomagają w tym sporządzane mapy znaczeń (jak w przypadku jednostek *mít rád* i *milovat* czy *závidět i žárlit*), gdzie łączone są poszczególne znaczenia danych jednostek.

Krok drugi – automatyczna ekstrakcja par ekwiwalentów

Automatyczną ekstrakcję można przeprowadzić na kilka sposobów. Najprostszym jest wygenerowanie listy ekwiwalentów za pomocą serwisu Treq, przy czym należy pamiętać o ograniczeniach doboru danych, z których czerpana jest lista ekwiwalentów. Innym sposobem jest samodzielne zastosowanie narzędzia GIZA^{++} (Och i Ney 2003)[248], który jest wolno dostępnym programem, wytworzonym w celu statystycznego przekładu maszynowego i zawierającym moduł wiązania segmentów na poziomie wyrazu – *word-to-word alignment* (Kaczmarska i Rosen 2013). Właśnie to wiązanie umożliwia stworzenie listy par ekwiwalentów, która tworzy swoisty słownik (Skoumalová 2008; Jirásek 2011).[249] Warto przypomnieć, iż GIZA^{++} może być stosowana również na pewnym zbiorze tekstów, ściśle określonej części korpusu, np. tekstach beletrystycznych z rozróżnieniem na kierunek przekładu.

Automatycznie wygenerowany słownik można wykorzystać jako uzupełnienie konwencjonalnego dwujęzycznego słownika[250]; dla użytkownika znającego

[247] Natomiast czasowniki, które nie odnalazły swoich ekwiwalentów na danym etapie, przechodziłyby do następnego.

[248] Narzędzie GIZA^{++} wykorzystywane jest również przez Treq.

[249] Próbę automatycznej ekstrakcji par ekwiwalentów w oparciu o dane z korpusów porównywalnych podjęli Darja Fišer i Nikola Ljubešić (Fišer i Ljubešić 2013).

[250] Por. też wykorzystanie korpusu równoległego jako źródła ekwiwalentów do słownika chorwacko-czeskiego (Jirásek 2011).

obydwa języki i poszukującego różnych znaczeń danego leksemu automatycznie wygenerowany słownik może mieć duże znaczenie (choćby w przypadku badań translatorskich czy frekwencyjnych) (Kaczmarska i Rosen 2013: 120). Tego typu słownik może okazać się również niezastąpionym narzędziem w badaniach niektórych leksemów np. zdrobnień, choć w tym przypadku korzystniejsze byłoby wykorzystanie sortowania a tergo (Rosen, Kaczmarska i Škodová 2014; Škodová, Rosen i Kaczmarska 2016; Škodová, Kaczmarska i Rosen 2017).

Zgromadzone w ten sposób potencjalne odpowiedniki stanowią niejako klaster ekwiwalentów (Kaczmarska 2017B); badania wykazują bowiem, iż często nie jest możliwe odnalezienie jednego jednowyrazowego czasownikowego ekwiwalentu pokrywającego wszystkie znaczenia (Lewandowska-Tomaszczyk 2010A, 2013; Kaczmarska 2014A, 2014B; Kaczmarska i Rosen 2013, 2014A). Zdarza się, iż na liście pojawia się cały szereg czasowników, różniących się znaczeniowo, ale pretendujących do miana ekwiwalentu danej jednostki; żaden z tych odpowiedników może nie być jednak tym stuprocentowym (Lewandowska-Tomaszczyk 1984, 2013)[251].

Wcześniejsze badania pokazują, iż niektóre czasowniki już na tym etapie odnajdują trafne ekwiwalenty; co ciekawe, są to często czasowniki wyrażające negatywne odczucia (Kaczmarska 2016). Analizowany we wcześniejszym rozdziale czasownik *závidět* znalazł swój odpowiednik *zazdrościć* już na liście wygenerowanej przez Treq, jednakże analiza wymagała porównania zakresów użycia wzmiankowanego czasownika i jednostki *žárlit*. W przypadku konieczności przejścia do następnych etapów algorytmu automatycznie wygenerowana lista hipotetycznych ekwiwalentów wskazuje, które z nich są najczęstsze i można traktować jako potencjalnie najtrafniejsze.

Krok trzeci – analiza walencyjna

Drugim krokiem jest ustalenie wymagań walencyjnych[252]. Uzasadnione wydaje się założenie, iż zbieżność struktur składniowych i podobieństwo semantyczne obiektów łączących się z danymi czasownikami może być kluczem do konkretyzacji znaczenia i wskazania ekwiwalentu na tym etapie (Levin 1993; Greń 1994).

[251] Według Barbary Lewandowskiej-Tomaszczyk takie stuprocentowe ekwiwalenty w zasadzie nie istnieją; można jednak znaleźć zbiór semantycznie bliskich jednostek (*a cluster of equivalents*) i wykorzystać je w procesie przekładu (Lewandowska-Tomaszczyk 1984, 2013).

[252] Krok ten ze względu na odniesienia do innych metod i teorii wydaje się kluczowym dla tego algorytmu, z tego powodu został opisany szerzej niż inne etapy.

Walencja jest znanym od dawna nurtem językoznawstwa. Powszechnie uznaje się, iż ojcem terminu *walencja* jest Lucien Tesnière (Tesnière 1959)[253]. Pisze o tym Danuta Rytel-Kuc (Rytel-Kuc 1989):

> *Za ojca teorii walencyjnej uważany jest* (...) *francuski językoznawca Lucien Tesnière, którego główna praca na ten temat pt. „Éléments de syntaxe strukturale" ukazała się w roku 1959. On też stworzył podstawową terminologię tego kierunku w lingwistyce, poczynając od samego terminu walencja, który odniósł do czasownika jako elementu centralnego w zdaniu. Tesnière podzielił czasowniki na awalentne, monowalentne, diwalentne i triwalentne, zależnie od ilości przyjmowanych przez nie elementów zależnych – aktantów, obok których wyróżnił elementy walencyjne niezwiązane – cirkumstanty.*

Jak jednak zauważają Michał Smułczyński (Smułczyński 2013: 173) i Adam Przepiórkowski (Przepiórkowski 2017: 20–21), idea wymagań składniowych pojawiła się także w opracowaniu Karla Bühlera z 1934 r., gdzie wspomina on o konotacji (Bühler 2004: 180)[254], natomiast o teorii łączliwości składniowej pisał w tym samym czasie Kazimierz Ajdukiewicz (Ajdukiewicz 1935)[255].

Sam termin walencja – *valence* – jak pisze Adam Przepiórkowski (Przepiórkowski 2017: 21), pojawia się jednak u Tesnière'a jeszcze przed rokiem 1959 r.:

> *Mało znany jest natomiast fakt użycia terminu valence przez Tesnière'a już w pracy Esquisse d'une Syntaxe Structurale (Tesnière, 1953: 9) oraz to, że sama metafora walencji, w znaczeniu bardzo zbliżonym do Tesnière'owskiego, pojawiła się w lingwistyce jeszcze co najmniej trzykrotnie przed wydaniem Elementów… (w tym dwukrotnie przed Esquisse…), a mianowicie w pracach: Hockett (Hockett, 1958) (ang. valence), de Groot (de Groot, 1949) (hol. valentie) oraz – najwcześniej – Kacnel'son (Kacnel'son, 1948) (ros. valentnost').*

Ruth Konvalinková choć stwierdza za Bühlerem (Bühler 1934), iż niektórym zjawiskom, opisywanym później przez teorie walencyjne, uwagę poświęcali już średniowieczni scholastycy, za pierwszego prawdziwego prekursora

[253] Przekład dzieła *Éléments de syntaxe structurale* (Tesnière 1959) na angielski – *Elements of Structural Syntax* (Tesnière 2015).

[254] Oryginalne dzieło Bühlera zostało opublikowane w 1934 r. w Jenie – *Sprachtheorie. Die Darstellungsfunktion der Sprache* (Bühler 1934). Karl Bühler opisuje walencję w szerokim znaczeniu jako zdolność wszystkich części mowy do otwierania pustych miejsc.

[255] Dostęp online do pełnego tekstu ze stron Biblioteki Śląskiej: https://www.sbc.org.pl/dlibra/show-content/publication/edition/21615?id=21615#dcId=1545929943481&p=1 (30.08.2018)

Na ostatniej stronie widoczny jest zapis po łacinie, świadczący o tym, że artykuł został złożony 15 lipca 1934 r. Artykuł ten dostępny jest też w wersji angielskiej – *Syntactic Connexion* (Ajdukiewicz 1967) w książce *Polish Logic, 1920–1939* (McCall 1967); tekst przekładu (tłum. H. Weber) odnosi się do języka angielskiego w przeciwieństwie do oryginału, gdzie Ajdukiewicz przytacza przykłady niemieckie.

walencji uważa Johanna Wernera Meinera (1723–1789), który sformułował myśl o centralnej roli predykatu w procesie organizacji zdania oraz zauważył swoistą niesamodzielność (niem *Unselbständigkeit*) niektórych typów słów (czasowników, przymiotników), czyli w języku współczesnej terminologii – ich walencję (Konvalinková 2011: 78). Posługując się stale językiem współczesnej lingwistyki, Ruth Konvalinková stwierdza, iż Johann Werner Meiner wskazał również różne typy predykatów i dokonał ich klasyfikacji z uwzględnieniem wymaganych przez nie uzupełnień. Vilmos Ágel uważa nawet pracę Meinera (*Versuch einer an der menschlichen Sprache abgebildeten Vernunftlehre* z 1781 r.)[256] za kamień milowy w rozwoju teorii walencyjnej i dowodzi jej nowoczesności (Ágel 2000: 21–23). Chociaż Meiner nie zdołał przeforsować swoich poglądów na temat centralnej roli predykatu, ponieważ w owym czasie dominującą była koncepcja zdania jako struktury dwuczłonowej o równym statusie podmiotu i orzeczenia, jego spostrzeżenia odbijają się w wielu współczesnych koncepcjach lingwistycznych, np. Lexical Functional Grammar (LFG), Head Driven Phrase Structure (HPSG) czy Principles and Parameters (P&P) (Konvalinková 2011: 78)[257].

Walencja jest zjawiskiem opisywanym w wielu opracowaniach, wspominając choćby relatywnie najnowsze (Čermáková 2009; Urbańczyk-Adach 2011; Skwarska i Kaczmarska 2016; Zaron i Skwarska 2017, 2018). Nie jest jednak walencja pojęciem ostrym. Jak pisze Zofia Zaron, myśląc o zależnościach semantyczno-syntaktycznych, a w szczególności o wymaganiach składniowych, należy mieć świadomość przebogatej literatury na ten temat i – niestety – niejednoznaczności terminów (Zaron 2012: 673–674):

> *W tym miejscu warto przywołać różnorakie koncepcje opisu zależności składniowych od tradycyjnych począwszy [por. np. w polszczyźnie (Klemensiewicz, 1937)] po nowsze związane z nazwiskami Bühlera (Bühler, 1934), Tesniére'a (Tesnière, 1959), Gołąba (Gołąb, 1967), Szwedowej (Szwedowa, 1970), Mielczuka [(Mielczuk, 1974, 1988) etc.], Karolaka (Karolak, Składnia wyrażeń predykatywnych, 1984), Bogusławskiego*

[256] W roku 2010 praca ta ukazała się jako faksymile (Meiner 2010).

[257] Teoria walencyjna bardzo dobrze przyjęła się w językoznawstwie niemieckim, gdzie głównie rozwinął się syntaktyczny kierunek badań walencyjnych. Na gruncie językoznawstwa polskiego teoria walencyjna nie znalazła jednak żywego oddźwięku; szerzej na ten temat w artykule Zbigniewa Grenia (Greń 2016). W roku 1967 ukazał się artykuł Zbigniewa Gołąba (Gołąb 1967), który przez długi czas był jedyną pozycją poświęconą temu zagadnieniu; później powstawały kolejne prace (Buttler 1976; Dębski 1982; Rytel 1989; Rytel-Kuc 1991). Teorie walencyjne znalazły swoje odbicie również w pracach dotyczących języka czeskiego i słowackiego (Ružička 1968; Daneš 1971; Daneš, Hlavsa i Kořenský 1973; Panevová 1975, 1980, 1994, 1998).

(Bogusławski 1978). Teorie te posługują się pojęciem konotacji / walencji / kompletywności / determinacji czy akomodacji wreszcie[258].

W sprawie doprecyzowania terminów *konotacja* i *walencja* głos zabierało wielu językoznawców. Natallia Yaumen wręcz postulowała zachowanie obu terminów, choć jak sama pisze, obydwa dotyczą tego samego zjawiska; termin *konotacja* rezerwuje jednak dla wymagań semantyczno-składniowych leksemów, natomiast terminem *walencja* określa „otwieranie miejsc" dla pozycji semantycznych (Yaumen 2013: 16). Kłóci się to z pojęciem konotacji w rozumieniu Bühlera, który właśnie konotację widzi jako zdolność wyznaczania tzw. pustych miejsc (Zaron 2012: 674)[259].

W tym opracowaniu autorka posługuje się terminem *walencja* rozumianym jako zjawisko językowe odnoszące się do liczby argumentów wymaganych przez predykat werbalny[260] wraz z morfosyntaktycznymi i semantycznymi cechami tych argumentów (Kaczmarska, Rosen, Hana i Hladká 2015: 153). Schematy walencyjne określane są zarówno dla czeskiego oryginału, jak i polskich potencjalnych odpowiedników, dzięki temu znaczenie analizowane jest dwukrotnie: raz przez pryzmat walencji w ramach jednego języka oraz drugi raz przez pryzmat ekwiwalentu. Przekład staje się więc niejako semantyczną anotacją, innym sposobem wyrażenia znaczenia oryginału[261]. Dzieje się to na dowolnym poziomie: segmentów, zdań, części zdania, fraz, słów; na tych poziomach w korpusie paralelnym wyraża się też związek pomiędzy ekwiwalentami przekładowymi, czyli wiązanie (ang. *alignment*, czes. *zarovnání*). Wiązane elementy dostarczają też wzajemnie o sobie informacji, a dane uzyskane na podstawie kontekstu o danym wyrażeniu w jednym języku można wykorzystać do anotacji jego ekwiwalentu w drugim języku[262].

---

[258] Już w przypisie Zofia Zaron odsłania tajemnicę powiązań nazwisk lingwistów i terminów (Zaron 2012: 674):

Por. pojęcie konotacji u Bühlera (Bühler 1934), a także konotacji formalno-gramatycznej – u Gołąba (Gołąb 1967), walencji – u Tesniére'a (Tesnière, 1959), kompletywności – u Bogusławskiego (Bogusławski 1978) czy akomodacji – u Karolaka (Karolak 1972) i u Saloniego, Świdzińskiego (Saloni i Świdziński 1985), u Szwedowej (Szwedowa 1970) zaś – determinacji. W opisach składniowych zachodnich, a także rosyjskich preferowany jest termin Tesniére'a walencja.

[259] Zofia Zaron (Zaron 2012: 674) wyjaśnia, iż jedna jednostka języka otwiera wokół siebie miejsca dla innych jednostek języka, które muszą spełnić dwa warunki, aby mogły to miejsce zająć: 1) przyjąć odpowiednią formę gramatyczną (o ile są odmienne), 2) mieć określone znaczenie (element znaczenia wspólny dla klasy wyrażeń, które mogą zająć wskazane miejsce).

[260] Por. (Daneš i Hlavsa 1987; Rytel 1989; Čermáková 2009; Urbańczyk-Adach 2011).

[261] *If we assume that we may find the meaning of a textual element through its paraphrase, which is also a text, then we may describe parallel corpora as repositories for such paraphrases* (Teubert 2001).

[262] Np. jeżeli w języku A pojawia się leksem wieloznaczny, a w wiązanym segmencie w języku B pojawia się ekwiwalentny leksem jednoznaczny, wówczas można ujednoznacznić w danym kontekście leksem w języku A.

W ten sposób w równoległych tekstach można uzupełniać anotację na różnych poziomach – od morfologii po składnię, z anotacją opartą na informacjach o walencji włącznie (Kaczmarska i Rosen 2016A: 321–323; Kaczmarska i Rosen 2016B: 157–159)[263].

Własności walencyjne leksemów są podstawą większości współczesnych teoretycznych i praktycznych podejść do składni – od formalnych teorii (Sgall 1998; Pollard i Sag 1994; Bresnan 2001), przez leksykografię (Kettnerová 2014), po komputerowe przetwarzanie języka naturalnego czy analizę sentymentalną (Šindlerová, Veselovská i Hajič jr. 2014; Taboada 2016)[264].

Wykorzystanie walencji w badaniu kontrastywnym i przy przetwarzaniu tekstów równoległych nie jest już tak powszechne, jednak korespondencja pomiędzy argumentami predykatu oryginału i ekwiwalentu może być wykorzystana do przygotowania słowników przekładowych, przekładu maszynowego, wyszukiwania w korpusie paralelnym oraz badań czysto teoretycznych (Dyvik, Meurer, Rosén i De Smedt 2009; Šindlerová i Bojar 2009, 2010)[265].

Wykorzystanie w algorytmie analizy wymagań walencyjnych jednostek oryginalnych (tu – czeskich) i jednostek w języku docelowym (tu – polskich) może ułatwić identyfikację ekwiwalentu przy założeniu, iż w niektórych przypadkach ekwiwalent może być ustalony na podstawie zbieżności wymagań walencyjnych. Autorka proponuje następującą procedurę postępowania:

- analiza manualna wybranych czasowników (wraz z wiązanymi segmentami),

[263] Zob. np. (Yarowsky i Ngai 2001; Padó i Lapata 2009).

[264] Szerzej na temat metod ekstrakcji informacji o walencji z anotowanych syntaktycznie tekstów w języku czeskim i polskim zob. np. (Bojar 2003; Benešová i Bojar 2006; Przepiórkowski 2009). Informacje o typowym wypełnieniu pozycji walencyjnych konkretnymi leksemami można uzyskać przy pomocy tzw. profili kolokacyjnych i leksykalnych (Word Skeches). Mowa o tym w dalszej części opracowania.

[265] W wyrównanym i językowo anotowanym korpusie równoległym można w idealnej sytuacji w przypadku każdego wyrażenia określić jego formę, anotację i kontekst (walencję) zarówno w języku źródłowym, jak i docelowym. Jeżeli nieznana jest anotacja w języku docelowym, można ją rzutować na podstawie języka źródłowego. Podobnie jeśli chodzi o niemożność ustalenia wyrażenia w języku docelowym, można go uściślić poprzez rzutowanie – wybór z możliwych ekwiwalentów wyrażenia w języku źródłowym na podstawie jego kontekstu. To rzutowanie informacji z jednego języka na drugi jest możliwe dzięki zastosowaniu różnych metod, które się wzajemnie uzupełniają, m. in. ręczna analiza paralelnych konkordancji, wyrównanie „słowo do słowa" z automatyczną ekscerpcją leksykalnych ekwiwalentów, automatyczna analiza składniowa, leksykalne i kolokacyjne profile (Bojar i Prokopová 2006). Na temat rzutowania anotacji z jednego języka na drugi także w pracach Romana Roszko (Roszko 2016; Koseska-Toszewa i Roszko 2016).

- określenie w każdym segmencie, jak dany czeski czasownik został przetłumaczony na język polski, ile argumentów wiąże i jakiego rodzaju są to argumenty[266] oraz w jaki sposób są z danym czasownikiem połączone,
- ustalenie, jakie argumenty łączą się z polskimi ekwiwalentami,
- syntaktyczna i semantyczna analiza argumentów wiążących się z czeskimi czasownikami i ich polskimi ekwiwalentami[267].

W badaniach pilotażowych analiza walencyjna była przeprowadzana jedynie częściowo (dotyczy to np. czasowników *trápit*, *soucítit*, *závidět*); bardziej szczegółowa analiza została przeprowadzona z udziałem czeskiego czasownika *toužit* w następnym rozdziale.

Teorie walencyjne znajdują swoje odniesienia także w teorii Fillmore'a (Fillmore 1968, 1977). Danuta Rytel-Schwarz przedstawia klasyfikację różnych typów walencji. Obok walencji syntaktycznej wyróżnia również walencję logiczną i semantyczną. Dwa ostatnie typy ściśle się łączą, tworząc model walencji logiczno-semantycznej. Danuta Rytel-Schwarz pisze, iż tak rozumiana walencja (logiczno-semantyczna) zlokalizowana poza językiem, na uniwersalnej płaszczyźnie logiczno-semantycznej była przedmiotem badań gramatyki przypadków głębokich Fillmore'a (Rytel 1989: 237–247). Fakt ten skłania do przypuszczenia, iż teoria Fillmore'a jest teorią bliską walencji rozumianej jako zdolność czasowników do otwierania wokół siebie jednego lub więcej pustych miejsc. Elementy (lub element), dla których czasownik otwiera miejsce, porównać można do argumentów w teorii Fillmore'a, które są wykładnikami powierzchniowymi przypadków głębokich. Właśnie teoria Fillmore'a wykorzystywana jest na kolejnym etapie tej analizy.

Krok czwarty – gramatyka przypadków głębokich

Trzecim krokiem jest identyfikowanie przypadków głębokich – roli, jakie pełnią elementy wiążące się z danym czasownikiem. Na tym poziomie relewantne

[266] Np. fraza nominalna – jeśli tak, to jakiego typu: obiekt realny, abstrakcyjny, nazywający istotę ludzką.

[267] Schematy walencyjne ustalone na podstawie ekscerpowanego z InterCorpu materiału porównywane są ze schematami ujętymi w słownikach walencyjnych: VALLEX (Kettnerová, Lopatková i Bejček 2012; Lopatková, Kettnerová, Bejček, Vernerová i Žabokrtský 2016), PDT-Vallex (Urešová, Štěpánek, Hajič, Panevová i Mikulová 2014), *Walenty* (Hajnicz, Patejuk, Przepiórkowski i Woliński 2016; Hajnicz i Nitoń 2017).

jest ustalenie, czy jednostki kandydujące na ekwiwalenty łączą się z obiektami reprezentującymi te same kategorie ról semantycznych[268].

Teoria przypadków głębokich Fillmore'a wychodzi naprzeciw konfrontatywnej analizie struktury semantycznej zdań[269]. Teoria ta zakłada istnienie pewnego poziomu semantycznego, na którym czasownik charakteryzowany jest poprzez frazy, z którymi się łączy. Frazy te to przypadki semantyczne[270], które pełnią określone funkcje, jednakże nie mogą być utożsamiane z powierzchnią[271]. Ośrodkiem zdania w tej teorii jest czasownik i tylko od niego zależy wybór odpowiednich przypadków. Liczba wyróżnianych przypadków nie jest stała; różnice pojawiają się także w samych pracach Fillmore'a[272].

Poszukiwania idealnego ekwiwalentu może ułatwić identyfikacja obiektu o tej samej wartości, łączącego się zarówno z badanym czasownikiem czeskim, jak i potencjalnym ekwiwalentem polskim. Na początku warto jednak zauważyć, iż ze względu na swój charakter wszystkie analizowane czasowniki wyrażające stany psychiczne wiążą argumenty o wartości Experiencer[273], np[274].

Nemohu říct, že by mě to nemrzelo. (MVPnK2003cz)
Nie powiem, żeby mi się to podobało. (MVSnK2007pl)

---

[268] Krok ten dotyczy sprawdzania zgodności przypadków semantycznych we frazach z badanym czasownikiem w języku wyjściowym i jego potencjalnym ekwiwalentem w języku docelowym; podstawę teoretyczną tego etapu analizy tworzyły prace Charlesa Fillmore'a (Fillmore 1968, 1977), Michaela Hollidaya (Halliday 1985) i Małgorzaty Korytkowskiej (Korytkowska 1984, 1992, 1993).

[269] Fillmore utrzymuje w swych pracach, iż gramatyka przypadków semantycznych jest gramatyką uniwersalną; analizując najgłębszy poziom struktury wyrażania, Fillmore oczekiwał pojawienia się cech wspólnych dla wszystkich języków. Z praktyki jednak wynika, iż o pełnej uniwersalności nie można mówić.

[270] Termin – „przypadek" – niezręczny dla języków fleksyjnych.

[271] Jeden z wyróżnionych przypadków to Agentive, który na powierzchni może, ale nie musi okazać się podmiotem.

[272] Przypadek semantyczny to inaczej funkcja, rola, która jest obecna na poziomie struktury głębokiej. Wykładnikiem tej roli w strukturze powierzchniowej jest wyrażenie argumentowe.

[273] Wyodrębnienie argumentu o wartości Experiencer wymagało ustalenia, jakie predykaty otwierają miejsce dla tego argumentu. Klasa predykatów otwierających miejsce dla argumentu o wartości Experiencer została wyodrębniona na potrzeby tego opracowania na podstawie następujących kryteriów: 1) otwierają miejsce dla argumentu o cesze [+Hum] (wtórnie również dla argumentu o cesze [+Anim]); 2) odnoszą się do stanów psychosomatycznych i zdarzeń, które nie podlegają obiektywnej obserwacji, ponieważ dotyczą sfery wewnętrznej jednostki; 3) określają stany i zdarzenia zachodzące bez udziału woli jednostki, która jest nimi objęta; oznaczają stany i zdarzenia (nie czynności!), które jeśli wymagają pojawienia się „sprawcy zdarzenia", to jest to „sprawca" nie działający świadomie i celowo (Kaczmarska, 2001: 177–179). Por. *Experiencer odczuwa uczucia, emocje, przeżywa różne stany psychiczne, mimowolnie odbiera różne bodźce zmysłami* (Hajnicz i Nitoń 2017). Zob. też: Experiencer (Korytkowska 1984, 1992, 1993; Kaczmarska 2001, 2010), Prožitatel (*Příruční mluvnice češtiny* 1996), Senser (Halliday 1985).

[274] Experiencer – podkreślenie linią ciągłą, Stimulus – linią falowaną.

Pacientku dlouho trápila prasklá ovariální cysta. (MVRpz2001cz)
Pacjentka długo cierpiała z powodu pękniętej torbieli na jajniku. (MVPdk2005pl)

Lucie mne miluje a její zdraví závisí na mé lásce! (MKZ1991cz)
Łucja mnie kocha i jej zdrowie zależy od mojej miłości! (MKZ1999pl)

Ale snad o to víc ho mrzelo, že petici nepodepsal. (MKNlb1985cz)
Ale tym bardziej chyba żałował, że nie podpisał petycji. (MKNlb1992pl)

Występowanie obiektów o tej samej wartości jest tylko pozornym ułatwieniem, nie rozstrzyga bowiem kwestii wyboru ekwiwalentów. Warto się więc przyjrzeć pozostałym obiektom obecnym w strukturze danych czasowników.

Jak zostało już wcześniej wspomniane, partycypantem pojawiającym się w każdym z analizowanych przypadków jest argument o wartości Experiencera. Drugim argumentem występującym przy czasownikach tego typu jest pewnego rodzaju źródło lub bodziec[275]. Należy oczekiwać, iż w strukturach z potencjalnymi ekwiwalentami również musi się znaleźć istota przeżywająca dany stan (tu – Experiencer) i (na ogół) bodziec ten stan wywołujący.

Ustalenie ról semantycznych argumentów łączących się z badanymi czasownikami i ich ekwiwalentami z uwagi na powtarzalność tych samych przypadków głębokich może nie pomóc w wyłonieniu adekwatnego ekwiwalentu, choć kluczowe mogą się tu okazać elementy dodatkowe występujące w zdaniu. Warto też rozważyć przeprowadzenie głębszej analizy realizacji powierzchniowej danych argumentów[276].

Mimo iż ten etap analizy nie przyniósł do tej pory przełomu w badaniu, nie zostanie usunięty z finalnej wersji algorytmu. Analizowane jednostki (wyrażające różne emocje i uczucia) są dość jednolite pod względem łączliwości z argumentami o pewnych wartościach – zawsze znajdzie się tu przeżywający czy źródło i wyłącznie na samej tej podstawie nie sposób dokonać zróżnicowania znaczenia. Krok ten może być jednak wykorzystany w badaniu innych grup czasowników, gdzie rola semantyczna oryginału i ekwiwalentu może być istotna przy różnicowaniu znaczenia jak choćby w przypadku zdań:

Przejechał las.
Przejechał psa.

W przypadku wyrazu „las" chodzi o miejsce (Location), natomiast „pies" to w terminologii słownika walencyjnego *Walenty* rola semantyczna Theme. Na tej

[275] W różnych opracowaniach nazywany Impuls, Stimulus, Źródło. FrameNet wymienia szereg ról występujących z Experiencerem: Stimulus, Circumstances, Degree, Empathy_target, Frequency, Manner, Reason, Time, State, Duration, Topic, Content (Baker, Fillmore i Lowe 1998).

[276] Próba takiej analizy zostanie przeprowadzona w następnym rozdziale.

podstawie można ustalić dwa znaczenia czasownika *przejechać* i spróbować ustalić ekwiwalenty.

Krok piąty – Pattern Grammar

Czwarty krok to wykorzystanie metod teorii Pattern Grammar (Francis, Hunston i Manning 1996; Hunston i Francis 2000; Ebeling i Ebeling 2013). Ten etap ma w założeniu sprawdzić, czy istnieje zależność pomiędzy znaczeniem wyrazu a otoczeniem, w jakim występuje. Hunston i Francis twierdzą, że jeżeli jakieś słowo jest wieloznaczne, a jednocześnie pojawia się w kilku wzorcach, to każdy wzorzec pojawi się z jednym z jego znaczeń częściej niż z pozostałymi, czyli dany wzorzec wskaże najbardziej prawdopodobne znaczenie słowa.

> *If a word has several senses, and is used in several patterns, each pattern will occur more frequently with one of the senses than the others, such that the patterning of an individual example will indicate the most likely sense of the word in that example* (Hunston i Francis 2000: 20).

Według tej teorii każdy wyraz ma przypisany zestaw wzorców, opisujących typowe konteksty, w których wyraz ten się pojawia. Często są to konteksty odrębne dla różnych znaczeń[277]. Chociaż wydaje się, iż najprościej byłoby tą metodą identyfikować związki frazeologiczne, co zresztą badacze również czynią, jednak podejście to może być stosowane do różnych wyrazów, np. kontrastywna analiza angielsko-norweska czasownika angielskiego *cause* (Ebeling S.O. 2013)[278]. Podobną analizę można przeprowadzić dla jednostek czeskich np. *způsobit*, *vykašlat se*, *po krk*, czy polskiej jednostki *spowodować* czy *bardzo*[279].

Czasowniki wyrażające stany psychiczne są w większości polisemiczne. Poprzez śledzenie i identyfikację ich wzorców oraz ustalenie prozodii semantycznej można połączyć konkretne znaczenie z typem wzorca (rozumianym jako powtarzalna kombinacja słów).

---

[277] Kontekst ów określa się najczęściej poprzez ustalenie otoczenia prawo- i lewostronnego; np. sytuując daną jednostkę centralnie, bada się kontekst obejmujący trzy pozycje na lewo i na prawo od tej jednostki.

[278] Por. angielsko-norweskie opracowanie jednostek: ang. *big deal*, *found*, *out of the ordinary, get hold of*, *back and forth, in and out*, norweskiej: *få tak,* (Ebeling i Ebeling 2013: 97–207), ang. *way* (Johansson 2009), ang. *well* (Johansson 2006) czy ang. *cause* (Ebeling S.O. 2013), por. (Stubbs 1995).

[279] Zwłaszcza w przypadku czasowników takich jak *způsobit* czy *spowodować* warto ustalić prozodię semantyczną, czyli relację pomiędzy daną jednostką a wyrazami, które najczęściej z nią się pojawiają.

A pattern can be identified if a combination of words occurs relatively frequently, if it is dependent on a particular word choice, and if there is a clear meaning associated with it (Hunston i Francis 2000: 37).

Podjęto próbę ustalenia, czy w rzeczywistości istnieje taka powtarzalność w poświadczeniach korpusowych dla analizowanych tu czasowników. Większość badań prowadzonych w tym nurcie dotyczy języka angielskiego; z uwagi na to, że jest to język pozycyjny i w znacznym stopniu izolujący, można założyć, iż ma mniej swobodny szyk zdania niż język polski czy czeski, zatem wzorce w nim występujące są łatwiej identyfikowane[280]. W przypadku języka czeskiego czy polskiego badanie oparte na kontekście linearnym może być więc bardziej skomplikowane, ewentualnie może wymagać dużo większej liczby danych[281].

Ręczna analiza oparta na materiale z korpusu InterCorp ukazała między innymi dwa wzorce czeskiej jednostki *být líto* związanych z dwoma znaczeniami ('być przykro', 'żałować'). Jeżeli jednostka *být líto* łączy się z dwoma frazami nominalnymi (w celowniku i dopełniaczu), wówczas taki jej wzorzec odpowiada polskiemu ekwiwalentowi *żal*. Jeżeli natomiast *být líto* łączy się tylko z celownikową frazą nominalną (oraz ewentualnie z elementem *to*), wzorzec ten odpowiada znaczeniu polskiej jednostki *(być) przykro* (Kaczmarska 2017A: 187–188). Ilustruje to poniższa tabelka.

Tabela 20. Wzorce czeskiej jednostki *být líto* (*żal, być przykro*)

*Jak **mi ho** bylo líto!*	*Pak **mi** je líto.*
Jakże mi go było żal!	*Wobec tego, przykro mi!*
*Je **mi ho** samozřejmě líto.*	*Potom **nám** to bylo oběma líto.*
Jest mi go oczywiście żal…	*Potem nam obu było przykro.*
*Přišlo **mi jí** prostě líto.*	*…nabídne mi sisinku a já si vezmu, protože by **mu** bylo líto, kdybych si nevzala…*
Po prostu zrobiło mi się jej żal.	*...zaprasza mnie na cuksa i ja biorę, bo byłoby mu przykro, gdybym nie wzięła…*
být líto** + NP_{DAT} + NP_{GEN} = **żal	***být líto** + NP_{DAT} + **to** / **Ø** = (**być**) **przykro***

[280] Jeśli w danym języku o funkcji gramatycznej i składniowej wyrazu decyduje głównie jego pozycja w zdaniu, a każda część zdania ma swoje ustalone miejsce, które pozwala na zachowanie sensu wypowiedzi, można przypuszczać, iż w tym języku będzie można identyfikować powtarzające się różne schematy.

[281] Autorzy niektórych analiz leksemów angielskich ustalali powtarzające się kombinacje słów na podstawie stosunkowo niewielkich danych, np. ang. *big deal* – 32 poświadczenia (Ebeling i Ebeling 2013: 97), około 200 poświadczeń czasownika *cause* w badaniu angielsko-norweskim (Ebeling S.O. 2013)

W przypadku czasownika *kochać* ręczna analiza pozwoliła na zidentyfikowanie jednego wzorca mogącego odpowiadać jakiejś konkretnej jednostce czeskiej – *jak Boga kocham*[282]; odpowiedniki czeskie znalezione w poświadczeniach to: *namouduši* (6), *na mou duši* (3), *proboha* (2), *to si piš* (1), *propána* (1), *fakt mi věřte* (1), *věř mi* (1), pominięcie w tłumaczeniu (1). Kontekst linearny oparty na dużo większej liczbie danych mógłby być oceniony dzięki analizie stochastycznej.

Krok szósty – inne metody

Ostatni krok to miejsce na inne rozwiązania, często przyszłościowe takie jak np. wykorzystanie narzędzia Sketch Engine dla obu języków na tym samym korpusie[283], zaawansowane badania stochastyczne, czy połączenie już dostępnych narzędzi. Być może dzięki nim, na podstawie badania kontekstu i rozróżnienia łączliwości z różnymi obiektami, powiedzie się ustalenie ekwiwalentów do wszystkich znaczeń czasowników, które nie znajdą swoich odpowiedników na wcześniejszych poziomach algorytmu (Kaczmarska, Rosen, Hana i Hladká 2015).

Word Sketch

Sketch Engine[284] to narzędzie umożliwiające przedstawienie obrazu kolokacyjnego danego słowa – tzw. Word Sketch – również z odniesieniem do gramatyki. Narzędzie to zostało wykorzystane w trakcie analizy czasownika *soucítit* i może się wydawać uniwersalnym narzędziem do analizy kolokacji i połączeń wyrazowych rozumianych zarówno w kategoriach walencji, jak również pośrednio – przypadków semantycznych. Kolokaty analizowanej jednostki są pogrupowane bowiem według relacji gramatycznych, w których się pojawiają. Istotna w przypadku tego narzędzia jest konieczność bazowania na tym samym korpusie dla wszystkich (tu – dwóch) badanych języków. W przypadku korpusu równoległego InterCorp narzędzie Sketch Engine jest dostępne dla języka polskiego.

[282] Z innowacjami – *jak ciocię kocham*, *jak babcię kocham*. Słabo wyodrębnił się też wzorzec *kochać nad życie* (3 poświadczenia) i *kochać nade wszystko* (3 poświadczenia).

[283] O możliwości wykorzystania narzędzia Sketch Engine pisał już też Tadeusz Piotrowski (Piotrowski 2011: 53–54).

[284] *A word sketch is a one-page, automatic, corpus-derived summary of a word's grammatical and collocational behavior* – https://www.sketchengine.eu/user-guide/user-manual/word-sketch/ (dostęp 14.11.2018).

W przypadku języka czeskiego narzędzia tego nie można zastosować na tekstach z InterCorpu; oprogramowanie pobiera teksty z całego korpusu języka czeskiego, co sprawia, że otrzymane wyniki czesko-polskie są nieporównywalne. Paralelne dwujęzykowe Word Sketches dla materiału z InterCorpu to kwestia przyszłości.

Wykorzystanie *Słowosieci* i narzędzia *WoSeDon*

WoSeDon to narzędzie przeznaczone do ujednoznaczniania znaczeń słów (ang. *Word Sense Disambiguation*). Zostało zaprojektowane dla języka polskiego, jednak po odpowiedniej zmianie tagsetu, możliwe byłoby działanie również na innych językach. Głównym algorytmem wykorzystanym w tym narzędziu, który został zaimplementowany do rozstrzygania niejednoznaczności jest PageRank. *WoSeDon* zawiera różne modyfikacje tego algorytmu oraz umożliwia manipulacje wieloma parametrami, które mają wpływ na jakość ujednoznaczniania znaczeń leksykalnych (Kędzia, Piasecki i Orlińska 2016; Janz, Kędzia i Kaszewski 2018).

Słowosieć (z ang. *wordnet*) to, jak piszą jej twórcy, relacyjny słownik semantyczny, który odzwierciedla system leksykalny języka polskiego[285] (Maziarz, Piasecki i Rudnicka 2014; Piasecki, Kędzia i Orlińska 2016; Dziob i Piasecki 2018). Pojedyncze znaczenia w *Słowosieci* połączone są wzajemnymi relacjami. Tak powstaje sieć, w której każdy wyraz jest zdefiniowany poprzez odniesienie do innych wyrazów, np. *koń* to zwierzę; *koń* stanowi całość, na którą składają się np. *głowa*, *kopyto*, *wrąb*, *kłąb*, *noga*, *pęcina*, natomiast jego wyrazami bliskoznacznymi są *ogier*, *klacz*, *szkapa*. *Słowosieć* to także bardzo ważny zasób, biorący udział w komputerowym przetwarzaniu języka i badaniach nad sztuczną inteligencją, znajdującym m.in. zastosowanie w automatycznych tłumaczeniach Google Translate.

Według autorki tego opracowania istnieje możliwość wykorzystania narzędzia WoSeDon i *Słowosieci* w procesie ujednoznaczniania znaczeń podczas badań dwujęzycznych. Dzięki narzędziu *WoSeDon* czasownikom zostają przypisane synsety[286], innymi słowy narzędzie analizuje, w jakich znaczeniach zostały użyte wskazane czasowniki – np. czasownik *lubić*[287] i generuje informację zwrotną (obok synsetów – przykłady użycia).

[285] Polska *Słowosieć* jest już największym wordnetem na świecie i nieustannie się rozrasta. Por. *WordNet* (Miller 1995; Fellbaum 1998).

[286] Ujednoznacznienie znaczenia w WoSeDon działa również dla wszystkich części mowy. Po wprowadzeniu tekstu narzędzie może przypisać synsety wszystkim częściom mowy, bądź tylko niektórym; zależy to od ustawienia filtrów.

[287] *Słowosieć* przedstawia 5 znaczeń czasownika *lubić* – http://plwordnet.pwr.wroc.pl/wordnet/516c3334-dcfa-11e8-9074-ebaa873339db (dostęp 15.11.2018).

Na podstawie ujednoznacznionego znaczenia czasownika można dopasować ekwiwalent (np. czeski); może nim być jakieś jedno konkretne znaczenie czasownika, a nie „cała" wieloznaczna jednostka. Wiąże się to ze stwierdzeniem Tadeusza Piotrowskiego, iż tłumacz poszukuje ekwiwalentu dla określonego kontekstu, natomiast leksykograf poszukuje „uogólnionego" ekwiwalentu. Jak pisze Tadeusz Piotrowski, tłumacz ma całkowitą swobodę w kształtowaniu tekstu tłumaczenia, leksykograf zaś do swego zadania potrzebuje zbioru wzorców z informacją dla najczęstszych kontekstów, przy czym informacja ta powinna obejmować także ich kontekst sytuacyjny (Piotrowski 2011: 55). W proponowanym badaniu zostałyby zebrane przez *WoSeDon* konteksty użycia dla poszczególnych znaczeń danego czasownika. Wówczas, gdyby paralelne teksty były opracowane podobnie[288], możliwe by było wykorzystanie danych z InterCorpu, znalezienie odpowiadających sobie znaczeń i w konsekwencji dopasowanie ekwiwalentów.

lubić	lubić.1(35:cczuj)	lubić.1(35:cczuj)	Chciałbym kontynuować tę rolę w organizacji, którą ogromnie lubię.
lubić	lubić.3(35:cczuj)	lubić.3(35:cczuj) gustować.1(35:cczuj)	Tak, bardzo lubię Nicka Dentona – jest bardzo szczery, a przy tym skromny jeśli chodzi o zakres swoich działań.
lubić	lubić.1(35:cczuj)	lubić.1(35:cczuj)	Internautom, którzy lubią się bać, podpowiada, czym mogą uraczyć zmysły tuż przed 31 października – prezentuje na stronie listę „strasznych" filmów i „makabrycznych" historii.
lubić	lubić.3(35:cczuj)	lubić.3(35:cczuj) gustować.1(35:cczuj)	Gdy, jednak dysponuję wolnym czasem, lubię grać w golfa, jeździć na snowboardzie, a także nurkować.
lubić	lubić.1(35:cczuj)	lubić.1(35:cczuj)	To jest dziedzina sportu, którą bardzo lubię.
lubić	lubić.3(35:cczuj)	lubić.3(35:cczuj) gustować.1(35:cczuj)	Dzięki księdzu Wojtkowi nasze dzieci i wnuki lubią przychodzić w niedzielę na mszę o godzinie 11.00, bo ksiądz Drozdowicz dopuścił je do ołtarza.
lubić	lubić.1(35:cczuj)	lubić.1(35:cczuj)	Jednocześnie dodał, że wobec całego tego zamieszania wcale nie dziwi się, że politycy nie są w Polsce lubiani.

[288] Język czeski również posiada swój WordNet (Pala i inni 2011).

Eksperymentalnie przeanalizowano polski czasownik *cierpieć*[289]; poniżej kilka przykładów.

Nicola powiedziała też, że niedawno straciła pierwszego syna, Logana, który również cierpiał na porażenie mózgowe. Jednostka: **cierpieć.1**; synset: niedomagać.1, cierpieć.1, chorować.1.

Cierpiał na niewydolność układu oddechowego i krwionośnego. Jednostka: **cierpieć.1**; synset: niedomagać.1, cierpieć.1, chorować.1.

Alicja Tysiąc cierpiała od wielu lat na poważną wadę wzroku. Jednostka: **cierpieć.1**; synset: niedomagać.1, cierpieć.1, chorować.1

Wielu z nas cierpiało, ponieważ sąd zaakceptował chorą, opartą na błędnych przesłankach teorię Parental Alienation Syndrome (PAS). Jednostka: **cierpieć.2**; synset: tolerować.1, cierpieć.2.

Wszystkie nasze mamy traktowano jak winowajczynie, a my wszystkie dzieci cierpiałyśmy z powodu sposobu, w jaki traktowano nasze mamy. Jednostka: **cierpieć.2**; synset: tolerować.1, cierpieć.2.

Claiborne od wielu lat cierpiała na chorobę nowotworową. Jednostka: **cierpieć.4**; synset: męczyć_się.3, zadręczać_się.1, cierpieć.4[290].

Michael Schiavo, mąż Terri, wnosił do sądu Florydy o odłączenie jego żony od tuby, aby nie cierpiała i zmarła. Jednostka: **cierpieć.4**; synset: męczyć_się.3, zadręczać_się.1, cierpieć.4.

Już sam synset jest pierwszą wskazówką i konkretyzacją znaczenia. Umożliwia nam znalezienie czeskiego ekwiwalentu konkretnego znaczenia czasownika np. *cierpieć*$_1$. Trafnym odpowiednikiem byłby czasownik *trpět* w swoim drugim znaczeniu: *trpět*$_2$ (čím: nač) 'být nemocen chorobou': *pacient trpící těžkým revmatismem* (Havránek i inni 1989)[291].

Ta część algorytmu jest otwarta na inne metody, które dałyby się zaaplikować do badania języków czeskiego i polskiego. Można sobie wyobrazić

[289] Dzięki uprzejmości Kolegów z Politechniki Wrocławskiej – dr. inż. Macieja Piaseckiego i mgr. inż. Pawła Kędzi.

[290] W tym przypadku narzędzie błędnie oznaczyło synset.

[291] Przy czym warto zaznaczyć, iż zarówno w języku polskim, jak i czeskim pojawia się w tym kontekście najczęściej poważna choroba. Analiza czeskiego materiału została wykonana jednak ręcznie.

dodanie jakiegokolwiek okienka z kolejną metodą, ale nie wszystkie mogą być aplikowane[292].

Opisany algorytm obejmuje pięć etapów analizy odnoszących się zarówno do poszukiwań leksykologicznych, jak i analiz ręcznych i automatycznych, a także konkretnych teorii językoznawczych. Wybrane zostały trzy podejścia językoznawcze – walencja, gramatyka przypadków głębokich i Pattern Grammar. W zasadzie wszystkie bazują na możliwościach łączenia się wyrazów; w przypadku walencji i ról semantycznych istotne są zależności, dla Pattern Grammar relewantny jest też kontekst liniowy. Krok szósty to badania eksperymentalne i metody przyszłościowe. W następnym rozdziale przeanalizowany zostanie czasownik *toužit*, który przejdzie przez wszystkie etapy proponowanego algorytmu.

[292] Kolejnym krokiem mogłoby być np. zastosowanie metod gramatyki kognitywnej; wykorzystanie korpusów w gramatyce kognitywnej jest już znane (Gries 2007; Lewandowska-Tomaszczyk 2012). Na tym poziomie można by odkodować znaczenie słowa w kontekście konceptualizacji zjawiska nim nazwanego. W badaniach pilotażowych próbowano analizować jednostki *mít rád* i *milovat* w oparciu o elementy metod kognitywnych, jednakże bardzo trudno implementować je do algorytmu, ponieważ przeprowadzane są w większości całkowicie manualnie, a badany kontekst nierzadko musiałby obejmować cały tekst (niemożliwe w przypadku automatycznej ekscerpcji z korpusu InterCorp). W analizie manualnej w przypadkach problemowych możemy odnieść się także do eksplikacji i naturalnego metajęzyka semantycznego (Wierzbicka 1980, 2001) lub skonstruować skalę intensywności właściwości wyrażanej przez dany czasownik (Bratman 1987; Mikołajczuk 1997, 1999). Każda z tych metod jest również przeprowadzana ręcznie, dla każdej jednostki oddzielnie.

Przypadek 6 – *toužit*

Jednostką analizowaną według opracowanego algorytmu jest czasownik *toužit*, który doczekał się dwóch artykułów (Kaczmarska i Rosen 2013; Kaczmarska, Rosen, Hana i Hladká 2015), a także kilku opracowań, w których jest jedną z kilku omawianych jednostek, m.in. (Kaczmarska 2013, 2017; Kaczmarska i Rosen 2014A, 2015B, 2016)[293]. Czasownik *toužit* jest jedną z tych wieloznacznych jednostek, dla których niezwykle trudno wskazać ekwiwalent.

Ekwiwalenty słownikowe

Definicja słownikowa

W cytowanym już słowniku języka czeskiego przedstawione są trzy znaczenia czasownika *toužit*: 1) 'mít touhu', 'oddávat se touze'; 2) 'stěžovat si', 'naříkat si', 'stýskat si'; 3) 'naříkat', 'bědovat', 'hořekovat', przy czym drugie i trzecie znaczenie autorzy słownika uznają za przestarzałe (Havránek i inni 1989). Nowszy (mniej obszerny) słownik języka czeskiego – *Slovník spisovné češtiny pro školu a veřejnost* (SSCSV) – podaje tylko jedno znaczenie: *mít touhu* z synonimem *přát si* (Filipec, Daneš, Machač i Mejstřík 1994).

Słownik dwujęzyczny

Czasownik *toužit* w słowniku czesko-polskim ma swoje trzy odpowiedniki: *tęsknić, pragnąć, marzyć* (Siatkowski i Basaj 2002: 809). Wszystkie trzy jednostki dotyczą sfery uczuć czy emocji, ale nie są dla siebie najbliższymi synonimami.

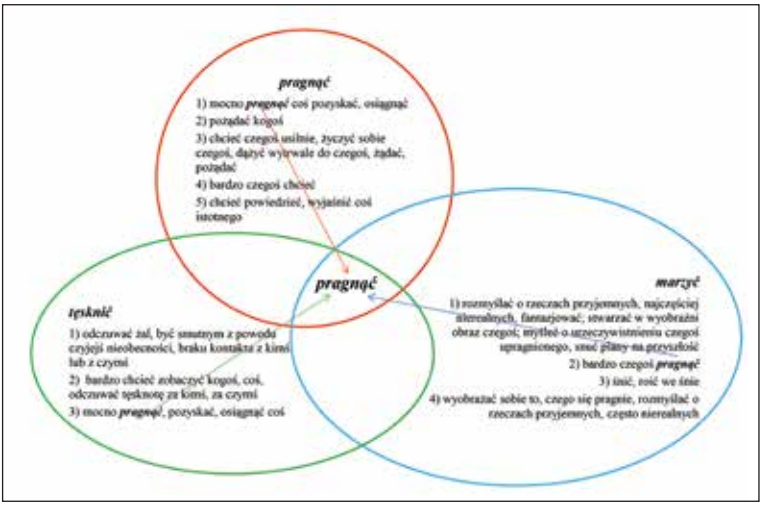
[293] Do tych badań autorka w tym rozdziale nawiązuje.

Jak podaje *Słownik języka polskiego* (Szymczak 1995: 466) czasownik *tęsknić* ma dwa znaczenia: 1) 'bardzo chcieć zobaczyć kogoś, coś, odczuwać tęsknotę za kimś, za czymś'; 2) 'mocno pragnąć coś pozyskać, osiągnąć'[294].

Słownik języka polskiego (Szymczak 1995: 110) odnotowuje jedno znaczenie współczesne czasownika *marzyć* – 1) rozmyślać o rzeczach przyjemnych, najczęściej nierealnych, fantazjować; stwarzać w wyobraźni obraz czegoś; myśleć o urzeczywistnieniu czegoś upragnionego, snuć plany na przyszłość, bardzo czegoś pragnąć' oraz jedno znaczenie przestarzałe – 2) 'śnić, roić we śnie'[295].

Czasownik *pragnąć* jest definiowany jako 'chcieć czegoś usilnie, życzyć sobie czegoś, dążyć wytrwale do czegoś, żądać, pożądać' (Szymczak 1995: 867)[296].

Analiza definicji tych trzech polskich czasowników pozwala ustalić, iż mają one część wspólną – element wolitywny „pragnąć"[297].



Obraz 27. Czasowniki *pragnąć, marzyć, tęsknić* – znaczenia słownikowe

[294] Por. słownik języka polskiego online: 1) 'odczuwać żal', 'być smutnym z powodu czyjejś nieobecności, braku kontaktu z kimś lub z czymś'; 2) 'mocno pragnąć pozyskać, osiągnąć coś'. Dostęp: https://sjp.pwn.pl/szukaj/t%C4%99skni%C4%87.html 15.02.2019

[295] Por. słownik języka polskiego online: 1) 'wyobrażać sobie to, czego się pragnie, rozmyślać o rzeczach przyjemnych, często nierealnych'; 2) 'bardzo czegoś pragnąć'; 3) dawniej – 'śnić, roić we śnie'. Dostęp: https://sjp.pwn.pl/szukaj/marzy%C4%87.html 15.02.2019

[296] Por. słownik języka polskiego online: 1) 'bardzo czegoś chcieć'; 2) 'pożądać kogoś'; 3) 'chcieć powiedzieć, wyjaśnić coś istotnego'. Dostęp: https://sjp.pwn.pl/szukaj/pragn%C4%85%C4%87.html 15.02.2019. Trzecie znaczenie pojawiające się w tym słowniku nie pojawia się w cytowanym już *Słowniku języka polskiego*. Nie notuje go też *Słownik języka polskiego* pod red. Doroszewskiego: https://sjp.pwn.pl/doroszewski/pragnac;5481062.html

[297] Warto zwrócić uwagę na fakt, iż czasownik *pragnąć*, ze względu na swój wolitywny odcień, nie jest typowym czasownikiem odnoszącym się emocji, uczuć czy przeżyć.

Jednocześnie pojawia się szereg pytań: czy owa część wspólna jest jądrem znaczenia polskiego ekwiwalentu czeskiego czasownika, czy jednostka *toužit* reprezentuje pojęcie nieznane językowi polskiemu i do tego wspólnego znaczenia zbliża się cały klaster czasowników polskich, czy może chodzi tu o wieloznaczność czeskiego czasownika uchwytną dopiero poprzez język drugi (tu – polski)? W rozwiązaniu tego problemu pomocna może się okazać analiza korpusowa.

Automatyczna ekstrakcja ekwiwalentów z czesko-polskiego korpusu równoległego

Korpus równoległy InterCorp stanowi bazę do automatycznej ekstrakcji par ekwiwalentów; lista takich par to słownik wygenerowany metodą automatyczną z czesko-polskiej części korpusu. W pierwszych badaniach wykorzystywana była wspomniana wcześniej GIZA++[298], a materiał językowy pochodził z wersji 6 InterCorpu[299]. Aktualnie listę par ekwiwalentów można wygenerować dzięki usłudze Treq oferowanej przez Czeski Korpus Narodowy.

Automatyczna ekstrakcja par ekwiwalentów dostarcza cały szereg potencjalnych odpowiedników, nie bierze jednak pod uwagę obiektów łączących się z nimi, ani ich wymagań walencyjnych; jest oparta na kontekście[300]. Badanie schematów walencyjnych stanowi kolejną część analizy.

[298] GIZA++ to program stworzony do statystycznego przekładu maszynowego, który zawiera moduł wiązania segmentów na poziomie wyrazu (word-to-word alignment). Por. http://www.statmt.org/moses/giza/GIZA++.html (Och i Ney 2003).

[299] Na wstępie czeskie i polskie teksty były w lematyzowanej formie. Uwzględniona została jedynie beletrystyka (bez rozróżnienia na czeskie, polskie czy obce oryginały) – w tym teksty w języku czeskim: 11,885 mln słów i teksty w języku polskim 11,860 mln słów. Wiązanie segmentów (zdań) zostało ograniczone do 1:1, co zwiększa wiarygodność wiązania na poziomie zdania i wyrazu; to oznacza, że były wykorzystane tylko segmenty wiązane 1:1. Metodą tą wyekstraktowano 8,651 mln par lematów. Po złączeniu identycznych par (z zachowaniem informacji o ich liczbie) powstało 528 tysięcy haseł dwujęzycznych (Kaczmarska i Rosen 2013).

[300] W dalszej fazie można by zamiast par leksemów ekstraktować pary istotnych kolokacji, natomiast w odleglejszej perspektywie wykorzystać syntaktycznie anotowane teksty (równoległy bank drzew – parallel treebank).

Tabela 21. Lista ekwiwalentów wygenerowana przez Treq[301]

Lista ekwiwalentów czasownika *toužit* wygenerowana przez Treq		
frekwencja	**%**	**potencjalny ekwiwalent**
716	32,1%	*pragnąć*
476	21,3%	*chcieć*
221	9,9%	*tęsknić*
164	7,3%	*marzyć*
108	4,8%	*pożądać*
62	2,8%	*ochota*
38	1,7%	*pragnienie*
24	1,1%	*zależeć*
21	0,9%	*życzyć*
20	0,9%	*szukać*
19	0,9%	*zapragnąć*
18	0,8%	*potrzebować*
14	0,6%	*tęsknota*
13	0,6%	*spragniony*
13	0,6%	*czekać*
12	0,5%	*woleć*
10	0,4%	*domagać*
10	0,4%	*chęć*
9	0,4%	*upragniony*
9	0,4%	*dążyć*
8	0,4%	*zatęsknić*
8	0,4%	*aspirować*
6	0,3%	*potrzebny*
6	0,3%	*żądny*
5	0,2%	*doczekać*
5	0,2%	*oczekiwać*
5	0,2%	*potrzeba*
5	0,2%	*szaleć*

[301] W tabeli ujęte zostały tylko te ekwiwalenty, które pojawiy się minimalnie pięć razy.

Analiza walencyjna

Wstępem do analizy walencyjnej była automatyczna ekscerpcja poświadczeń (wraz z całymi segmentami) czasownika *toužit* z korpusu równoległego InterCorp (wersja 10) oraz badanie semantyczno-leksykalne tych poświadczeń. Leksem *toužit* wyszukiwany był w części czesko-polskiej trzonu korpusu (*core*) w dwóch odrębnych ekscerpcjach: w czeskich oryginałach wyrównanych z tekstami polskimi oraz w czeskich tłumaczeniach polskich oryginałów. Chcąc przedstawić rozmiar materiału poddanego ekscerpcji, należy dokonać rozróżnienia na oryginały czeskie, polskie i inne (obce), występujące w języku czeskim i polskim. Dane liczbowe dla wersji 10 korpusu InterCorp przedstawiają się następująco (wszystkie dane liczbowe oprócz frekwencji czasownika podane w milionach słów):

Tabela 22. Liczby słów (w milionach) w beletrystycznych tekstach w czesko-polskiej części korpusu InterCorp (wersja 10)[302]

InterCorp wersja 10	czeskie oryginały	polskie oryginały	obce oryginały	Σ
w języku czeskim	2,325	2,726	18,289	23,34
w języku polskim	2,44	2,658	18,14	23,238
f(*toužit*) – liczba poświadczeń	272	250	1,954	2,476

Wyszukiwanie: czeskie oryginały → polskie przekłady (cs → pl)

W wersji dziesiątej InterCorpu (v10) odnaleziono 272 poświadczenia czasownika *toužit*[303] (Tabela 23).

[302] Badania pojawiające się w poprzednich publikacjach opierały się na materiale z wersji 5, 6, 7 InterCorpu. Por. liczby słów (w milionach) w beletrystycznych tekstach w czesko-polskiej części korpusu InterCorp (wersja 5):

InterCorp wersja 5	czeskie oryginały	polskie oryginały	obce oryginały	Σ
w języku czeskim	1,709	1,757	4,917	8,384
w języku polskim	1,783	1,698	4,903	8,384
f(*toužit*) – liczba poświadczeń	248	146	418	812

[303] Poświadczenia te zostały przebadane ręcznie; każdy z segmentów był manualnie otagowany i został w nim wskazany element będący odpowiednikiem czasownika *toužit* w danym segmencie.

Tabela 23. Ekwiwalenty polskie czasownika *toužit* – rozkład procentowy

Ekwiwalenty w języku polskim	**Liczba wystąpień**	**Udział procentowy**
pragnąć	133	48,88%
marzyć	38	13,95%
tęsknić	34	12,49%
chcieć	22	8,08%
pożądać	11	4,04%
mieć ochotę	4	1,47%
zapragnąć	4	1,47%
dążyć	2	0,74%
pragnienie	2	0,74%
zatęsknić	2	0,74%
ciągnąć do	1	0,37%
czekać	1	0,37%
dbać	1	0,37%
imponować	1	0,37%
marzący	1	0,37%
myśleć	1	0,37%
pożądanie	1	0,37%
pożądany	1	0,37%
szukać	1	0,37%
tęskno	1	0,37%
tęsknota	1	0,37%
upragniony	1	0,37%
zachciewać się	1	0,37%
żądać	1	0,37%
żądny	1	0,37%
życzyć sobie	1	0,37%
żywić miłość	1	0,37%
zmiana struktury	2	0,37%
błąd korpusu	2	0,74%
RAZEM	**272**	100,00%

Analiza ekwiwalentów znalezionych w InterCorpie ukazała, iż skupiają się one wokół odpowiedników zaproponowanych przez słownik dwujęzyczny, tworząc grupy semantyczne[304] (Tabela 24).

[304] W rozdziale 6; odpowiedniki czasownika *zdát se* zostały skupione w grupach wyrazów bliskoznacznych. Por. na temat wspólnych wniosków wobec *chcieć* i *pragnąć* (Kaczmarska 2014B: 51–53) oraz klasyfikowania *chcieć* jako synonimu *pragnąć* (Kaczmarska 2015A: 136).

Tabela 24. Ekwiwalenty polskie czasownika *toužit* w podziale na grupy semantyczne

Ekwiwalenty w języku polskim		**Liczba wystąpień**	**Udział procentowy**
pragnąć	*pragnąć*	185	68,01%
	chcieć		
	pożądać		
	mieć ochotę		
	zapragnąć		
	dążyć		
	pragnienie		
	pożądanie		
	pożądany		
	upragniony		
	zachciewać się		
	żądać		
	żądny		
	życzyć sobie		
marzyć	*marzyć*	39	14,32%
	marzący		
tęsknić	*tęsknić*	38	13,97%
	zatęsknić		
	tęskno		
	tęsknota		
inne	*ciągnąć do*	7	2,59%
	czekać		
	dbać		
	imponować		
	myśleć		
	szukać		
	żywić miłość		
	zmiana struktury	1	0,37%
	błąd korpusu	2	0,74%
	RAZEM	**272**	100,00%

pragnąć

Rozkład procentowy ukazuje, iż najczęściej wybieranym ekwiwalentem jest czasownik *pragnąć* oraz jego jednostki bliskoznaczne (wśród nich: *chcieć, pożądać, mieć ochotę, zapragnąć, dążyć, pragnienie, pożądanie, upragniony, zachciewać się, żądać, żądni, życzyć sobie*); czasowniki *chcieć* czy *pożądać* sklasyfikować można jako synonimy *pragnąć* (Kaczmarska 2015A: 136). Należy jednak zwrócić uwagę, iż czasowniki w tej grupie mają różne wymagania składniowe, np. *chcieć* łączy się z obiektem w bierniku, *pragnąć* i *pożądać* – w dopełniaczu. W sumie jednostki te stanowią 68,01% wszystkich znalezionych poświadczeń, np.:

Toužila, aby přišel, toužila, aby ji k sobě pozval! MKNLB1985cz
***Pragnęła**, żeby przyszedł, pragnęła, aby zaprosił ją do siebie!* MKNLB1992pl

Toužila, aby s ní sdílelo její samotu alespoň nějaké zvířátko. MKN1993cz
***Chciała** dzielić z kimś swą samotność, choćby z jakimś zwierzątkiem.* MKN2008pl

marzyć

Marzyć i jednostki temu czasownikowi bliskie (*marzący*) – 14,32%:

Netoužila o něm dlouze rozprávět. MKVNR1997cz
*Nie **marzyła** o długiej rozmowie na ten temat.* MKWP2001pl

tęsknić

Jednostki zgromadzone wokół czasownika *tęsknić* (*tęsknić, zatęsknić, tęskno, tęsknota*) stanowią wraz z nim 13,97% poświadczeń:

Ale právě po ní vždycky toužil. MKNLB1985cz
*Ale do niej właśnie **tęsknił** od zawsze. MKNLB1992pl*

Grupy *marzyć* i *tęsknić* reprezentowane są przez podobną ilość poświadczeń; różni je kilkadziesiąt setnych procenta.

Inne

W części „inne”[305] figuruje sześć poświadczeń, reprezentujących jednostki niezaklasyfikowane do wymienionych wcześniej grup (*ciągnąć do, czekać, dbać, imponować, myśleć, szukać, żywić miłość*):

[305] W monografii poświęconej konfrontacji języka czeskiego z językami romańskimi autorzy, prezentując badania oparte na danych z korpusu InterCorp, omawiają problem czeskich odpowiedników dla wybranych zjawisk gramatycznych pojawiających się w językach: portugalskim, hiszpańskim, włoskim i francuskim. Również w ich badaniu pojawia się kategoria „przełożono inaczej” – czes. *přeloženo jinak* (Čermák i Nádvorníková 2015).

Anebo vám připomenu tu nádhernou scénu z Fromentinova Dominika: oba zamilovaní, kteří po sobě léta toužili a nikdy se jeden druhého nedotkli… MKN1993cz
*Przypomnijcie sobie także wspaniałą scenę z Dominika Fromentina: kochankowie, którzy przez lata **żywili** do siebie **miłość**, lecz się nie dotykali…* MKN2008pl

Vždycky jsem byl nadšen počátkem, ale za krátký čas jsem znechuceně toužil po něčem jiném. VRR1944cz
*Zawsze byłem zachwycony początkiem, ale, zniechęcony, zaczynałem wkrótce **myśleć** o czym innym.* VRK1954pl

(…) *její idealisovaná postava je utvořena do značné míry podle jiné idealisované postavy, jejího otce, jejž od dětství zbožňuje. S klady, po nichž touží, i s výhradami, jež si stanovila.* VRR1944cz
(…) *jej wyidealizowany kochanek jest stworzony do pewnego stopnia na wzór innego wyidealizowanego mężczyzny – jej ojca, którego ubóstwiała od dzieciństwa, mężczyzny z cechami, które jej **imponowały**, i z wadami, które z góry określiła.* VRK1954pl

Zmiana struktury

W tabeli widoczna jest również pozycja dotycząca zmiany struktury przekładu w stosunku do tekstu oryginału; odnaleziono jeden taki przykład:

Marxovi se na procházku nechtělo, toužil být s Jenny a ne s Bettinou. MKN1993cz
Marks nie miał na przechadzkę najmniejszej ochoty, nad towarzystwo Bettiny stawiał towarzystwo Jenny. MKN2008pl

Błąd korpusu

W jednym z dwóch przykładów (~0,74%) oznaczonych w tabeli jako „błąd korpusu" został odnaleziony homonim – nazwa własna; w drugim – fragment polski, który miał być ekwiwalentem tekstu czeskiego, okazał się innym tekstem[306].

Nedaleko Ničína je vesnice Toužím a u ní kopec Palice, kde padly poslední výstřely druhé světové války. IDHB1998cz

Niedaleko od Niczina jest wioska Toużim, a koło niej góra Palica, gdzie padły ostatnie strzały drugiej wojny światowej. IDB2003pl

W cytowanym przykładzie nazwa miejscowości jest homonimiczna wobec formy pierwszej osoby liczby pojedynczej czasu teraźniejszego czasownika *toužit*[307].

[306] Podobna sytuacja miała miejsce podczas analizy czasownika *zdát se*; tam również została wyodrębniona grupa zawierająca poświadczona obarczona błędem korpusu. Szerzej na temat błędów w korpusach oraz ich przyczynach – (Hebal-Jezierska, Kaczmarska i Rosen 2016; Rosen 2016).

[307] Korpus powinien wychwycić nazwę własną również ze względu na nietypowy zapis – wielką literą.

Jak znaleźć ekwiwalent – analiza walencyjna (cs → pl)

Jednym z kryteriów w wyborze właściwego odpowiednika mogłyby być wymogi walencyjne czasownika *toužit* oraz obiekt, z jakim się łączy. Warto więc zbadać, czy istnieje zależność znaczenia czasownika od jego zachowania składniowego[308]. W tym celu wyekscerpowane poświadczenia zostały zapisane w postaci edytowalnej tabelki, otagowane pod względem walencji czeskiego czasownika, następnie pogrupowane pod względem walencji i łączliwości z określonym obiektem.

intercorp_cs	intercorp_pl	przekład	czeski	polski
Ve skutečnosti ==touží== 80% celosvětové populace po stejné životní úrovni, jako má 20% nejbohatších.	W praktyce 80% ludności naszej planety dąży do osiągnięcia takiego samego poziomu życia co 20% ludności najbogatszej.	dążyć	touzit po Oabstr	dążyć do Oabstr
Pořád ==toužím== ze všeho na světě nejvíc po tom, abych ji ještě uviděl, ale nevím, jestli bychom se mohli domluvit, kdybychom se setkali .	Nadal najbardziej ze wszystkiego pragnę ją jeszcze zobaczyć, ale nie wiem, czy zdołalibyśmy się porozumieć, gdybyśmy się spotkali .	pragnąć	touzit po Oabstr	pragnąć inf
Celý život jsem ==toužila== po skutečném domově, a zatím žiju v předsálí nesrozumitelného chrámu, jehož pachy pronikají škvírami nábytku a prosakují všechny předměty.	Całe życie tęskniła m za prawdziwym domem, a tymczasem żyję w przedsionku niepojętej świątyni, której smrody przenikają przez szczeliny mebli i przesiąkają wszystko.	tęsknić	touzit po OR	tęsknić za OR
Mladý muž ==touží== po vlastním divadle.	Młody Mężczyzna marzył o własnym teatrze.	marzyć	touzit po OR	marzyć o OR
Nedaleko Ničína je vesnice ==Toužím== a u ní kopec Palice, kde padly poslední výstřely druhé světové války.	Niedaleko od Niczina jest wioska Toużim, a koło niej góra Palica, gdzie padły ostatnie strzały drugiej wojny światowej.	MIASTO / NAZWA WŁASNA		
Kornet byl rozmrzen a ==netoužil== po vysvětlení, jehož se obával, jako by tento dům skrýval tajemství hrozná a zlá.	Kornet był zły i nie pragnął nawet wyjaśnienia, którego się obawiał, jakby ten dom ukrywał tajemnicę straszną i złą.	pragnąć	touzit po Oabstr	pragnąć Oabstr

Obraz 28. Edytowalna tabelka do tagowania walencji czasownika *toužit* i jego polskich ekwiwalentów

[308] Szerzej o zależności znaczenia jednostki i jej zachowania składniowego w pracy Beth Levin (Levin 1993).

Wyróżnione zostały następujące schematy, w których pojawia się dany czasownik[309]:

- *toužit po* + obiekt abstrakcyjny (dalej – Oabstr)
- *toužit po* + obiekt ludzki lub inny poddany antropomorfizacji (dalej – Ohum)
- *toužit po* + obiekt realny, materialny, rzecz (dalej – OR)
- *toužit do* + obiekt realny, materialny, rzecz (OR)
- *toužit* + bezokolicznik (dalej – inf)
- *toužit, aby* + zdanie podrzędne (dalej – S)
- *toužit po tom, aby* + zdanie podrzędne (S)[310]

[309] Autorka zdaje sobie sprawę, iż schematy te łączą wymagania składniowe i semantyczne. W pierwszej (niepublikowanej wersji tego badania z roku 2012) schematy te obejmowały wyłącznie wymagania składniowe: *toužit po + Obiekt, toužit + bezokolicznik, toužit + zdanie (S)*; wyniki tej analizy wskazały na konieczność rozróżnienia pod względem semantycznym obiektów łączących się z tym czasownikiem, stąd kombinacja podziału semantycznego i składniowego. Podobne schematy semantyczno-składniowe zostały ustalone we wcześniejszej publikacji; tam też zostały podane pierwsze wyniki analizy oparte na materiale z wersji 5 InterCorpu (Kaczmarska i Rosen 2013).

[310] Czasownik *toužit*, jak wynika ze zgromadzonego materiału, łączy się z frazą zdaniową na dwa sposoby – bezpośrednio – *toužit, aby* oraz pośrednio (ze wskaźnikiem zespolenia) – *toužit po tom, aby* por. (Kaczmarska i Rosen 2013: 108).

rodzaj schematu	typ łączliwości	Przykład
***toužit po* Oabstr**	obiekt abstrakcyjny	***Netoužím po** tomhle slizkém **bratrství***[311]. MKZ1991cz
***toužit po* Ohum**	obiekt ludzki	*Jsi krásná, nepřestanu **po tobě toužit** a bát se tvé krásy, miluji tě až do smrti, ale ještě více tě nenávidím*[312]. JDB1993cz
***toužit po* OR**	obiekt realny	*Když jsem si to uvědomil, přepadlo mne málem zoufalství: cítil jsem se tu jako trosečník a **toužil jsem** najednou žíznivě po Praze, po své práci, **po** psacím **stolku** ve svém bytě, **po knihách***[313]. MKZ1991cz
***toužit do* OR**	obiekt realny	*…to tělo již **toužilo do hrobu***[314]. JDB1993cz
***toužit* + inf**	bezokolicznik	*Každý z nás **touží překročit** erotické konvence, erotická tabu, a vstoupit v omámení do království Zakázaného*[315]. MKN1993cz
toužit, aby	zdanie podrzędne	*Všechno, co se kolem ní dělo, ji obtěžovalo a rušilo a ona **toužila, aby** se nedělo nic*[316]. MKN1993cz
toužit po tom, aby	zdanie podrzędne	*Nesmírně **toužila po tom, aby** se s ní oženil, ale bála se, že kdyby k tomu došlo příliš záhy, cítil by se spoután a ztratila by ho pak tím jistěji*[317]. VRR1944cz

Każda z grup została następnie przefiltrowana pod kątem polskiego ekwiwalentu (oraz jego walencji). Wyniki zostały przedstawione w tabelkach poniżej.

[311] *Nie tęsknię za takim oślizłym braterstwem.* MKZ1999pl

[312] *Jesteś piękna, nigdy nie przestanę cię pragnąć i bać się twojej urody, kocham cię aż do śmierci, ale bardziej cię nienawidzę.* JDZ1959pl

[313] *Kiedy to sobie uprzytomniłem, ogarnęła mnie niemalże rozpacz; czułem się tu jak rozbitek i nagle gorąco zatęskniłem za Pragą, za swoją pracą, za biurkiem w swoim mieszkaniu, za książkami.* MKZ1999pl

[314] (…) *to ciało tęskniło już do grobu.* JDZ1959pl

[315] *Każdy z nas pragnie przekroczyć erotyczne konwencje, tabu, i w upojeniu wstąpić w królestwo Zakazu.* MKN2008pl

[316] *Wszystko, co działo się dookoła, zawadzało jej, męczyło, i pragnęła, by nie działo się nic.* MKN2008pl

[317] *Pragnęła bardzo, żeby się z nią ożenił, ale obawiała się, że o ile stanie się to za wcześnie, będzie się czuł spętany, a ona straci go tym prędzej i tym pewniej.* VRK1954pl

toužit po Oabstr

Tabela 25. Czasownik *toužit po Oabstr* w przekładzie na język polski

	ekwiwalent	**liczba**	
toužit po Oabstr	*pragnąć Oabstr*	31	**98**
	marzyć o Oabstr	22	
	tęsknić do Oabstr	11	
	tęsknić za Oabstr	7	
	pożądać Oabstr	5	
	pragnąć inf	4	
	dążyć do Oabstr	2	
	czekać na Oabstr	1	
	imponować Oabstr	1	
	marzący o Oabstr	1	
	marzyć o Ohum	1	
	myśleć o Oabstr	1	
	pożądać Oabstr	1	
	pragnąć + S	1	
	pragnienie Oabstr	1	
	szukać Oabstr	1	
	tęskno za Oabstr	1	
	tęsknota za Oabstr	1	
	upragniony Oabstr	1	
	zapragnąć inf	1	
	zatęsknić Oabstr	1	
	żądać Oabstr	1	
	żądni Oabstr	1	

Schemat ***toužit po Oabstr*** występuje w 98 poświadczeniach (36,5% wszystkich poświadczeń). Jest to najliczniejszy i najbarwniejszy typ pod względem przekładu. W grupie tej, jak wynika z tabelki, najwyższą frekwencję jako ekwiwalent ma połączenie ***pragnąć + Oabstr*** (31 poświadczeń).

Ale zatím chce, abych život snášel a po smrti toužil. JDB1993cz
A tymczasem chce, bym życie znosił, a ***śmierci pragnął****.* JDZ1959pl

Podział semantyczny wszystkich poświadczeń należących do tego typu ukazuje, iż najliczniej reprezentowany jest właśnie polski czasownik *pragnąć* wraz z bliskimi mu jednostkami bliskoznacznymi (49 poświadczeń), co prezentuje poniższa tabelka:

Tabela 26. Podgrupa „pragnąć" jako realizacja frazy *toužit po Oabstr*

pragnąć Oabstr	31	49
pożądać Oabstr	5	
pragnąć inf	4	
dążyć do Oabstr	2	
pożądać Oabstr	1	
pragnąć + S	1	
pragnienie Oabstr	1	
upragniony Oabstr	1	
zapragnąć inf	1	
żądać Oabstr	1	
żądny Oabstr	1	

W podgrupie tej występują też jako ekwiwalenty inne jednostki zilustrowane w tabelce:

Ale Španělé, kteří toužili po slávě zdánlivé, uvěřili tlachu o příchodu císařského vojska z Německa (…). JDB1993cz

*Ale **żądni** pozornej sławy Hiszpanie uwierzyli plotkom o mającej przybyć odsieczy wojsk cesarskich z Niemiec* (…). JDZ1959pl

Jeden znamenal všednost, druhý sen, jeden ji chtěl, druhý nechtěl, jednomu chtěla uniknout, po druhém toužila. MKVNR1997cz

*Jeden uosabiał powszedniość, drugi marzenie, jeden chciał jej, drugi nie chciał, od jednego pragnęła się uwolnić, drugiego **pożądała**.* MKWP2001pl

Milování s ní bylo vždycky dlouhé, skoro nekonečné, protože dosahovala orgasmu, po kterém lačně toužila, jen s velkým úsilím. MKN1993cz

*Ich uściski zawsze trwały długo, niemal bez końca, gdyż G mogła osiągnąć gorąco **upragniony** orgazm po bardzo długim wysiłku.* MKN2008pl

Przybliżoną ilość poświadczeń mają w grupie ***toužit po Oabstr*** także czasowniki *marzyć* i *tęsknić* wraz z formami bliskoznacznymi. Podgrupy te w sumie zbliżają się liczbowo do wyniku wyżej omawianej podgrupy.

Tabela 27. Podgrupa „marzyć” jako realizacja frazy *toužit po Oabstr*

marzyć o Oabstr	22	25
marzący o Oabstr	1	
marzyć o Ohum	1	
myśleć o Oabstr	1	

Tabela 28. Podgrupa „tęsknić” jako realizacja frazy *toužit po Oabstr*

tęsknić do Oabstr	11	21
tęsknić za Oabstr	7	
tęskno za Oabstr	1	
tęsknota za Oabstr	1	
zatęsknić Oabstr	1	

Na uwagę zasługuje fakt, iż w tej podgrupie ”marzyć” znalazło się najwięcej jednostek realizujących schemat ***marzyć o Oabstr***.

(…) *byli jsme zastříkáni, unaveni, promočeni a toužili jsme po odpočinku.* MKZ1991cz
(…) *byliśmy zachlapani, zmęczeni, przemoczeni i* ***marzyliśmy o odpoczynku****.* MKZ1999pl

W podgrupie ”tęsknić” najwięcej poświadczeń odnotowują dwie jednostki: ***tęsknić do Oabstr*** (11 poświadczeń) i ***tęsknić za Oabstr*** (7 poświadczeń).

(…) *a myslím, že není hříchem, toužíme-li v srdci svém po odpočinku a pokoji.* JDB1993cz
(…) *i sądzę, że nie jest grzechem, jeśli w głębi serca* ***tęsknimy za odpoczynkiem i spokojem****. JDZ1959pl*

Toužila po světě, kde lidé mluví jinou řečí než on. MKVNR1997cz
Tęskniła do świata*, w którym ludzie mówią innym językiem niż on. MKWP2001pl*

W grupie tej pojawiają się także przykłady zawierające inne jednostki:

Tabela 29. Podgrupa „inne” jako realizacja frazy *toužit po Oabstr*

czekać na Oabstr	1	3
imponować Oabstr	1	
szukać Oabstr	1	

Po posledním pomazání ve Vojenské nemocnici toužili dva (…). JHODV1996cz
Na ostatnie namaszczenie czekali w szpitalu dwaj ranni (…) JHPDW1957pl

Víte, Ludvíku, já bych byla strašně nešťastná, kdybyste si myslel, že jsem panička, jako jsou ostatní paničky, co se nudí a touží po dobrodružství. MKZ1991cz
Wie pan, Ludwiku, czułabym się strasznie nieszczęśliwa, gdyby pan sobie pomyślał, że jestem taka, jak te znudzone damulki, co to ***szukają przygód****.* MKZ1999pl

Typ ten został określony jako najliczniejszy i zarazem najbarwniejszy pod względem przekładu; stało się tak również za sprawą samych obiektów abstrakcyjnych, które obejmują szeroką gamę różnych zjawisk. Nie zaobserwowano jednak żadnej reguły łączenia się jakiegoś konkretnego typu obiektu z jednym z ekwiwalentów.

toužit po Ohum

Tabela 30. Czasownik *toužit po Ohum* w przekładzie na język polski

	ekwiwalent	liczba	
toužit po Ohum	*pragnąć Ohum*	12	**41**
	tęsknić za Ohum	6	
	pożądać Ohum	5	
	tęsknić do Ohum	5	
	marzyć o Ohum	2	
	pragnąć Oabstr	2	
	ciągnąć do Ohum	1	
	mieć ochotę + inf	1	
	pożądany Ohum	1	
	pragnąć + inf	1	
	tęsknić do + S	1	
	tęsknić do Oabstr	1	
	zapragnąć Oabstr	1	
	zatęsknić za Ohum	1	
	żywić miłość do Ohum	1	

Schemat ***toužit po Ohum*** występuje w 41 poświadczeniach, co stanowi 15,25% wszystkich wystąpień. Zgromadzone dane prezentują różne możliwości przekładu na język polski czasownika *toužit* łączącego się z obiektem ludzkim lub obiektem

antropomorfizowanym. Jako ekwiwalent najczęściej pojawia się jednostka *pragnąć* (również *zapragnąć*) oraz wyrazy bliskoznaczne (*pożądać, mieć ochotę*):

Tabela 31. Podgrupa „pragnąć" jako realizacja frazy *toužit po Ohum*

pragnąć Ohum	12	23
pożądać Ohum	5	
pragnąć Oabstr	2	
mieć ochotę + inf	1	
pożądany Ohum	1	
pragnąć + inf	1	
zapragnąć Oabstr	1	

Viděla jsem to a zalila mne horkost z radosti, že po mně touží, a že po mně touží o to víc, když mu připomínám, že jsem vdaná, protože se mu tím vzdaluju, a člověk touží vždycky nejvíc po tom, co se mu vzdaluje (…). MKZ1991cz

Dostrzegłam to i zrobiło mi się gorąco z radości, że ***mnie pragnie****, że pragnie mnie tym więcej, bo przypominam mu, iż jestem mężatką, tym samym bowiem oddalam się od niego, a człowiek zawsze najbardziej pragnie tego, co mu się wymyka* (…). MKZ1999pl

W poświadczeniach pojawiają się również pozostałe, proponowane przez słowniki odpowiedniki; wśród nich najczęściej – *tęsknić*:

Tabela 32. Podgrupa „tęsknić" jako realizacja frazy *toužit po Ohum*

tęsknić za Ohum	6	14
tęsknić do Ohum	5	
tęsknić do + S	1	
tęsknić do Oabstr	1	
zatęsknić za Ohum	1	

(…) *toužil jsem po ní, jako se touží po něčem, co je definitivně ztracené.* MKZ1991cz

(…) ***tęskniłem za nią****, jak tęskni się za czymś, co utraciło się na zawsze.* MKZ1999pl

Jedynie dwa poświadczenia zawierały czasownik *marzyć*:

Tabela 33. Podgrupa „marzyć" jako realizacja frazy *toužit po Ohum*

marzyć o Ohum	2

Vždycky jsem toužila po člověku, který by byl prostý a přímý. MKZ1991cz

Zawsze ***marzyłam o człowieku****, który byłby prosty i bezpośredni.* MKZ1999pl

Pojedynczo pojawiają się także przykłady zawierające inne jednostki (m.in. *żywić miłość*).

Tabela 34. *Podgrupa „inne" jako realizacja frazy toužit po Ohum*

ciągnąć do Ohum	1	2
żywić miłość do Ohum	1	

Anebo vám připomenu tu nádhernou scénu z Fromentinova Dominika: oba zamilovaní, kteří po sobě léta toužili a nikdy se jeden druhého nedotkli (…) MKN1993cz

*Przypomnijcie sobie także wspaniałą scenę z Dominika Fromentina: kochankowie, którzy przez lata **żywili do siebie miłość**, lecz się nie dotykali* (…) MKN2008pl

To by musela mít jinak rostlý oči, aby si všimla, že nejenom mužský toužej po oblejch stehnech a kulatejch ňadrech do dlaní. PHPMP2002cz

*Musiałyby jej inaczej wyrosnąć oczy, żeby zauważyła, że nie tylko chłopów **ciągnie do** obłych ud i krągłych piersi, w sam raz pasujących do dłoni.* PHCCG2007pl

Na podstawie przeprowadzonego badania można by stwierdzić, że czasownik *toužit*, łączący się z obiektem ludzkim, jest najbliższy znaczeniowo polskiemu czasownikowi *pragnąć* i *pożądać*. Istnieje jednak niebezpieczeństwo, iż dysponując niewielką liczbą danych, opracuje się całą regułę w oparciu o język jednego tłumacza. Warto więc przyjrzeć się przykładom, a w nich powtarzającym się kontekstom, w których czasownik *toužit* się pojawia[317].

Pierwszą powtarzającą się sytuacją jest afekt i kontekst erotyczny połączony z ekwiwalentem *pragnąć* czy *pożądać*:

Měl pocit, jako by právě po ní toužil už po mnoho let, jako by ji celou tu dobu hledal po celém světě. MKN1993cz

*Poczuł, że to jej właśnie **pragnął** przez całe lata, że to jej szukał po całym świecie.* MKN2008pl

Kdysi toužil jen po nových ženách. MKN1993cz
*Niegdyś **pożądał** tylko nowych kobiet.* MKN2008pl

Toužila po něm. MKNLB1985cz
***Pragnęła** jego obecności.* MKNLB1992pl

Nemohl jsem najednou vůbec pochopit, proč jsem tak nepříčetně toužil po jejím těle (…) MKZ1991cz

[317] Por. przykłady kontekstów z czasownikiem *toužit* (Kaczmarska i Rosen 2013: 109–110).

Nie mogłem naraz absolutnie pojąć, dlaczego tak niepoczytalnie ***pragnąłem*** *jej ciała* (…) MKZ1999pl

Pojawiają się też sytuacje, w których występują czasowniki *pragnąć* i *tęsknić* połączone z obiektem wyrażającym części ciała, które „reprezentują" całą osobę[318]:

Toužila po jeho prací poničených rukách, po jeho temně opáleném zátylku, po vůni laciné březové vody v jeho vlasech. MVUZ2001cz

Pragnęła *jego zniszczonych pracą rąk, ciemnego, opalonego karku, zapachu taniej wody brzozowej we włosach.* MVW2008pl

Oči, po kterých touží, jsou oči Tomáše. MKNLB1985cz
Oczy, do których ***tęskni****, to oczy Tomasza.* MKNLB1992pl

Toužila po svém vlastním těle, náhle objeveném, nejbližším i nejcizejším a nejvíc vzrušujícím. MKNLB1985cz

Tęskniła *do własnego ciała, nagle odkrytego, najbliższego i najbardziej obcego, najbardziej podniecającego.* MKNLB1992pl

Użycie czasownika *pragnąć* jako ekwiwalentu *toužit* wiąże się też z sytuacjami, kiedy wyrażane jest pragnienie / gorące chcenie:

Copak tys nikdy netoužil po dítěti? MKVNR1997cz
Czy nigdy nie ***pragnąłeś*** *mieć dziecka?* MKWP2001pl

choć czasem też w takich sytuacjach pojawia się czasownik *marzyć*:

Vždycky jsem toužila po člověku, který by byl prostý a přímý. MKZ1991cz
Zawsze ***marzyłam o człowieku****, który byłby prosty i bezpośredni.* MKZ1999pl

toužit po OR

Schemat *toužit po OR* powtórzył się w wyekscerpowanym materiale tylko 13 razy.

[318] Tęsknotę za osobą (istotą), ale bez kontekstu seksualnego wyrażają poświadczenia, w których pojawia się również czasownik *tęsknić*:
Tys neznala Boha, Lucie, ale toužila jsi po něm. MKZ1991cz
Tyś nie znała Boga, Łucjo, ale ***tęskniłaś*** *do Niego.* MKZ1999pl

Proč by měl Goethe toužit po Herderovi? MKN1993cz
Dlaczego zatem Goethe miałby ***tęsknić*** *do pośmiertnej obecności Herdera?* MKN2008pl

Dlouhé hodiny se jí v duchu zuřivě vysmíval a dlouhé hodiny po ní zoufale toužil. MVUZ2001cz
Całymi godzinami naśmiewał się z niej w duchu i całymi godzinami rozpaczliwie za nią ***tęsknił****.* MVW2008pl

Tabela 35. Czasownik *toužit po OR* w przekładzie na język polski

	ekwiwalent	**liczba**	
toužit po OR	*marzyć o OR*	5	**13**
	pragnąć OR	3	
	tęsknić za OR	2	
	marzyć + S	1	
	pożądanie	1	
	pragnąć + inf	1	

Tak mała liczba poświadczeń nie może być podstawą żadnej głębszej analizy. W przypadku tej grupy trudno mówić o jakiejś tendencji; najwięcej przykładów zawiera jednostkę *marzyć*, ale pozostałe czasowniki – *pragnąć* i *tęsknić* – proponowane przez słowniki, są również obecne:

Mladý muž touží po vlastním divadle. RDMM2009cz
*Młody mężczyzna **marzył** o własnym teatrze.* RDMM2009pl

Celý život jsem toužila po skutečném domově (…) MADM1993cz
*Całe życie **tęskniłam** za prawdziwym domem* (…) MAIN2005pl

Mé patro, vyprahlé po noci zpola probdělé a zpola neklidně prosněné, toužilo po jejím vřelém a mrazivě vonném doušku. VRR1944cz
*Moje podniebienie, wyschnięte po nocy na wpółprzemarzonej i na wpół prześnionej, **pragnęło** jej gorącego, orzeźwiająco wonnego łyku.* VRK1954pl

toužit inf

Tabela 36. Czasownik *toužit inf* w przekładzie na język polski

	ekwiwalent	**liczba**	
toužit inf	*pragnąć inf*	55	**95**
	chcieć inf	21	
	pragnąć Oabstr	5	
	marzyć o Oabstr	4	
	mieć ochotę inf	3	
	zapragnąć	2	
	dbać o Oabstr	1	
	pragnąć + S	1	
	pragnienie inf	1	
	tęsknić za + S	1	
	zachciewać się Oabstr	1	

Schemat ***toužit inf*** występuje w 95 poświadczeniach (35,32% wszystkich wystąpień). Jak wynika z analizy i tabeli, jest to typ dość jednolity, jeżeli chodzi o przekład na język polski. W zdecydowanej większości przykładów pojawia się czasownik *pragnąć* lub jednostki mu bliskie.

Tabela 37. Podgrupa „pragnąć" jako realizacja frazy *toužit inf*

pragnąć inf	55	89
chcieć inf	21	
pragnąć Oabstr	5	
mieć ochotę inf	3	
zapragnąć inf	2	
pragnąć + S	1	
pragnienie inf	1	
zachciewać się Oabstr	1	

Netouží být po smrti ani s Paulem ani s Brigitou (…)
*Nie **pragnęła** również **spotkać się** z Paulem ani z Brigitte* (…)

Každý z nás touží překročit erotické konvence, erotická tabu, a vstoupit v omámení do království Zakázaného. MKN1993cz
*Każdy z nas **pragnie przekroczyć** erotyczne konwencje, tabu, i w upojeniu wstąpić w królestwo Zakazu.* MKN2008pl

Začal znovu skicovat do náčrtníku obrazy, které toužil namalovat. MKN1993cz
*Zaczął znowu szkicować w notatniku obrazy, które **pragnął namalować**.* MKN2008pl

Já si taky umím představit, že si člověk touží vzít život. MKN1993cz
*Ja również mogę zrozumieć, że człowiek **chce** ze sobą **skończyć**.* MKN2008pl

Nemohu než to ještě jednou zdůraznit: netoužila vidět pohlaví cizího muže. MKNLB1985cz
*Nie mogę nie podkreślić tego jeszcze raz: nie **chciała oglądać** organów płciowych obcego mężczyzny.* MKNLB1992pl

Když on toužil mluvit sprostě, ona zarputile mlčela. MKN1993cz
*Gdy jemu **zachciewało się wulgarnych słów**, ona uparcie milczała.* MKN2008pl

W przypadku tej podgrupy można zaryzykować stwierdzenie o wpływie walencji na znaczenie i rozumienie tego czasownika. Czasowniki *pragnąć, chcieć* oraz formy formalnie im bliskie (także derywaty) łączące się z bezokolicznikiem stanowią 86,4% wszystkich ekwiwalentów w tej grupie, co dokumentują poniższe tabele:

Tabela 38. Podgrupa „pragnąć” + inf jako ekwiwalent *toužit inf*

pragnąć inf	55	82
chcieć inf	21	
mieć ochotę inf	3	
zapragnąć inf	2	
pragnienie inf	1	

Pozostałe przykłady to pojedyncze użycia trzech czasowników; żaden z nich nie powiela schematu z bezokolicznikiem.

Tabela 39. Podgrupa „marzyć” jako realizacja frazy *toužit inf*

marzyć o Oabstr	4

Teprve během těch tří dnů s ním mohla být tak, jak s ním vždycky toužila být. MKN1993cz
Te trzy dni były jedynymi, które udało jej się spędzić przy ojcu w warunkach, ***o*** *jakich zawsze* ***marzyła****.* MKN2008pl

Tabela 40. Podgrupa „tęsknić” jako realizacja frazy *toužit inf*

tęsknić za + S	1

Nakonec toužíte poznat někoho obyčejnýho. MVPNK2003cz
W końcu ***tęskni*** *się* ***za tym, by*** *poznać kogoś zwyczajnego.* MVSNK2007pl

Tabela 41. Podgrupa „inne” jako realizacja frazy *toužit inf*

dbać o Oabstr	1

Toužil být milován láskou absolutní (…) MKN1993cz
Dbał o *miłość absolutną…* (…) MKN2008pl

toužit, aby +S / toužit po tom, aby +S

W obu pozostałych typach pojawiają się w sumie 22 poświadczenia, w których czasownik *toužit* łączy się ze zdaniem podrzędnym. Przeważająca liczba przykładów zawiera jednostkę *pragnąć* (również *chcieć* i *życzyć sobie*), *marzyć* pojawia się trzy razy, natomiast *tęsknić* w zebranym materiale nie pojawia się w ogóle, co ilustrują poniższe tabelki:

Tabela 42. Czasownik *toužit aby +S* w przekładzie na język polski

	ekwiwalenty	**liczba**	
toužit, aby + S	*pragnąć + S*	11	**17**
	pragnąć + inf	3	
	marzyć + S	2	
	życzyć sobie + S	1	

Touží, aby ji někdo odnaučil být anachronická! MKNLB1985cz
***Marzy**, by ktoś nauczył ją, jak nie być anachroniczną!* MKNLB1992pl

Toužila tehdy, aby už jednou skončila ta nebezpečná cesta zrad. MKNLB1985cz
***Pragnęła** wówczas, żeby raz wreszcie skończyła się ta niebezpieczna droga zdrad.* MKNLB1992pl

Toužila, aby přišel, toužila, aby ji k sobě pozval! MKNLB1985cz
***Pragnęła**, żeby przyszedł, pragnęła, aby zaprosił ją do siebie!* MKNLB1992pl

Nesmírně toužila po tom, aby se s ní oženil. VRR1944cz
***Pragnęła** bardzo, żeby się z nią ożenił.* VRK1954pl

Tabela 43. Czasownik ***t**oužit po tom, aby + S* w przekładzie na język polski

	ekwiwalent	**liczba**	
toužit po tom, aby + S	*pragnąć + S*	2	**5**
	chcieć + S	1	
	marzyć się + Nom	1	
	pragnąć inf	1	

Netoužil jsem nikdy po tom, aby se za mnou ohlíželi s obdivem jako Vilém Haba, chtěl jsem jen být mezi nimi doma. VRR1944cz
*Nigdy nie **pragnąłem**, jak Haba, żeby oglądali się za mną z podziwem, chciałem tylko czuć się wśród nich jak w domu.* VRK1954pl

Wyszukiwanie: czeskie przekłady ← polskie oryginały (cs ← pl)

Ta część analizy to ekscerpcja z tekstów czeskich przekładów wyrównanych z polskimi tekstami oryginalnymi. Poszukiwane były polskie frazy, których czeskim ekwiwalentem był czasownik *toužit*. Znaleziono 243 poświadczeń, z czego 10 nie zostanie

poddanych analizie; siedem z nich to celowe błędy polskiego autora, a dwa to błędy wyrównania[319]. Dalsza metoda postępowania była podobna do technik wykorzystanych w pierwszym etapie, jednakże zastosowana w wersji uproszczonej. Prowadząc badanie w odwrotnym kierunku (polski oryginał i jego przekład na język czeski), autorka próbowała wyszukać, przekładem jakich jednostek języka polskiego jest czasownik *toužit*.

Tabela 44. Polskie frazy tłumaczone jako *toužit* (polskie oryginały / czeskie przekłady

polskie frazy tłumaczone jako *toužit*	**liczba poświadczeń**	**suma poświadczeń**
pragnąć inf	44	
pragnąć Oabstr	37	
chcieć inf	11	
tęsknić do Oabstr	11	
marzyć o Oabstr	10	
pragnąć OR	8	
pragnąć Ohum	7	
pożądać Oabstr	5	
spragniony Oabstr	5	
chcieć Oabstr	4	
pragnąć X	4	
pragnienie Oabstr	3	
tęsknić do Ohum	3	
tęsknić za Oabstr	3	
chcieć OR	2	
dążyć do Oabstr	2	
pożądać Ohum	2	
pragnąć, aby + S	2	
tęsknić za OR	2	
złakniony Oabstr	2	
chcieć Ohum	1	
chcieć X	1	
chęć Oabstr	1	
chęć, żeby + S	1	
chętnie by + V	1	
chodzić o Oabstr	1	
dbać o Oabstr	1	
łaknące się Ohum	1	
łaknąć Oabstr	1	

[319] W poprzednim badaniu w podobnym wyszukiwaniu odnaleziono jedynie 149 poświadczeń, przy czym z całkowej liczby 149 przykładów 6 nie podlegało analizie. Były to błędy celowe autora oryginału polskiego (5 przykładów) oraz błąd wiązania segmentów tekstu – 1 przykład (Kaczmarska i Rosen 2013).

polskie frazy tłumaczone jako *toužit*	**liczba poświadczeń**	**suma poświadczeń**
łaknąć Ohum	1	
łaknąć OR	1	
marzący o Oabstr	1	
marzenie inf	1	
marzyć inf	1	
marzyć o Ohum	1	
marzyć o OR	1	
marzyć o tym, aby + S	1	
marzyć się Oabstr	1	
marzyć, aby + S	1	
marzyć, że + S	1	
mieć ochotę inf	1	
nie chciało się Oabstr	1	
odechcieć się Oabstr	1	
postanowić inf	1	
pożądające Ohum	1	
pożądanie OR	1	
pożądany Oabstr	1	
pragnący inf	1	
pragnienie inf	1	
pragnienie, aby + S	1	
próbować inf	1	
przepadać za OR	1	
spróbować inf	1	
tęskniąc do Oabstr	1	
tęsknić za to, żeby + S	1	
tęsknić za tym, żeby + S	1	
upragniony Oabstr	1	
usiłować inf	1	
wzdychać, aby + S	1	
zachcieć się Ohum	1	
zamarzyć inf	1	
zapragnąć inf	1	
zapragnąć Oabstr	1	
żądać OR	1	
żądny Oabstr	1	
zmiana struktury	18	
błąd	10	
frazeologizm	3	243

W wyekscerpowanym materiale przeważają poświadczenia z polskimi jednostkami z podgrupy „pragnąć"; znaleziono 164 rekordy, w których *toužit* jest odpowiednikiem polskich jednostek.

Tabela 45. Polskie jednostki podgrupy „pragnąć" tłumaczone na język czeski jako *toužit*

pragnąć inf	44	
pragnąć Oabstr	37	
chcieć inf	11	
pragnąć OR	8	
pragnąć Ohum	7	
pożądać Oabstr	5	
spragniony Oabstr	5	
chcieć Oabstr	4	
pragnąć X	4	
pragnienie Oabstr	3	
chcieć OR	2	
dążyć do Oabstr	2	
pożądać Ohum	2	
pragnąć, aby + S	2	
złakniony Oabstr	2	
chcieć Ohum	1	
chcieć X	1	
chęć Oabstr	1	
chęć, żeby + S	1	
chętnie by + V	1	
łaknące się Ohum	1	
łaknąć Oabstr	1	
łaknąć Ohum	1	
łaknąć OR	1	
mieć ochotę inf	1	
nie chciało się Oabstr	1	
odechcieć się Oabstr	1	
postanowić inf	1	
pożądające Ohum	1	
pożądanie OR	1	
pożądany Oabstr	1	
pragnący inf	1	
pragnienie inf	1	
pragnienie, aby + S	1	
upragniony Oabstr	1	
usiłować inf	1	
zachcieć się Ohum	1	
zapragnąć inf	1	
zapragnąć Oabstr	1	
żądać OR	1	
żądny Oabstr	1	**164**

Mniej liczna jest podgrupa „tęsknić". Należą do niej 22 jednostki:

Tabela 46. Polskie jednostki podgrupy „tęsknić" tłumaczone na język czeski jako *toužit*

tęsknić do Oabstr	11	**22**
tęsknić do Ohum	3	
tęsknić za Oabstr	3	
tęsknić za OR	2	
tęskniąc do Oabstr	1	
tęsknić za to, żeby + S	1	
tęsknić za tym, żeby + S	1	

Dwadzieścia rekordów, w których *toužit* jest odpowiednikiem polskich jednostek, to jednostki należące do podgrupy „marzyć".

Tabela 47. Polskie jednostki podgrupy „marzyć" tłumaczone na język czeski jako *toužit*

marzyć o Oabstr	10	**20**
marzący o Oabstr	1	
marzenie inf	1	
marzyć inf	1	
marzyć o Ohum	1	
marzyć o OR	1	
marzyć o tym, aby + S	1	
marzyć się Oabstr	1	
marzyć, aby + S	1	
marzyć, że + S	1	
zamarzyć inf	1	

Jedynie sześć rekordów, w których *toužit* jest odpowiednikiem polskich jednostek, to jednostki należące do podgrupy „inne".

Tabela 48. Polskie jednostki podgrupy „inne" tłumaczone na język czeski jako *toužit*

chodzić o Oabstr	1	**6**
dbać o Oabstr	1	
próbować inf	1	
przepadać za OR	1	
spróbować inf	1	
wzdychać, aby + S	1	

Stosunkowo liczną grupę stanowią również przykłady, w których czeski czasownik *toužit* okazał się „wspólnym mianownikiem" dla różnych struktur polskich. W strukturach tych nie sposób ustalić, której konkretnie jednostce miałby odpowiadać czeski czasownik[320]. Takich poświadczeń odnotowano 18, np.:

Sensacyjne morderstwo, popełnione pod lampą w Aller, sprawiło, że stanowiły towarzystwo szczególnie atrakcyjne. JCWC2001pl
Vzhledem k senzačnímu zločinu, jaký se odehrál pod lampou v Allerodu, teď zřejmě kdekdo ***toužil*** *pobývat v naší společnosti.* JCVC2005cz

Twierdziła*, że musi tam być, że będzie sama, bez żywej znajomej polskiej duszy, i koniecznie chce, żebyśmy przyszły, jeśli już nie dla wrażeń artystycznych, to przynajmniej dla towarzystwa.* JCWC2001pl
Tvrdila, že se jí musí účastnit, ale bude se prý cítit opuštěná, bez podpory živé duše neboli svých krajanů, a proto jí moc záleží na tom, abychom se tam dostavili. Jestli ***netoužíme*** *po uměleckém zážitku, aspoň bychom jí mohli dělat společnost.* JCVC2005cz

Czekaliśmy w napięciu, zasłuchani w niezwykłą opowieść. JCWC2001pl
Napjatě jsme čekali, co z něho vypadne, a ***toužili*** *slyšet pokračování nezvyklého příběhu. JCVC2005cz*

Odrębną – nową – grupę stanowią polskie frazeologizmy, które zostały przetłumaczone jako *toužit*. Jest to jednak bardzo nieliczna grupa, zawierająca jedynie trzy poświadczenia[321] (trzy różne frazeologizmy) i jako taka nie będzie dalej analizowana.

Ale doradź generalnie: jak go traktować... on w ogóle ***nie ma nosa*** *do postępu.* ERA1973pl
Ale poraď generelně: jak se k němu chovat... on vůbec ***netouží*** *po pokroku. ERV1977cz*

W wyekscerpowanym materiale pojawiły się też nowe schematy walencyjne zarówno polskich, jak i czeskich jednostek, np. *chcieć X, pragnąć X* czy *toužit X*, gdzie X oznacza niewypełnioną pozycję walencyjną[322].

I zawsze: „A dlaczego?" jakby samo to, że ludzie ***pragną****, dążą i cierpią, nie wystarczało.* CMDI1981pl
Vždycky „a pročpak?", jako by nestačilo jen to, že lidé ***touží****, snaží se a trpí.* CMUI1993cz

[320] Struktury te nie mają wspólnych rysów, nie są też podobne ani pod względem semantycznym, ani syntaktycznym. Nie będą brane pod uwagę w dalszej analizie w tym rozdziale.
[321] Są to trzy różne frazeologizmy, z trzech różnych źródeł.
[322] Wszystkie schematy przedstawione zostaną w kolejnych tabelach.

toužit po Oabstr

Tabela 49. Polskie jednostki tłumaczone na język czeski jako *toužit po Oabstr*

toužit po Oabstr	**93**	*pragnąć Oabstr*	35
		marzyć o Oabstr	9
		tęsknić do Oabstr	9
		pożądać Oabstr	5
		chcieć Oabstr	4
		spragniony Oabstr	4
		tęsknić za Oabstr	3
		pragnąć inf	2
		złakniony Oabstr	2
		chcieć inf	1
		chcieć X	1
		dążyć do Oabstr	1
		dbać o Oabstr	1
		chodzić o Oabstr	1
		łaknąć Oabstr	1
		marzyć się Oabstr	1
		tęskniąc do Oabstr	1
		upragniony Oabstr	1
		zapragnąć inf	1
		zapragnąć Oabstr	1
		żądny Oabstr	1
		struktura	5
		frazeologizm	3

W wersji 10 korpusu InterCorp wyszukano 93 poświadczenia ze schematem *toužit po Oabstr*, będącym przekładem polskiej jednostki. W dwóch trzecich przykładów w polskiej wersji (oryginał) pojawia się jednostka *pragnąć* (oraz *chcieć*, *pożądać, spragniony*).

Tabela 50. Podgrupa polskich fraz oryginalnych „pragnąć", których czeskim przekładem jest fraza *toužit po Oabstr*

pragnąć Oabstr	35	60
pożądać Oabstr	5	
chcieć Oabstr	4	
spragniony Oabstr	4	
pragnąć inf	2	
złakniony Oabstr	2	
chcieć inf	1	
chcieć X	1	
dążyć do Oabstr	1	
łaknąć Oabstr	1	
upragniony Oabstr	1	
zapragnąć inf	1	
zapragnąć Oabstr	1	
żądny Oabstr	1	

Ludność była zmęczona, po udanej rewolucji ***pragnęła*** *spokoju i wytchnienia.* GHGDPN1993pl

Lid byl unaven, po zdařilé revoluci ***toužil*** *po klidu a vydechnutí.* GHGDP1995cz

W oryginałach występują jednostki z grupy „tęsknić" (*tęsknić do, tęsknić za, tęskniąc*). Poświadczeń takich znaleziono jedynie 13 poświadczeń:

Tabela 51. Podgrupa polskich fraz oryginalnych „tęsknić", których czeskim przekładem jest fraza *toužit po Oabstr*

tęsknić do Oabstr	9	13
tęsknić za Oabstr	3	
tęskniąc do Oabstr	1	

Tęskniłabym *tylko do wilgoci, wystawiałabym moje ciało do mgieł i deszczu, skraplałabym na sobie mokre powietrze.* OTDD1998pl

Toužila *bych jen po vlhku, své tělo bych vystavovala mlze a dešti.* OTDD2002cz

Dziesięć poświadczeń zawiera jednostki z grupy "marzyć" (*marzyć o, marzyć się*):

Tabela 52. Podgrupa polskich fraz oryginalnych „marzyć", których czeskim przekładem jest fraza *toužit po Oabstr*

marzyć o Oabstr	9	10
marzyć się Oabstr	1	

Jednakowoż wiemy teraz na pewno, że gdy po planetach będą się przechadzali pierwsi wysłannicy Ziemi, inni jej synowie nie o wyprawach takich będą ***marzyć****, lecz o kawałku chleba.* SLGP1984pl

Přesto však nyní víme s jistotou, že až se po planetách budou procházet první vyslanci Země, jiní její synové nebudou ***toužit*** *po takových výpravách ale po kousku chleba.* SLPH1981cz

W poświadczeniach pojawiają się też dwa przykłady, w których czeski czasownik jest odpowiednikiem polskich jednostek *dbać o* i *chodzić o* (podgrupa „inne"):

Tabela 53. Podgrupa polskich fraz oryginalnych „inne", których czeskim przekładem jest fraza *toužit po Oabstr*

dbać o Oabstr	1	2
chodzić o Oabstr	1	

Popychał swój wózeczek, popychany przez zbiegowisko, a ja wycofywałem się pospiesznie, przerażony, z jego pobliża, ***dbając*** *o to tylko, żeby jak najszybciej wyjść, a właściwie uciec.* SLGP1984pl

Strkal svůj vozík, sám postrkován davem, a já jsem spěšně vycouval, jako by mě jeho blízkost vyděsila. ***Netoužil*** *jsem po ničem jiném, než abych co nejrychleji odešel, vlastně utekl.* SLPH1981cz

toužit po Ohum

Tabela 54. Polskie jednostki tłumaczone na język czeski jako *toužit po Ohum*

toužit po Ohum	**19**	*pragnąć Ohum*	7
		pożądać Ohum	2
		tęsknić do Ohum	2
		chcieć Ohum	1
		łaknące się Ohum	1
		łaknąć Ohum	1
		marzyć o Ohum	1
		pożądające Ohum	1
		pożądany Oabstr	1
		zachcieć się Ohum	1
		struktura	1

W wyekscerpowanych poświadczeniach znalazło się 19 rekordów, w których występuje schemat *toužit po Ohum* jako odpowiednich polskich fraz. Większość z nich to jednostki z podgrupy „pragnąć”

Tabela 55. Podgrupa polskich fraz oryginalnych „pragnąć”, których czeskim przekładem jest fraza *toužit po Ohum*

pragnąć Ohum	7	15
pożądać Ohum	2	
chcieć Ohum	1	
łaknące się Ohum	1	
łaknąć Ohum	1	
pożądające Ohum	1	
pożądany Oabstr	1	
zachcieć się Ohum	1	

*Pogardzał sobą, że jej ciągle **pragnął**, zapominał o pogardzie, kiedy ją miał w łóżku, później w momencie najbardziej fizycznym otwierała się ranka w duszy, i przyjemność diabli brali.* MKGO1961pl

*Pohrdal sebou, že po ní stále **touží**, zapomínal na pohrdání, když ji měl v posteli, ale potom v okamžiku vrcholně fyzickém se mu v duši otevírala ranka a všechnu rozkoš vzal čert.* MKOH1971cz

Polski schemat *tęsknić do Ohum* pojawia się tylko w dwóch poświadczeniach.

Tabela 56. Podgrupa polskich fraz oryginalnych „tęsknić", których czeskim przekładem jest fraza *toužit po Ohum*

tęsknić do Ohum	2

A ona między drzewa, co na placu były jak Łasica wbiegła i, w ich cień się schroniwszy, stamtąd ***tęsknić*** *do niego i wzdychać zaczęła.* WGTA2000pl

A ona vběhla jako Lasička mezi stromy, ukryla se do jejich stínu a odtamtud po něm začala ***toužiti*** *a vzdychati.* WGTA2007cz

Natomiast schemat *marzyć o Ohum* pojawia się tylko raz.

Tabela 57. Podgrupa polskich fraz oryginalnych „marzyć", których czeskim przekładem jest fraza *toužit po Ohum*

marzyć o Ohum	1

(…) *właśnie z powodu tej niemęskiej dykcji mówił cicho, jakby się wstydząc, że ktoś pomyśli o nim jak o homoseksualiście, a przecież brat starego K.* ***marzył*** *o kobietach.* WKG2003pl

(…) *právě díky své zženštilé dikci mluvil potichu, jako by se styděl, že si o něm lidé pomyslí, že je homosexuál, bratr starého K. však* ***toužil*** *po ženách.* WKS2009cz

toužit po OR

Tabela 58. Polskie jednostki tłumaczone na język czeski jako *toužit po OR*

toužit po OR	**17**	*pragnąć OR*	8
		chcieć OR	2
		tęsknić za OR	2
		łaknąć OR	1
		marzyć o OR	1
		pożądanie OR	1
		przepadać za OR	1
		żądać OR	1

W wyekscerpowanym materiale znalazło się 17 poświadczeń, w których występuje schemat *toužit po OR* jako odpowiednich polskich fraz. Trzynaście z nich to jednostki z podgrupy „pragnąć"

Tabela 59. Podgrupa polskich fraz oryginalnych „pragnąć", których czeskim przekładem jest fraza *toužit po OR*

pragnąć OR	8	13
chcieć OR	2	
łaknąć OR	1	
pożądanie OR	1	
żądać OR	1	

Jezusie Nazareński, jeszcze nigdy nie ***chciało*** *mi się tak ani garnuszka wody* (…) TNAJK1968pl

Kriste Pane, ještě nikdy jsem ***netoužil*** *tolik po hrníčku vody* (…) TNABK1984cz

Późną noc przeznaczam na czytanie książek, które, niestety, nie zawsze są takie jakich bym ***pragnął****.* WGD1988pl

Pozdní noc je určena na čtení knih, které bohužel nejsou zrovna těmi, po kterých bych ***toužil****.* WGD1994cz

Pozostałe rekordy ze schematem *toužit po OR* w przekładzie na język czeski to pojedyncze poświadczenia zawierające w oryginale jednostki *tęsknić za* (2), *marzyć* (1) i *przepadać za* (1). Poniżej tabelki wraz z przykładami:

Tabela 60. Podgrupa polskich fraz oryginalnych „tęsknić", których czeskim przekładem jest fraza *toužit po OR*

tęsknić za OR	2

Czasami przejechał piękny, opływowy mercedes, ale Dianka już nie ***tęskniła*** *za mercedesami, tylko za własnym pokojem w bloku, w Bratysławie, za kolacją zrobioną przez mamę, za herbatą i odrabianiem lekcji.* MWL2005pl

Občas kolem projel pěknej aerodynamickej mercedes, jenže Dianka už ***netoužila*** *po mercedesech, toužila po svým pokoji v paneláku, v Bratislavě, po večeři od mámy, čaji a psaní úloh.* MWC2007cz

Tabela 61. Podgrupa polskich fraz oryginalnych „marzyć", których czeskim przekładem jest fraza *toužit po OR*

marzyć o OR	1

Chciała kupić sobie szal, o którym od dawna ***marzyła****.* AKKOS1970pl

Chtěla si koupit šál, po kterém odedávna ***toužila****.* AKKOS1975cz

Tabela 62. Podgrupa polskich fraz oryginalnych „inne", których czeskim przekładem jest fraza *toužit po OR*

przepadać za OR	1

Trurl, który ***przepadał*** *za ciepłymi krajami, myślał o Ogniolii, państwie Płomienionogów, Klapaucjusz natomiast, chłodniejszego usposobienia, wybrał galaktyczny biegun zimna, kontynent czarny wśród gwiazd lodowatych.* SLC2002pl

Trurl ***toužil*** *po teplých krajinách a navrhoval Thermonii, zemi planoucích plameňáků, zatímco Klapacius, chladnějšího založení, hodlal odcestovat k intergalaktickému pólu zimy, pustému to kontinentu uprostřed zledovatělých hvězd.* SLK1983cz

toužit X

Tabela 63. Polskie jednostki tłumaczone na język czeski jako *toužit X*

toužit X	**5**	*pragnąć X*	4
		struktura	1

W wyekscerpowanych poświadczeniach znaleziono zaledwie 5 rekordów ze schematem *toužit X*, będących przekładem polskiej jednostki, przy czym 4 z nich to *pragnąć X*. W tych przypadkach ani w języku polskim, ani w języku czeskim nie pojawia się obiekt.

Nie tęsknić, nie ***pragnąć****.* OTDD1998pl

Netěšit se, ***netoužit****.* OTDD2002cz

Jeden przykład odnosi się do sytuacji, w której jako *toužit X* została przełożona inna polska struktura:

Nie, nie mściwym, ale sponiewierania, ale swojej utoczonej krwi katolickiej – nie daruję. WRZO1957pl

Ne, netoužím, ale zneuctění a prolitou katolickou krev - jim neodpustím. WRZZ1967cz

Szerszy kontekst jednak ukazuje, iż segment ten to odpowiedź na pytanie i jest on elipsą:

Takiś to mściwy? - Nie, nie mściwym, ale sponiewierania, ale swojej utoczonej krwi katolickiej - nie daruję. WRZO1957pl

To tolik toužíš po pomstě? Ne, ***netoužím****, ale zneuctění a prolitou katolickou* krev *- jim neodpustím.* WRZZ1967cz

toužit inf

Tabela 64. Polskie jednostki tłumaczone na język czeski jako *toužit inf*.

toužit inf	**87**	*pragnąć inf*	41
		chcieć inf	8
		pragnienie Oabstr	3
		tęsknić do Oabstr	2
		chęć Oabstr	1
		chęć, żeby + S	1
		chętnie by + V	1
		dążyć do Oabstr	1
		marzący o Oabstr	1
		marzenie inf	1
		marzyć o Oabstr	1
		marzyć, że + S	1
		marzyć inf	1
		mieć ochotę inf	1
		odechcieć się Oabstr	1
		postanowić inf	1
		pragnący inf	1
		pragnąć Oabstr	1
		pragnienie inf	1
		próbować inf	1
		spragniony Oabstr	1
		spróbować inf	1
		tęsknić do Ohum	1
		usiłować inf	1
		wzdychać, aby + S	1
		zamarzyć inf	1
		struktura	11

W wyekscerpowanym materiale znalazło się 87 rekordów, w których występuje schemat *toužit inf* jako odpowiednich polskich fraz. Przeważająca liczba to jednostki z podgrupy „pragnąć”; jest ich w sumie 63.

Tabela 65. Podgrupa polskich fraz oryginalnych „pragnąć", których czeskim przekładem jest fraza *toužit inf*

pragnąć inf	41	**63**
chcieć inf	8	
pragnienie Oabstr	3	
chęć Oabstr	1	
chęć, żeby + S	1	
chętnie by + V	1	
dążyć do Oabstr	1	
mieć ochotę inf	1	
odechcieć się Oabstr	1	
pragnący inf	1	
pragnąć Oabstr	1	
pragnienie inf	1	
spragniony Oabstr	1	
usiłować inf	1	

***Pragnę** zrujnować mu jego dzieciństwo.* WGD1988pl
***Toužím** mu zničit jeho dětství.* WGD1994cz

*Czasami **pragnął** już nawet, po prostu, zamknąć oczy i umrzeć, i nigdy nie zmartwychwstać.* HGZW1965pl
*Někdy už dokonce **toužil** prostě zavřít oči a zemřít, a nikdy už nevstat z mrtvých.* HGZV1996cz

*Bardzo **chciał** ją objąć, ale przez przekorę nie zrobił tego.* ASMP1992pl
***Toužil** ji obejmout, ale nakonec to s lítostí přece jen neudělal.* ASMO1993cz

Warto zwrócić uwagę, iż fraza *toužit inf* najczęściej odpowiada polskim frazom *pragnąć inf* (41 przykładów).

Fraza *toužit inf* pojawia się też w kilku przypadkach jako przekład jednostki *marzyć* (3), również *marzący* (1), *marzenie* (1) i *zamarzyć* (1).

Tabela 66. Podgrupa polskich fraz oryginalnych „marzyć", których czeskim przekładem jest fraza *toužit inf*

marzący o Oabstr	1	**6**
marzenie inf	1	
marzyć o Oabstr	1	
marzyć, że + S	1	
marzyć inf	1	
zamarzyć inf	1	

Jest to aptekarz zagnany tutaj przez wojnę, ***marzący*** *o powrocie do rodziny.* CMRE1990pl
Je to lékárník, kterého sem zahnala válka a který ***touží*** *vrátit se k rodině.* CMRE1997cz

A drugim Stefanem Zweigiem ***marzy*** *pan być raczej w cudzysłowie i nie ze wszystkimi szczegółami, mam nadzieję?* JPMPS2006pl
A doufám, že druhým Stefanem Zweigem ***toužíte*** *být jaksi v uvozovkách, nikoli ve všech podrobnostech. ???* JPMPS2009cz

Wśród poświadczeń ze schematem *toužit inf* znajdują się też przykłady z jednostkami z grupy „inne" w wersjach oryginalnych: *postanowić* (1), *próbować* (1), *spróbować* (1), *wzdychać* (1):

Tabela 67. Podgrupa polskich fraz oryginalnych „inne", których czeskim przekładem jest fraza *toužit inf*

postanowić inf	1	**4**
próbować inf	1	
spróbować inf	1	
wzdychać, aby + S	1	

A poza tym ***postanowiłaś*** *skończyć z niepewnością, czy jesteś z nią na pani, czy na ty.* JCWC2001pl
A kromě toho jsi konečně ***toužila*** *definitivně vyřešit otázku, jestli si spolu vykáte, nebo tykáte.* JCVC2005cz

Kiedy zamykał za sobą drzwi konfesjonału, przez chwilę miał bezsensowne uczucie, że ***próbuje*** *zasiąść we wnętrzu swojego zegara.* ASOG2006pl
Když za sebou zavíral dveře zpovědnice, měl chvíli nesmyslný pocit, že ***touží*** *usednout do nitra svých hodin.* ASHP2001cz

Tam właśnie, w czasie dłużącego się nabożeństwa, odczuwaliśmy wspólnotę losu, więź, która jest nam potrzebna, choć wielokroć ***wzdychamy****, aby się jej pozbyć* (…) WTCR1981pl
Právě tam, po celou pomalu probíhající mši, jsme pociťovali společenství osudu, pouto, které je nám potřebné, i když se ho mnohokrát toužíme zbavit (…) WTCR1981cz

Najmniej liczną podgrupę stanowią rekordy z jednostką *tęsknić* w polskiej wersji oryginalnej; mowa o trzech przykładach.

Tabela 68. Podgrupa polskich fraz oryginalnych „tęsknić", których czeskim przekładem jest fraza *toužit inf*

tęsknić do Oabstr	2	**3**
tęsknić do Ohum	1	

Samotność Tomaszowi ciążyła, jednak do czego ***tęsknił****, było roztopieniem się i rozmową bez słów.* CMDI1981pl
Osamělost Tomáše tížila, a přece ***toužil*** *se rozplynout a rozmlouvat beze slov.* CMUI1993cz

Czuję się samotny i ***tęsknię*** *do Boga, chociaż go nie rozumiem.* OTB2007pl
Cítím se opuštěný a ***toužím*** *poznat Boha, protože mu nerozumím.* OTPAJC1999cz

toužit po tom inf

W wyszukiwaniach pojawiły się też dwa poświadczenia z konstrukcją obcą językowi polskiemu *toužit po tom inf*.

Tabela 69. Polskie jednostki tłumaczone na język czeski jako *toužit po tom inf*

toužit po tom inf	**2**	*chcieć inf*	1
		tęsknić za to, żeby + S	1

Nie bardzo ***chciałabym*** *zabierać nowe znajomości. WRZO1957pl*
Netoužím *po tom* **uzavírat** *nové známosti. WRZZ1967cz*

Poświadczenia, w których w wersji czeskiej pojawia się jednostka *toužit* łącząca się z frazą zdaniową, to jedynie 10 rekordów. Polskie frazy oryginalne, przekładane na język czeski reprezentują 9 różnych struktur (dwa razy pojawia się *pragnąć, aby + S*).

toužit, aby + S

Tabela 70. Polskie jednostki tłumaczone na język czeski jako *toužit, aby + S*

toužit, aby + S	**2**	*marzyć, aby + S*	1
		pragnąć, aby + S	1

W zgromadzonym materiale struktura *toužit, aby + S* jest odpowiednikiem dwóch fraz w wersji oryginalnej:

Leżąc przy nim w łóżku ***marzył*** *wstydliwie, że Celestyn jest kobietą.* OTDD1998pl
Když vedle něho ležel, ***toužil****, aby byl Celestýn žena.* OTDD2002cz

(…) *z całej duszy pragnął, by odwiedziny trwały jak najdłużej, z drugiej strony z całej duszy* ***pragnął****, by odwiedziny jak najprędzej się skończyły.* JPPMA2007pl
(…) *z celé duše si přál, aby jejich návštěva trvala co nejdéle, na druhou stranu však z celé duše* ***toužil****, aby návštěvy co nejdřív skončily.* JPUSA2007cz

toužit po tom, aby + S

Tabela 71. Polskie jednostki tłumaczone na język czeski jako *toužit po tom, aby + S*

toužit po tom, aby + S	**8**	*chcieć inf*	1
		marzyć o tym, aby + S	1
		nie chciało się Oabstr	1
		pragnąć inf	1
		pragnąć Oabstr	1
		pragnąć, aby + S	1
		pragnienie, aby + S	1
		tęsknić za tym, żeby + S	1

W tej grupie najwięcej przypadków dotyczy podgrupy „pragnąć".

Tabela 72. Podgrupa polskich fraz oryginalnych „pragnąć", których czeskim przekładem jest fraza *toužit po tom, aby + S*

chcieć inf	1	6
nie chciało się Oabstr	1	
pragnąć inf	1	
pragnąć Oabstr	1	
pragnąć, aby + S	1	
pragnienie, aby + S	1	

***Pragnienie** zaś duszy mojej takie: aby coś się stało.* WGTA2000pl
*Má duše **touží** po tom , aby se něco stalo.* WGTA2007cz

*Całe życie **chciała** studiować historię, ale dopiero na rencie znalazła na to czas.* MŁR2007pl
*Celý život **toužila** po tom, aby mohla studovat historii, teprve v důchodu si však udělala dostatek času.* MŁR2008cz

***Pragnie** jedynie, by wszystko zostało skrupulatnie odnotowane i zapłacone.* ASOG2006pl
***Touží** jen po tom, aby bylo všechno svědomitě zapsáno a zaplaceno.* ASHP2001cz

W materiale znalazł się też jeden rekord z jednostką *marzyć* i jeden z jednostką *tęsknić*:

Tabela 73. Podgrupa polskich fraz oryginalnych „marzyć", których czeskim przekładem jest fraza *toužit po tom, aby + S*

marzyć o tym, aby + S	1

***Marzą** o tym, żeby nic się nam nie udało (…)* SLGP1984pl
*Oni **touží** po tom, aby se nám nic nepovedlo (…)* SLPH1981cz

Tabela 74. Podgrupa polskich fraz oryginalnych „tęsknić", których czeskim przekładem jest fraza *toužit po tom, aby + S*

tęsknić za tym, żeby + S	1

*Któż nie **tęskni** za tym, żeby być dotkniętym, żeby być muśniętym przez Niewidzialną Rękę?* ESS1971pl
*Kdopak by **netoužil** po tom, aby se ho dotkla, aby o něho zavadila Neviditelná Ruka?* ESS1980cz

Wyniki ekscerpcji manualnej

Podwójna ekscerpcja pozwoliła na zgromadzenie par ekwiwalentów wybranych przez tłumaczy. Znaleziono ponad 60 różnych fraz polskich przekładanych na czeski jako *toužit* oraz 47 różnych fraz – ekwiwalentów czeskiej oryginalnej jednostki *toužit*.

Analiza przeprowadzona w tym rozdziale ukazuje, iż wpływ walencji na przekład czasownika *toužit* widoczny jest najbardziej w przypadku łączliwości analizowanego czasownika z bezokolicznikiem; w tekście polskim najczęściej pojawia się czasownik *pragnąć* (także *chcieć*) łączący się również z bezokolicznikiem: *toužit inf → pragnąć inf.*

Ekwiwalencję tych fraz potwierdza również badanie polskich oryginałów i przekładów na język czeski. Tylko sporadycznie pojawiają się w tym kontekście pozostałe odpowiedniki proponowane w słowniku tradycyjnym oraz pojedyncze inne odpowiedniki.

W przypadku występowania czasownika *toužit* w innych strukturach (z pozostałymi obiektami) trudno stwierdzić wpływ walencji na rozumienie i jego przekład, choć należy zaznaczyć, iż czasownik *pragnąć* oraz jednostki wobec niego bliskoznaczne pojawiają się jako ekwiwalent zawsze na szczycie listy potencjalnych odpowiedników. W wielu przypadkach, jak wynika z tabel, nie można wnioskować o przyczynach (czy okolicznościach) pojawiania się danych ekwiwalentów czeskiej jednostki, ponieważ liczby poświadczeń są zbyt niskie i trudno wskazać jakiekolwiek reguły czy preferencje semantyczne.

Jak wynika z analizy, wiele wyzwań stawiają też poświadczenia ze schematem *toužit po Oabstr* i *toužit po Ohum*. Grupy obiektów abstrakcyjnych i ludzkich łączących się z badanym czasownikiem są niezwykle różnorodne[323], a przytaczane wcześniej przykłady wskazują, iż obiekty te należy analizować dokładniej i badać ich łączliwość oraz funkcje. W takim wypadku dobrym rozwiązaniem wydaje się wykorzystanie gramatyki przypadków głębokich (Fillmore, 1968).

Przypadki głębokie

Algorytm zakłada wykorzystanie teorii przypadków głębokich i jak zostało wspomniane, poszukiwania ekwiwalentu może ułatwić identyfikacja obiektu o tej samej wartości, łączącego się zarówno z badanym czasownikiem czeskim, jak i możliwym ekwiwalentem w języku polskim.

Partycypantem pojawiającym się w każdym z analizowanych przypadków jest argument o wartości Experiencera (w przykładach poniżej podkreślony prostą linią). Drugim argumentem łączącym się z czasownikiem *toužit* jest źródło (bodziec) – Stimulus (w przykładach zaznaczony linią falowaną). Ów bodziec nie musi się znajdować w bezpośrednim sąsiedztwie czasownika czy Experiencera; może go ujawnić szerszy kontekst (jak w ostatnim przykładzie).

323 Różnorodność dotyczy też obiektów realnych (OR); ich poświadczeń jednak było znacznie mniej.

Stál jsem i nyní stále kus od ní, kdežto ona naopak toužila po rychlém příchodu teplých doteků, které by přikryly tělo vystavené chladnosti pohledu. MKZ1991cz

I teraz stałem nieco z dala od niej, podczas gdy ona, przeciwnie niż ja, tęskniła za szybkim dotknięciem ciepłych ramion, które osłoniłyby jej ciało wystawione na chłód spojrzeń. MKZ1999pl

Lucie mi za moje dopisy jen plaše děkovala a brzy začala toužit po tom, aby mi je něčím oplácela. MKZ1991cz

Za moje listy Łucja dziękowała tylko nieśmiało i wkrótce zapragnęła jakoś mi się zrewanżować. MKZ1999pl

Podařilo se učinit infantu vzácnou; podmínky císaři byly kruté, ale čím byly krutější, tím více prý toužil mladý král. JDZ1959pl

Warunki postawione cesarzowi były okrutne, ale im okrutniejsze, tym mocniej podobno młody król pragnął infantki. JDB1993cz

Przeanalizowano 272 pary segmentów z analizowanym czasownikiem *toužit* i jego potencjalnymi ekwiwalentami. Jak przypuszczano, w przypadku tego typu czasowników identyfikacja ról semantycznych nie ułatwia ustalenia trafnego ekwiwalentu.

Rodzaj obiektu i typ roli semantycznej a proces wyboru ekwiwalentu

Jak zostało już stwierdzone, analiza czasowników odnoszących się do uczuć i emocji oznacza wystąpienie semantycznej roli Experiencera; tak jest określona osoba przeżywająca dany stan. Obiekty, które ten stan powodują, mają przypisaną rolę Stimulus (bodziec, impuls). W poniższej tabeli przedstawione są prototypowe obiekty poświadczone w korpusie, które mogą pełnić funkcje Experiencera i Stimulusa.

Tabela 75. Semantyczne role argumentów czasownika *toužit* i przykłady ich leksykalnej realizacji

NP_{nom} – *toužit po* – NP_{loc}			
Experiencer	Stimulus		
	Oabstr	Ohum	OR
člověk	***láska***	*ona*	*divadlo*
národ	*pomsta*	*kluk*	*byt*
Evropská unie	*vysvětlení*	*dítě*	***domov***
stromy	*krása*	***tělo milence***	*snídaně*

Specyfikacja semantycznych ról przy wskazaniu najbliższego ekwiwalentu niestety nie pomaga, ponieważ polskie ekwiwalenty mają tę samą kombinację ról; paradoks polega na tym, że w danej grupie czasowników pojawiać się będą te same zestawy ról i nie będą one na ogół wskaźnikiem różnicowania znaczenia[324]. Analiza powierzchniowej realizacji roli Stimulus prowadzi do podobnych wniosków, jakie uzyskano podczas badania struktury walencyjnej. W obu typach analizy często trudno zdecydować, z jakim typem obiektu mamy do czynienia. Np. *láska* (‘miłość‘) może wyrażać – często w zależności od kontekstu – zarówno obiekt abstrakcyjny (Oabstr), jak i istotę ludzką (Ohum)[325]. Podobnie *domov* (‘dom‘) – Oabstr / OR. Trudność sprawia też klasyfikacja obiektów typu *tělo milence* (‘ciało kochanka‘) – czy jest to obiekt ludzki, czy obiekt realny. Takie przykłady stojące na granicy dwóch klas można by analizować odrębnie, jako przechodnią grupę obiektów; należałby do niej również obiekt z tabelki – *Evropská unie*. Chociaż niemożliwość automatycznej identyfikacji typu obiektu stanowi kolejną przeszkodę, nie niweczy jednak zupełnie badania łączliwości (Kaczmarska i Rosen 2016: 329–330).

Przeprowadzona została analiza hipotetycznych przykładów mająca na celu próbę ustalenia, z jakiego typu abstrakcyjnymi obiektami łączą się dane czasowniki *pragnąć, marzyć, tęsknić*. Wzięto pod uwagę dwa obiekty – *wielka miłość* / *velká láska*[326] i *egzotyczna podróż* / *exotická cesta* i sprawdzono możliwość ich łączliwości z wybranymi czasownikami[327]. Znak zapytania wskazuje na wątpliwą łączliwość.

Marzyć o wielkiej miłości / egzotycznej podróży.
Tęsknić za wielką miłością / egzotyczną podróżą (?)
Tęsknić do wielkiej miłości / egzotycznej podróży (?)
Pragnąć wielkiej miłości / egzotycznej podróży (?)[328]

Z analizy wynika, iż czeski czasownik dopuszcza połączenia z oboma obiektami (*toužit po velké lásce / po exotické cestě*). Możliwe jest także połączenie obu tych obiektów z czasownikiem *marzyć*. Niepoprawne jest ich użycie z czasownikiem *tęsknić*; w przypadku obiektu *wielka miłość* ta niepoprawność nie jest już taka oczywista, ponieważ możemy pod nim wiedzieć osobę – kogoś ukochanego

[324] O wyjątkach mowa była w rozdziale 11 prezentującym algorytm.

[325] Kombinacja OR i Ohum – w tabeli czcionka pogrubiona. Kombinacja OR i Oabstr – w tabeli podkreślenie i pogrubienie. Kombinacja Ohum i Oabstr – w tabeli pogrubienie i podkreślenie falą.

[326] Obiekt *wielka miłość* / *velká láska* może by interpretowany jako należący do klasy przejściowej (Oabstr / Ohum).

[327] Por. wcześniejsze badania (Kaczmarska 2014A: 196–198; Kaczmarska 2014B: 51–53; Kaczmarska, Rosen, Hana i Hladká 2015: 162–163; Kaczmarska i Rosen 2016: 328–331).

[328] Pomysł głębszej analizy łączliwości obiektów abstrakcyjnych poddała dr hab. Elżbieta Hajnicz (IPI PAN).

i wówczas zdania mogą być uznane za poprawne. Czasownika *pragnąć* raczej nie można kombinować z obiektem *egzotyczna podróż*.

Podsumowując badania dotyczące łączliwości analizowanego czasownika z różnymi obiektami (*toužit po Oabstr*, *Ohum*, *OR*), warto zauważyć, iż przykłady wskazują na to, że rolę przy wyborze ekwiwalentu grać mogą różne faktory[329], np. istnienie „obiektu" w momencie nadania komunikatu, w przeszłości bądź w przyszłości, ewentualnie fakt, że podmiot miał już kontakt z obiektem. Można zakładać, iż czasownik *marzyć* łączy się z obiektem, który do podmiotu nie należy (i nigdy nie należał), bądź też nigdy nie miał do dyspozycji, czy wręcz nie miał z obiektem kontaktu.

Marzę o małym domku nad oceanem.
Marzę o pierścionku z szafirem.
Marzę o dużej rodzinie.

Można by również wnioskować o czasowniku *tęsknić*, iż pojawi się w sytuacji, gdy wystąpi obiekt co prawda znany, ale przestrzennie odległy.

Tęsknię za moim synem.
Tęsknię za swoim domem.
Tęsknię za spokojem.

Przykłady z korpusu nie potwierdzają jednak tak prostych zależności, np. czasownik *marzyć* łączy się z obiektem, z którym przeżywający już na pewno miał kontakt[330].

Marzę o szklance wody.

Kolejnym dowodem na złożoność tych relacji jest zachowanie czasownika *pragnąć*, który może się łączyć z obiektami różnych typów:

Pragnę dziecka.
Pragnę go z całego serca.
Pragnął łatwej wygranej.
Pragnęła spokoju.

Właśnie z punktu widzenia łączliwości tłumacze mogliby uważać ten czasownik za uniwersalny i przy przekładzie go preferować. Taki wybór jest jednak często

[329] Zob. preferencje semantyczne w algorytmie ustalania podobieństwa pomiędzy jednostkami pochodzącymi z różnych języków – *congruence* / *pattern* / *collocation* / *colligation* / *semantic preference* / *semantic prosody* (Ebeling i Ebeling 2013: 80); por. model Sinclaira (Sinclair 1996).

[330] Analizując zdanie *Marzę o wycieczce do Lizbony.* stwierdzimy raczej, iż chodzi o pierwszą wizytę w Lizbonie; możliwe jest jednak rozumienie, iż byłby to kolejny wyjazd do tego miasta: *Marzę o (ponownej) wycieczce do Lizbony*. Tu dodatkowe znaczenie (lub konkretyzację) może wnieść dalsze otoczenie.

związany z przesunięciem znaczeniowym. Mogłoby się też zdarzyć, iż powstaną konkstrukcje wątpliwe, np.:

**Pragnęła książki z obrazkami.*

To oznacza, iż każdy przekład z udziałem czasownika *toužit* obarczony być może uproszczeniem i stratą części znaczenia, a także, że potrzebna jest dokładniejsza – szersza – analiza semantyczno-syntaktyczna obejmująca tego typu obiekty, prowadzona na dużo bogatszym materiale językowym, lub analiza kontekstu. Stąd pomysł wykorzystania podejścia Pattern Grammar oraz metod stochastycznych.

Pattern Grammar

Materiał do tego badania wyekscerpowano z wersji 11 InterCorpu: czesko-polska część – czeskie oryginały z trzonu wiązane z polskimi tłumaczeniami (Rosen, Vavřín i Zasina 2018). Uzyskano 278 równoległych poświadczeń, dla których wybrano manualną formę analizy[331]. W poświadczeniach tych nie znaleziono żadnego wzorca, który przez swoją powtarzalność mógłby wskazywać na konkretne znaczenie i co za tym idzie na konkretny ekwiwalent w języku polskim. Badanie powtórzono więc na wszystkich tekstach czeskich (bez rozróżnienia na oryginał i na przekład) z wersji 11 InterCorpu, a KonText wyszukał 4244 segmenty równoległe. Automatycznie zostały wyszukane powtarzające się kolokacje; niestety ponownie, mimo ręcznej analizy wszystkich segmentów, nie udało się wyodrębnić żadnego wzorca, który mógłby mieć konkretny odpowiednik w języku polskim. Uwagę autorki zwróciły jednak dwie powtarzające się częściej frazy: *toužit po lásce* (22 poświadczenia) oraz *toužit mít* (84 poświadczeń). Wzorzec *toužit po lasce* nie ma konkretnego ekwiwalentu w języku polskim, co obrazuje też poniższa ilustracja; być może jest po prostu dość częstą kolokacją.

[331] Przy tak ograniczonej liczbie równoległych poświadczeń nie wykorzystano typowych n-gramów; analiza nie wychodziła jednak ponad pięć elementów kontekstu prawo- i lewostronnego. Por. badania na n-gramach (Chlumská 2017).

Monica	***toužila***	***po lásce**, a proto úmzslně bez mého vědomí otěhotněla.*
Monica chciała być kochana, oszukała mnie więc i zaszła w ciążę.		
Takže už jsem v dobré kondici, výborně oblečený a	***toužím***	***po lásce** a sexu.*
Byłem wtedy w dobrej formie, dobrze ubrany, spragniony miłości i seksu.		
To bylo gesto sentimentálního buržoz, který	***toužil***	***po lásce** prostých lidí.*
Był to gest wrażliwego bourgeois, dbającego o to, by ludzie prości kochali go.		
Já	***toužím***	***po lásce**.*
Ja tęsknię do miłości.		
stejně jako	***toužila***	***po lásce**, kterou se v této chvíli bála před ním prozradit.*
jak łaknęła miłości, z którą się bała zdradzać przed nim w tej chwili, powstrzymana jedynie jakimś odruchem kobiecej wstydliwej bierności.		
také	***toužím***	***po lásce!***
i ja pragnę miłości!		

Obraz 29. Wzorzec *toužit po lásce* wraz z odpowiednikami w języku polskim

Natomiast w przypadku frazy *toužit mít* zauważalna jest powtarzalność omijania przy przekładzie czasownika *mít*, jak w przypadku

Toužil mít prázdniny. MKNLB1985cz
Pragnął wakacji. MKNLB1992pl

W przypadku tego czasownika podejście nurtu Pattern Grammar nie przyniosło oczekiwanego rezultatu; niestety rodzi to przypuszczenie, iż niesatysfakcjonujące dla tej jednostki wyniki przyniosą również inne badania opierające się na kontekście liniowym[332].

[332] Część czasowników wyrażających stany emocjonalne można jednak z powodzeniem analizować metodą Pattern Grammar, np. wspomniana wcześniej fraza *být líto*, czy wzorzec *cítit s* wyodrębniona przy analizie czasownika *cítit*; jednostka *cítit s* ma w języku polskim znaczenie 'współczuć', 'być przykro' (szczególnie łatwe do ustalenia, kiedy badany czasownik jest w 1. osobie liczby pojedynczej):
***Cítím s** tebou, že musíš zpívat na takových místech.*
Współczuję ci, że musisz śpiewać w takich miejscach.
*Vernone, víš asi, jak **cítím s** Gwen.*
Vernon, wiesz, jak mi przykro z powodu Gwen.

Inne metody

Analiza stochastyczna

W celu ustalenia trafnych ekwiwalentów czasownika *toužit* wykorzystane zostały również metody stochastyczne[333]. W oparciu o teksty równoległe zgromadzone w korpusie InterCorp, dzięki użyciu programu GIZA++ zidentyfikowanych zostało 10 milionów par wyrównanych „słowo do słowa" (*word-to-word*). Badanie zakładało wykorzystanie dwóch typów kontekstu w języku oryginału (tu – w języku czeskim): linearnego (kilka słów otaczających dany leksem) oraz syntaktycznego (relacje zależnościowe). Podobnie jak w poprzednich badaniach wykorzystywane były lematyzowane teksty w celu ominięcia problemu rozproszenia danych[334]. Użyta została dziesięciokrotna walidacja, żeby ocenić wybrane podejście (Kondas 2016B). Żadne dane nie były korygowane ręcznie; analiza opierała się na wiązaniu „słowo do słowa" (*word-to-word*), co mogło spowodować pewien szum[335].

Jak wspomniano powyżej, wykorzystywane były *cechy* (*features*) w badaniu kontekstu liniowego i składniowego.

Cechą nazywana jest własność kontekstu – tu dana struktura z analizowanym czasownikiem, przy czym struktura, w której pojawia się polski odpowiednik, nie jest badana; np. różnymi *cechami* są:

Tabela 76. Wybrane *cechy – features* wraz z przykładowymi polskimi ekwiwalentami.

(1)	_ _ *vřele toužím* _ _	*pragnąć*
(2)	_ _ *vřele toužím* _ _	*marzyć*
(3)	_ _ _ *toužit* _ _ *odnedávna*	*tęsknić*
(4)	_ _ _ *toužit šileně* _ _	*pragnąć*

W obu przypadkach – zarówno analizy z kontekstem liniowym, jak i składniowym – wykorzystane zostały funkcje binarne, które umożliwiają stwierdzenie, czy element taki jak np. kategoria przysłówka lub konkretny leksem występuje w pobliżu predykatu czy nie. Liczba wystąpień konkretnego elementu w ramach jednego kontekstu była ignorowana. Na przykład nie miało znaczenia, czy predykat występuje z jednym czy dwoma przysłówkami. Tak więc wynik jest

[333] Badania stochastyczne przeprowadzone dzięki udziałowi dr. inż. Alexandra Rosena, dr. Jirki Hany i dr Barbory Hladkiej.

[334] Dane wykorzystywane do tego badania pochodzą z czesko-polskiej części trzonu InterCorpu (wersja 7), zawierającej 18 milionów czeskich tokenów.

[335] Szczegółowy opis tego podejścia – (Kaczmarska, Rosen, Hana i Hladká 2015: 165–171).

binarnym wektorem reprezentującym obserwowane *cechy* dla każdego wystąpienia analizowanego predykatu. Odległość i kierunek od predykatu były ignorowane. Jest to motywowane faktem, że w języku czeskim szyk jest determinowany bardziej pragmatycznie niż przez ograniczenia semantyczne czy składniowe, które mają z kolei większy wpływ na odpowiednie tłumaczenie. Ostatecznie brane były pod uwagę tylko te *cechy*, które wystąpiły w co najmniej 5 przypadkach. Dodatkowo – jeśli okazało się, że dana struktura odpowiadała wszystkim branym pod uwagę ekwiwalentom (tu – *pragnąć, marzyć, tęsknić*), *cecha* ta była usuwana jako nieróżnicująca.

Celem analizy było ustalenie, które *cechy* są relewantne i najbardziej decydują o wyborze ekwiwalentu.

Kontekst linearny

W tej metodzie wykorzystany został kontekst pięciu słów poprzedzających obserwowany leksem i pięciu słów następujących po nim, o ile słowa pojawiły się w tym samym zdaniu. W przeciwnym razie liczba słów w kontekście była ograniczona granicą zdania. Zatem dla czasownika *toužit* w zdaniu *Odedávna jsem toužila po komické roli. cechy* bazowały na wszystkich słowach w zdaniu.

Ten eksperyment niestety się nie udał; żaden kontekst nie wskazał jednego konkretnego ekwiwalentu[336].

Kontekst zależności

Zamiast kontekstu pięciu słów przed obserwowanym czasownikiem i po nim, tutaj wykorzystane zostały elementy zdania, które łączą się z danym czasownikiem.

W tej metodzie czeskie teksty parsowane były za pomocą stochastycznego parsera zależności, trenowanego na danych pochodzących z Prague Dependency Treebank (Nivre i Hall 2005; Jelínek 2014). Analiza składniowa miała ukazać, z jakimi elementami na zasadzie zależności łączy się analizowany czasownik, przy czym została zmodyfikowana tak, aby uwzględnić tylko rzeczywiste zależności i pominąć (czy raczej „obejść") słowa pomocnicze, jak przyimki czy spójniki, które mogłyby się znaleźć pomiędzy badanym czasownikiem a zależnym od niego obiektem (Štěpánek 2006). W przypadku przytaczanego wcześniej

[336] Dzięki tej metodzie najlepsze wyniki osiągnięte zostały w przypadku jednostek *být líto* i *trápit*; analiza tych czasowników przebiegała równolegle z czasownikiem *toužit*.

przykładu (*Odedávna jsem toužila po komické roli*.) „ominięty" zostaje przy analizie zależności przyimek „po".

Jak zostało to wcześniej stwierdzone, wygenerowane *cechy* reprezentują typ boolowski (ang. Boolean), czyli są „prawdziwe" albo 'fałszywe" w zależności od obecności lub braku konkretnych leksemów (lub ich kombinacji) we frazie odpowiadającej konkretnemu ekwiwalentowi (Kruszewski, Paperno i Baroni 2015). W każdym zdaniu z analizowanym czasownikiem sprawdzany jest bowiem odrębnie każdy element oraz kombinacja tych elementów[337].

Kolejnym etapem było przeprowadzenie procedury opisanej w artykule *Feature engineering in the NLI Shared Task 2013: Charles University submission report* (Hladká, Holub i Kríž 2013). *Cechy* przefiltrowane były z wykorzystaniem wartości zysku informacyjnego (Information Gain – IG). Wartość IG danej *cechy* wyraża, jak obniża się entropia (wątpliwość) związana z wyborem ekwiwalentu w momencie, gdy daną *cechę* znamy. Zatem IG *cechy* X wynosi:

$$IG(X) = H(Y) - H(Y|X)$$

gdzie H(Y) jest entropią zmiennej zależnej[338], a H(Y|X) jest entropią zmiennej zależnej ze znaną *cechą* X. Kluczowe jest, by parametr określający zysk informacyjny jakiejś danej *cechy* X [IG(X)] był najwyższy; stanie się to wówczas, kiedy H(Y|X) będzie równy zeru; jest to możliwe w sytuacji, gdy Y będzie całkowicie zdeterminowany przez X (czyli *cechę*). W poniższej tabeli przedstawiono najczęstsze *cechy* z najwyższym współczynnikiem IG, przy czym warto zaznaczyć, iż całkowita entropia zmiennej zależnej wynosi 2,45266. Rezultat ukazuje, iż do wyraźnego obniżenia entropii dochodzi przy kombinacji funkcji AuxP z przyimkiem *po* (0.095)[339].

[337] Metoda ta ma na celu znalezienie najważniejszych *cech*, czyli tych, które najlepiej rozróżniają kontekst z punktu widzenia wyboru ekwiwalentu. Wcześniej jednak nie jest wskazane, który element może być najważniejszy (np. podmiot w zdaniu). Program sam to ustala, co odpowiada uczeniu nienadzorowanemu – ang. *unsupervised machine learning*. Szerzej na temat różnych metod uczenia – (Brownlee 2016; Kondas 2016A).

[338] Oparta na wiedzy *a priori*, np. zakładając, że wiemy, ile jest ekwiwalentów i jaki jest ich rozkład procentowy; bez określenia *cechy*.

[339] https://ufal.mff.cuni.cz/pdt2.0/doc/manuals/cz/a-layer/html/ch03s03s07.html – objaśnienia wszystkich słów pomocniczych typu AUX.

Tabela 77. Lista *cech*, które najbardziej wpływają na wybór ekwiwalentu

Cecha	**Opis *cech***	**IG**
AuxP + *po*	leksem *po* w fumkcji przyimka	0,09482
po	leksem *po*	0,09446
AuxP	funkcja przyimka	0,07886
Obj	funkcja obiektu	0,05061
Adv	funkcja okolicznika	0,03
žena	leksem *žena*	0,0267
on	leksem *on*	0,02089
,	przecinek jako znak interpunkcyjny	0,01872
Obj + *být*	leksem *být* w funkcji obiektu	0,01564
Adv + *láska*	leksem *láska* w funkcji okolicznika	0,01471
nic	leksem *nic*	0,01359
život	leksem *život*	0,01341

Otrzymany zbiór *cech* był następnie użyty do trenowania klasyfikatora. Wypróbowane zostały maszyny wektorów nośnych[340], drzewa decyzyjne[341], losowe lasy[342], naiwny klasyfikator bayesowski[343]. Żaden z tych klasyfikatorów nie wygenerował lepszych wyników od prostego modelu, który zawsze wybierał najczęstszy polski ekwiwalent bez względu na kontekst. Nie udało się więc eksperymentalnie potwierdzić hipotezy, że wybór ekwiwalentu jest uzależniony od kontekstu czeskiego leksemu. Żadna z *cech* oparta na kontekście nie niesie w sobie bowiem wystarczająco dużo informacji dotyczących optymalnego wyboru.

[340] Maszyna wektorów nośnych (maszyna wektorów podpierających) – SVM ang. *support vector machine* – metoda machine learning, służąca przede wszystkim do klasyfikacji.

[341] Drzewa decyzyjne – czes. *rozhodovací stromy*, ang. *decision trees* – w dużym uproszczeniu – klasyfikator – technika pozyskiwania informacji z danych.

[342] Lasy losowe – czes. *náhodné lesy*, ang. *random forests* – algorytm zbudowany z wielu klasyfikatorów (drzew decyzyjnych – stąd nazwa „las"); każde drzewo budowane jest na podstawie innego, losowo wybranego podzbioru zbioru danych uczących. Każdy podział w drzewie jest wybierany jako najlepszy możliwy podział dla losowo wybranego małego podzbioru zmiennych – stąd nazwa „losowy" (Demski 2011).

[343] Naiwny klasyfikator bayesowski – czes. *naivní Bayesův klasifikátor*, ang. *naive Bayes classifiers* – model cech niezależnych, prosty klasyfikator probabilistyczny oparty na założeniu o wzajemnej niezależności zmiennych niezależnych (StatSoft 2006).

Word Sketch

Dane o typowym wypełnieniu pozycji walencyjnych konkretnymi leksemami można uzyskać dzięki narzędziu Sketch Engine, czyli tzw. profilom kolokacyjnym na podstawie morfologicznie lub syntaktycznie anotowanych tekstów (Word Sketch)[344]. Narzędzie Sketch Engine ukazuje kolokacyjny potencjał danego leksemu za pomocą automatycznie wygenerowanych list jego kolokatów (obiektów, z którymi się łączy), pogrupowanych według ich gramatycznych funkcji.

Wykorzystanie tego narzędzia umożliwia dokonanie dalszej analizy czeskiego czasownika *toužit*. Badanie zostało przeprowadzone z oparciu o czesko-polską część korpusu równoległego InterCorp. Pierwszym etapem była próba potwierdzenia tezy o ekwiwalencji *toužit* + *inf* i *pragnąć* + *inf*. Wyszukiwanym czasownikiem był czeski leksem *toužit*, a w języku polskim *pragnąć*, *marzyć* i *tęsknić*. Należało ustalić, który z ekwiwalentów łączy się z bezokolicznikiem; materiał korpusowy ukazał, iż polskie czasowniki *marzyć* i *tęsknić* nie wykazują takiej łączliwości. W tym momencie pojawił się kolejny problem: rezultaty badań dla języka czeskiego i polskiego są nieporównywalne.[345] Czeskie przykłady są ekscerpowane z olbrzymiego Czeskiego Korpusu Narodowego, a polskie przykłady – z korpusu równoległego InterCorp; dane są nieporównywalne pod względem wielkości. Obecnie wykorzystanie narzędzi Sketch Engine dla języka czeskiego jest niemożliwe z użyciem InterCorpu. Nie jest możliwe również wykorzystanie Narodowego Korpusu Języka Polskiego (NKJP); jego wielkość jest porównywalna z wielkością czeskiego korpusu, ale NKJP nie oferuje generowania profili typu Word Sketch. W czeskiej części InterCorpu profile kolokacyjne zostały więc zastąpione frekwencyjnymi statystykami konkordancji na podstawie zapytań równoważnych wyrażeniom regularnym w gramatyce profili kolokacyjnych[346].

[344] Zob. https://www.sketchengine.eu/user-guide/user-manual/word-sketch/ (Kilgarriff i inni 2014), (Benko 2014) (Didakowski i Geyken 2013). Również bez syntaktycznej czy morfologicznej anotacji można z korpusu ekscerpować statystycznie relewantne kolokacje, np. https://kontext.korpus.cz, http://nkjp.uni.lodz.pl/collocations.jsp i (Pęzik 2012).

[345] Innym problemem są też różnice w rodzajach tekstów zamieszczonych w tych korpusach – Czeskim Korpusie Narodowym (CNK) i czesko-polskiej części InterCorpu.

[346] Inną możliwością jest zestawienie wyników z CNK (*syn*) z porównywalnym polskim korpusem – Narodowym Korpusem Języka Polskiego (NKJP). Wyniki uzyskane dzięki narzędziu Kolokator, które jest najbliższe idei profili kolokacyjnych, niestety nie są porównywalne z rezultatami analiz z CNK ze względu na różne miary statystyczne i odmienną koncepcję. Jednak chociaż rezultatów nie można porównywać co do liczb, warto przeprowadzić to zestawienie i traktując eksperymentalnie, porównać wyniki jakościowe.

Tabela 78. Wyniki zastosowania narzędzia Sketch Engine do analizy połączenia *toužit + inf* i jego polskich ekwiwalentów

***toužit* + inf**						***pragnąć* + inf**		
syn			czesko-polska część *InterCorpu*					
			czeska część			polska część		
celkem	17 405	100,00%	**celkem**	336	100,00%	**celkem**	6 800	100,00%
mít	926	5,32%	*být*	39	11,61%	*podziękować*	805	11,84%
stát	864	4,96%	***mít***	12	3,57%	*podkreślić*	598	8,79%
poznat	382	2,19%	***poznat***	10	2,98%	*pogratulować*	379	5,57%
vidět	346	1,99%	*stát*	9	2,68%	*wyrazić*	391	5,75%
vrátit	333	1,91%	***jít***	8	2,38%	*zwrócić*	319	4,69%
hrát	333	1,91%	*slyšet*	7	2,08%	*przypomnieć*	165	2,43%
získat	332	1,91%	***vidět***	6	1,79%	*powiedzieć*	386	5,68%
dostat	311	1,79%	***dostat***	6	1,79%	*zauważyć*	97	1,43%
vyhrát	285	1,64%	*zbavit*	5	1,49%	*rozpocząć*	73	1,07%
žít	177	1,02%	***vrátit***	5	1,49%	*poruszyć*	58	0,85%
jít	176	1,01%	*opustit*	5	1,49%	*powtórzyć*	51	0,75%
najít	152	0,87%	*vědět*	4	1,19%	*skorzystać*	70	1,03%
udělat	143	0,82%	***udělat***	4	1,19%	*powitać*	43	0,63%
spatřit	132	0,76%	***spatřit***	4	1,19%	*dodać*	70	1,03%
uspět	124	0,71%	*proměnit*	4	1,19%	*zaznaczyć*	40	0,59%
dělat	105	0,60%	***najít***	4	1,19%	*przyłączyć*	33	0,49%
napravit	101	0,58%	***žít***	3	0,89%	*wezwać*	39	0,57%
zůstat	99	0,57%	***zůstat***	3	0,89%	*zapytać*	43	0,63%
pracovat	96	0,55%	*zemřít*	3	0,89%	*poinformować*	46	0,68%
podívat	92	0,53%	*vzít*	3	0,89%	*pochwalić*	26	0,38%

Tabelka ukazuje łączliwość *toužit* i *pragnąć* z bezokolicznikiem w oparciu o profile kolokacyjne z korpusu *syn*[347] oraz z wszystkich tekstów z czeskiej i polskiej części wersji 6. InterCorpu[348]. Mimo nierównowagi kwantytatywnej i jakościowej obu źródeł wyniki potwierdzają rezultaty wcześniejszych analiz; dwa pozostałe ekwiwalenty nie odnotowały żadnych kolokacji, co oznacza, że dla czeskiego czasownika *toužit* w połączeniu z bezokolicznikiem czasowniki *marzyć*

[347] 2,7 miliardów tokenów, wersja z 24.01.2014 roku.
[348] Około 59 milionów tokenów dla części polskiej.

i *tęsknić* nie będą preferowanym ekwiwalentem. Dla czasownika *pragnąć* jest natomiast połączenie z bezokolicznikiem bardzo częste.

Obrazy Word Sketch pokazują łączliwość czasownika z konkretnym obiektem oraz umożliwiają podzielenie leksemów na klasy semantyczne, np. po przyimku *po* w czeskim jednojęzycznym korpusie syn pojawiają się m.in. wyróżnione klasy: *láska, přátelství, vztah* (miłość, przyjaźń, związek)*; sláva, bohatství, popularita* (sława, bogactwo, popularność) czy *samostatnost, nezávislost, svoboda, volnost* (autonomia, niezależność, swoboda, wolność).

Tabela 79. Wyniki zastosowania narzędzia Sketch Engine do analizy połączenia *toužit po* + Obiekt i jego polskich ekwiwalentów

toužit po		***pragnąć***		***marzyć o***		***tęsknić za***		***tęsknić do***	
post_ _po	**23752**	**has_gen_ _obj**	**809**	**verb_o_ _noun**	296	**verb_za_ _noun**	**94**	**verb_do_ _noun**	**59**
dítě	697	*co*	76	*to*	70	*dom*	6	*spokój*	3
láska	599	*to*	52	*Europa*	14	*to*	5	*dom*	3
návrat	555	*Europa*	34	*powrót*	10	*junior*	2	*świat*	3
úspěch	493	*zachęcić*	26	*demokracja*	5	*mąż*	2	*słońce*	2
vítězství	457	*strona*	16	*wolność*	4	*żona*	2	*ciało*	2
změna	455	*coś*	15	*utopia*	3	*powrót*	2	*rzecz*	2
život	361	*powód*	14	*zemsta*	3	*ojciec*	2		
medaile	316	*region*	14	*domek*	3	*coś*	2		
pomsta	287	*śmierć*	13	*kariera*	3	*człowiek*	2		
klid	282	*demokracja*	12	*miłość*	3	*czas*	2		

Wyniki w tabeli X są niejednoznaczne, dlatego postanowiono przeprowadzić eksperymentalne badanie, wykorzystując obrazy typu Word Sketch dla języka polskiego (InterCorp), kolokator[349] – również dla języka polskiego (NKJP)[350] oraz profile Word Sketch dla języka czeskiego (CNK). Wyniki tej analizy przedstawia kolejna tabela.

[349] Kolokator dostępny na stronie: http://www.nkjp.uni.lodz.pl/collocations.jsp

[350] W przypadku kolokatora wyniki zostały zliczone ręcznie, a takiej analizie został poddany tylko jeden polski ekwiwalent – *pragnąć*.

Tabela 80. Wyniki eksperymentalnego badania z użyciem kolokatora i Sketch Engine

Word Sketches (InterCorp)		kolokator (NKJP)		Word Sketches (CNK)	
pragnąć	**809**	***pragnąć* + Gen**		***toužit po***	**23752**
has_gen_obj				**post_po**	
co	76	*on*	1460	*dítě*	697
to	52	*człowiek*	163	*láska*	599
Europa	34	*ty*	143	*návrat*	555
zachęcić	26	*życie*	110	*úspěch*	493
strona	16	*coś*	107	*vítězství*	457
coś	15	*bóg*	63	*změna*	455
powód	14	*kobieta*	60	*život*	361
region	14	*dziecko*	57	*medaile*	316
śmierć	13	*świat*	50	*pomsta*	287
demokracja	12	*nic*	47	*klid*	282

Rezultaty tego badania mają znaczenie wyłącznie poznawcze, a samego testu nie można implementować do algorytmu wyszukiwania ekwiwalentów. Program Sketch Engine wydaje się obiecującym narzędziem dla analizy. Optymalną sytuacją byłoby, gdyby analizę przeprowadzano na podstawie jednego korpusu i jednobrzmiących tekstów, dlatego należy dążyć do tego, by narzędzie współpracowało zarówno z czeską, jak i polską częścią InterCorpu. Autorka oczekuje, iż będzie można analizować nie tylko obiekty wyrażone frazami nominalnymi, ale także przysłówki łączące się z danym czasownikiem. Poniższa tabelka przedstawia rozkład przysłówków w odniesieniu do analizowanych jednostek[351].

[351] Jest to jednak badanie eksperymentalne; analiza dokonana została na nieporównywalnych korpusach.

Tabela 81. Łączliwość przysłówków – rezultat badania z wykorzystaniem Sketch Engine.

toužit		***pragnąć***		***tęsknić***		***marzyć***	
hodně	1099	*bardzo*	136	*bardzo*	34	*jedynie*	6
moc	1085	*gorąco*	40	*ogromnie*	3	*często*	5
tak	783	*jedynie*	19	*niesamowicie*	2	*bardzo*	5
už	778	*jednocześnie*	16	*okropnie*	2	*próżno*	4
tolik	773	*rozpaczliwie*	15	*straszliwie*	2	*długo*	4
vždycky	751	*rzeczywiście*	13	*strasznie*	2	*dobrze*	3
stále	543	*szczerze*	12	*szczególnie*	2	*niejasno*	2
dlouho	501	*mocno*	8			*naturalnie*	2
vždy	479	*wyraźnie*	8			*nieustannie*	2
také	468	*dużo*	7			*stale*	2

Wnioski

Prowadzona analiza czasownika *toužit* była wieloaspektowa. Badano definicje tego czasownika zarówno w słownikach jednojęzycznych, jak i dwujęzycznym, a także wymagania walencyjne, kolokacje i kontekst.

Ekwiwalenty słownikowe i automatyczna ekstrakcja ekwiwalentów

Analiza definicji słownikowych i automatycznie wygenerowanych par ekwiwalentów ukazuje, iż najczęściej wybieranym ekwiwalentem jest czasownik *pragnąć*; element znaczeniowy „pragnąć" jest też częścią wspólną trzech ekwiwalentów proponowanych przez słownik czesko-polski (Siatkowski i Basaj 2002). Biorąc pod uwagę jedynie statystykę, można by uznać jednostkę *pragnąć* za optymalny ekwiwalent analizowanego czeskiego czasownika.

Wart uwagi jest również fakt, iż wysokie statystyki osiągają również wyrazy bliskoznaczne wobec czasownika *pragnąć*: szczególnie *chcieć*, *pożądać* i *mieć ochotę*, co potwierdza również ranking najczęstszych potencjalnych ekwiwalentów czasownika *toužit* w analizie manualnej i automatycznej[352].

[352] Wyekstraktowane automatycznie pary porównano z wynikami ekscerpcji ręcznej, która prowadzona była na materiale czesko-polskim (wyłącznie czeskie oryginały i ich przekład na język polski). Takie porównanie może mieć na celu tylko sprawdzenie, które ekwiwalenty się powtarzają; wartości liczbowe nie mogą być zestawiane. Analiza automatyczna (Treq) obejmowała bowiem wszystkie teksty czeskie i polskie, bez względu na język oryginału, co oznacza, że materiał ten był dużo obszerniejszy niż materiał wykorzystany do analizy ręcznej. Przedstawiona tabela ukazuje, które pozycje zajmują ekwiwalenty na obu listach.

Tabela 82. Ranking najczęstszych potencjalnych ekwiwalentów czasownika *toužit* w analizie manualnej i automatycznej

potencjalny ekwiwalent	analiza ręczna	analiza automatyczna
pragnąć	1	1
marzyć	2	4
tęsknić	3	3
chcieć	4	2
pożądać	5	5
mieć ochotę	6	6
zapragnąć	6	11
dążyć	7	17
pragnienie	7	7
zatęsknić	7	18
ciągnąć do	8	x
czekać	8	14
dbać	8	x
imponować	8	x
marzący	8	x
myśleć	8	x
pożądanie	8	x
pożądany	8	x
szukać	8	10
tęskno	8	x
tęsknota	8	13
upragniony	8	17
zachciewać się	8	x
żądać	8	x
żądny	8	19
życzyć sobie	8	9
żywić miłość	8	x
życzyć	x	9
zależeć	x	8
potrzebować	x	12
spragniony	x	14
woleć	x	15
domagać	x	16
chęć	x	16
aspirować	x	18
potrzebny	x	19
doczekać	x	20
oczekiwać	x	20
potrzeba	x	20
szaleć	x	20

Powyższa tabela porównuje wyniki ekscerpcji ręcznej (cs → pl) i automatycznej[353]. Kolumny prezentują liczby rankingowe ekwiwalentów czasownika *toužit*; przy czym znak **x** oznacza, iż w analizie manualnej dany ekwiwalent w ogóle się nie pojawił, natomiast w analizie automatycznej nie pojawił się więcej niż cztery razy[354]. Ekwiwalenty na najwyższych pozycjach powtarzają się w wynikach obu badań, a pierwsza dziesiątka ma podobne pozycje. Zarówno analiza ręczna, jak i automatyczna wskazują te same jednostki jako pierwszych sześć odpowiedników czasownika *toužit*; są to: *pragnąć*, *marzyć*, *tęsknić*, *chcieć*, *pożądać*, *mieć ochotę* (kolejność według wyników analizy manualnej) i *pragnąć*, *chcieć*, *tęsknić*, *marzyć*, *pożądać*, *mieć ochotę* (kolejność według wyników analizy automatycznej). Wyniki ekscerpcji automatycznej odpowiadają w przybliżeniu rezultatom ekscerpcji manualnej. Oznacza to, iż automatyczna metoda ekscerpcji ekwiwalentów odniosła spodziewany efekt i należy oczekiwać, iż możliwe jest przeprowadzanie kolejnych badań z wykorzystaniem tej metody[355].

Analiza walencyjna

Analiza walencyjna obejmowała siedem schematów. Wyniki badania czeskich tekstów oryginalnych w przekładzie na język polski wskazują na trafny ekwiwalent tylko jednej struktury: *toužit inf* → *pragnąć inf*. Czasownik *pragnąć* oraz jednostki wobec niego bliskoznaczne pojawiają się najczęściej na szczycie listy potencjalnych odpowiedników również w przypadku pozostałych schematów. Należy jednak przypomnieć, iż czasowniki z tej grupy mają różne wymagania składniowe: *pragnąć* i *pożądać* łączą się z obiektem w dopełniaczu, *chcieć* – w bierniku, a *mieć ochotę* przyłącza frazę przyimkową *na* + obiekt w bierniku. Wymagania składniowe w przypadku czasownika *toužit* nie miały decydującego wpływu na wybór ekwiwalentu; nie można też tu mówić o podobieństwie składniowym potencjalnych odpowiedników i jednostki czeskiej, *toužit* przyłącza bowiem frazę przyimkową *po* + obiekt w miejscowniku.

[353] Ekscerpcja automatyczna przeprowadzona została w oparciu o wszystkie teksty znajdujące się w czesko-polskiej części korpusu.

[354] Leksemy mające taką samą liczbę poświadczeń, zajmują tę samą pozycję rankingową, stąd np. przy czasownikach *doczekać* i *oczekiwać* pojawia się 20 – ta sama pozycja rankingowa.

[355] Należy podkreślić, iż wyniki są podobne, mimo tego, iż w obu metodach – manualnej i automatycznej – korzystano z nieidentycznych korpusów; do ekscerpcji automatycznej wykorzystane zostały wszystkie teksty – oryginały i przekłady również z innych języków, jednak ciągle były to teksty beletrystyczne. Dodanie tekstów obcych wskazane było ze względu na to, iż wyniki metody automatycznej są bardziej wiarygodne przy większej ilości danych.

W przypadku pozostałych analizowanych schematów trudno stwierdzić wpływ walencji na wybór ekwiwalentu. Poza czasownikiem *pragnąć* jako możliwe ekwiwalenty czasownika *toužit* pojawiają się również jednostki *marzyć* (*o* + obiekt w miejscowniku) i *tęsknić* (*za* + obiekt w narzędniku / *do* + obiekt dopełniaczu). Jednak jak wynika z badania schematu *toužit po Ohum,* w tym przypadku mało prawdopodobne byłoby pojawienie się czasownika *marzyć* w roli ekwiwalentu.

Wyniki badania polskich tekstów oryginalnych w przekładzie na czeski, choć rozszerzonego o nowe grupy (m.in. schemat *toužit X* czy frazeologizmy), w sposób jednoznaczny potwierdzają jedynie tezę o ekwiwalencji struktur *pragnąć inf* i *toužit inf.*

Przypadki głębokie

Jak wykazała analiza 272 par segmentów z analizowanym czasownikiem *toužit* i jego potencjalnymi ekwiwalentami, wykorzystanie teorii przypadków głębokich i identyfikacja ról semantycznych w przypadku tego typu czasowników nie ułatwia ustalenia trafnego ekwiwalentu. Przeprowadzony w ramach tego badania eksperyment z hipotetycznymi przykładami, analizujący preferencje semantyczne, ukazał jedynie, iż uniwersalnym ekwiwalentem czasownika *toužit* mogłaby być jednostka *pragnąć*. Takie ujednoznacznienie wiązałoby się jednak z uproszczeniem i stratą części znaczenia w procesie przekładu.

Pattern Grammar

Manualna analiza wyekscerpowanych z wersji 11 InterCorpu (czeskie oryginały z trzonu korpusu wiązane z polskimi tłumaczeniami) 278 równoległych poświadczeń czasownika *toužit* nie wykazała żadnego wzorca, który przez swoją powtarzalność mógłby wskazywać na konkretne znaczenie i co za tym idzie na konkretny ekwiwalent w języku polskim. Również powtórzenie badania z wykorzystaniem poświadczeń ekscerpowanych ze wszystkich tekstów czeskich (bez rozróżnienia na oryginał i na przekład) nie zmieniło diametralnie wyników, choć KonText wyszukał 4244 segmenty równoległe. Zrodziło to przypuszczenie, iż w przypadku tej jednostki rezultatu nie przyniosą nie tylko badania nurtu Pattern Grammar, ale także inne badania opierające się na kontekście liniowym.

Analiza stochastyczna

Wyniki analizy stochastycznej opartej na kontekście linearnym potwierdziły wcześniejsze przypuszczenia; wykorzystując tę metodę nie udało się wskazać konkretnego ekwiwalentu czasownika *toužit*. Warto tu jednak powtórzyć, iż eksperyment ten uzyskał lepsze wyniki podczas analizy jednostki *být líto*.

Wyniki badań opartych na kontekście zależności nie potwierdziły natomiast hipotezy, iż wybór ekwiwalentu uzależniony jest od kontekstu czeskiego leksemu.

Word Sketch

Analiza z wykorzystaniem narzędzia Sketch Engine nie była w pełni możliwa, ponieważ jak zostało to już wspomniane, usługa generowania profili typu Word Sketch nie jest dostępna dla języka czeskiego poprzez InterCorp; czeskie przykłady ekscerpowane są z dużo obszerniejszego CNK. Eksperymentalne badania danych pochodzących z NKJP, CNK i InterCorpu potwierdziły tezę o ekwiwalencji fraz *toužit + inf* i *pragnąć + inf* oraz ukazały, iż polskie czasowniki *marzyć* i *tęsknić* nie wykazują naturalnej łączliwości z bezokolicznikiem i dla frazy *toužit + inf* jednostki te (*marzyć* i *tęsknić*) nie będą preferowanym ekwiwalentem.

Wnioski

Celem prezentowanego opracowania było sprawdzenie, czy w oparciu o dane z korpusu równoległego InterCorp można stwierdzić, jaka metoda jest najefektywniejsza przy ustalaniu polskich ekwiwalentów czeskich czasowników. Do analizy wybranych zostało kilka podejść badawczych ułożonych w swoisty algorytm. Zaprezentowane przykłady to wyselekcjonowane jednostki czeskie, sprawiające szczególne problemy przy przekładzie na język polski.

Podsumowanie studiów przypadków

Przedstawione czeskie jednostki okazały się po części podwójnie wieloznaczne; po pierwsze są polisemiczne już na gruncie języka czeskiego, po drugie poszczególne ich znaczenia odnotowane przez słowniki języka czeskiego mogą być interpretowane niejednoznacznie przez osobę polskojęzyczną. W oczywisty sposób utrudnia to zadanie stojące przed tłumaczem, który stara się oddać w tekście wszystkie elementy semantyczne reprezentowane przez czeską jednostkę: np. czasownik *zdát se* (rozdział 6), który obejmuje percepcję wzrokową, aspekt intelektualny, element emocji, komponent obiektywizmu i nieosobowości, a także ramę modalną czy czasownik *toužit* (rozdział 12), który zawiera w sobie pragnienie, tęsknotę i marzenie.

Poszczególne etapy opisanego algorytmu, jak wynika z przeprowadzonych analiz, różnie sobie radzą z ustalaniem trafnych ekwiwalentów. Wstępem do każdej analizy jest wyszukanie ekwiwalentów słownikowych. W procesie przekładu ta tradycyjna faza może mieć prymarne znaczenie, a ekwiwalent wskazany przez słownik może być tym optymalnym. Stało się tak m.in. w przypadku leksemu *závidět* (rozdział 9); czasownik *závidět* ma w języku polskim odpowiednik *zazdrościć*, który został wskazany już przez słownik, a w dalszych badaniach zostało to potwierdzone. Podobnie dzieje się z czasownikiem *milovat* (rozdział 10); słownik czesko-polski podaje trafny ekwiwalent – *kochać*. Trzeba przy tym zaznaczyć, iż odwołanie do źródeł leksykograficznych było integralną częścią każdego

badania, a ekwiwalenty słownikowe każdorazowo były porównywane z ekwiwalentami uzyskanymi z korpusu InterCorp.

Kolejny etap – automatyczna ekstrakcja par ekwiwalentów – jest obecnie, dzięki narzędziu Treq, najprostszą metodą wyszukiwania ekwiwalentów przekładowych. Wykorzystując Treq czy kompilując własny słownik na podstawie poświadczeń korpusowych, można wskazać w niektórych przypadkach trafne ekwiwalenty w oparciu o dane statystyczne: *soucítit – współczuć* (rozdział 8), *závidět – zazdrościć* (rozdział 9) *milovat – kochać* (rozdział 10).

Dalszy krok to analiza walencyjna jednostek czeskich i ich potencjalnych polskich ekwiwalentów; jak się okazywało, w niektórych sytuacjach była tylko częściowo możliwa. Podstawowym problemem bywała zbyt mała liczba poświadczeń korpusowych; jak w przypadku czasownika *trápit* (rozdział 7) czy *soucítit* (rozdział 8). Z tego powodu nie było możliwe ustalenie, czy pojawienie się lub brak którejś z pozycji walencyjnej lub pewna kombinacja argumentów wpływa na wybór ekwiwalentu. Jednak również ta metoda okazała się pożyteczna. W trakcie analizy walencyjnej ustalić można, iż fraza *žárlit na* łącząca się z obiektem w bierniku ma w języku polskim odpowiednik – frazę *zazdrosny o* również łączącą się z obiektem w bierniku (rozdział 9). Analiza ta pozwoliła również znaleźć ekwiwalent jednego ze schematów walencyjnych czasownika *toužit* (rozdział 12); schematowi *toužit inf* odpowiada fraza *pragnąć inf.* Ponieważ bezokolicznik w tym schemacie nie łączy się na ogół ani z czasownikiem *tęsknić*, ani z *marzyć*, warto rozpatrzeć umieszczenie informacji o takim ograniczeniu w słownikach.

Wykorzystanie w dalszej fazie analizy gramatyki przypadków głębokich ukazało, iż podejście to w przypadku tak określonej grupy czasowników nie przynosi zamierzonego efektu. Ze względu na występowanie tych samych zestawów przypadków głębokich nie można ustalić trafnych ekwiwalentów. Warto jednak przypomnieć, iż dla innych czasowników to podejście może być relewantne.

Zastosowanie podejścia Pattern Grammar i ręczna analiza automatycznie wyekscerpowanych poświadczeń badanych jednostek pozwalają w niektórych przypadkach na identyfikację wzorców mających określony ekwiwalent w języku polskim. Powiodło się to m.in. w przypadku *vyjádřit soustrast* (rozdział 8), którego odpowiednikiem jest jednostka wyrażająca werbalne przekazanie kondolencji czy wyrazów współczucia: najczęściej *składać kondolencje,* ale także *składać / przekazywać wyrazy współczucia*. Podobnie rzecz się miała z czasownikiem *žárlit* (rozdział 9); fraza *žárlit na* + O znalazła swój ekwiwalent w postaci polskiej frazy *zazdrosny o* + O[357].

[357] Dzięki temu podejściu odnajdują swe odpowiedniki również wspomniane wcześniej jedynie na marginesie frazy – *být lito* (w rozdziale 11) oraz *cítit s* (w rozdziale 12).

W opracowaniu tym wykorzystywano również inne metody, np. siatkę znaczeń zastosowaną dla czasowników *závidět* i *žárlit* a także *zazdrościć* i *być zazdrosny* (rozdział 9) oraz dla jednostek *mít rád* i *milovat* oraz *kochać* i *lubić* (rozdział 10). Siatki te, bazujące na definicjach słownikowych oraz poświadczeniach korpusowych, w sposób dość szczegółowy ukazują, które znaczenia danych jednostek są ekwiwalentne. Informacja o tych odpowiedniościach mogłaby być pożyteczna dla leksykografów.

Dokonano również próby „rozbicia" jednostek na znaczenia i powiązania ich z ekwiwalentami; mowa o czeskim czasowniku *trápit* i polskich leksemach *męczyć, dręczyć, gnębić, trapić* (rozdział 7). Oznaczenie, które znaczenia reprezentowane przez analizowane czasowniki pokrywają się ze znaczeniami potencjalnych polskich ekwiwalentów, może ponownie mieć wpływ na kształt haseł słownikowych.

Wykorzystano także elementy metody Semantic Mirroring (Dyvik 2005)[358] (Vandevoorde 2016; De Baets, Vandevoorde, De Sutter 2018). W przypadku jednostek *soucítit* i *mít rád* sprawdzano, jak na język czeski przekładane są ich najczęstsze ekwiwalenty i czy tłumaczenie „wraca" do analizowanych jednostek wyjściowych. Na tej podstawie można ocenić m.in. trafność przekładu[359].

W opracowaniu tym podjęto także próbę wykorzystania analizy stochastycznej. Rezultaty analizy czasownika *toužit* opartej na kontekście zależności nie potwierdziły hipotezy, iż wybór ekwiwalentu uzależniony jest od leksemów składniowo zależnych od czeskiej jednostki. Również badania oparte na podstawie kontekstu linearnego nie wskazało konkretnego ekwiwalentu czasownika *toužit*. Warto jednak nadmienić, iż eksperyment ten uzyskał lepsze wyniki podczas analizy jednostki *být líto*. Fakt ten skłania ku przypuszczeniu, iż w przypadku języków fleksyjnych badanie kontekstu liniowego oraz stosowanie metody Pattern Grammar w przypadku większości czasowników mogą być niecelowe. Analiza ta może zakończyć się sukcesem w przypadku jednostek werbo-nominalnych czy złożonych, a swoją rolę może odegrać również szyk wyrazów.

W języku czeskim są też jednak jednostki, wobec których wybrane podejścia okazały się bezradne. Tak się stało w przypadku jednostki *mít rád* (rozdział 5), której wieloznaczność trudno oddać w języku docelowym. I choć sam leksem wydaje się być łatwo przetłumaczalny (jako *lubić*) zarówno dzięki słownikowi, jak i poświadczeniom korpusowym, to jednak interpretacja wyznania – *mám tě*

[358] Metoda Semantic Mirroring opisywana była również we wcześniejszych pracach (Dyvik 1998, 2004).

[359] Zastosowanie tej metody w semantyce leksykalnej i leksykografii w oparciu o komputerowe przetwarzanie danych omówione jest w artykule *Computing Semantic Clusters by Semantic Mirroring and Spectral Graph Partitioning* (Eldén, Merkel, Ahrenberg, Fagerlund 2013).

rád / mám tě rada – wzbudza wątpliwości co do znaczenia, kiedy obiektem jest człowiek[360].

W świetle tych danych trudno wskazać jedną najlepszą metodę; optymalnym wydaje się powiązanie różnych metod, jak zostało to przedstawione w tej pracy. Niektóre podejścia się uzupełniają (np. definicje słownikowe i automatyczna ekstrakcja par ekwiwalentów), inne precyzują wyniki poprzednich (np. analiza walencyjna i Pattern Grammar). Zastosowanie różnorodnych metod może ułatwić ustalenie optymalnego ekwiwalentu.

Zaproponowany algorytm nie zamyka wachlarza możliwości badań; jest propozycją, którą można rozważyć na etapie poszukiwania ekwiwalentów, a także ją rozszerzyć. W oparciu o elementy tego algorytmu powstało już szereg prac dotyczących czesko-polskiej i polsko-czeskiej ekwiwalencji; powstaje także publikacja próbująca wykorzystać ten algorytm w badaniu gramatyczno-leksykalnym[361]. W świetle tych badań koniecznym wydaje się rozszerzenie inwentarza analizowanych czasowników zarówno z grupy tu opisywanej, jak i innych grup, np. czasowników ruchu ze szczególnym uwzględnieniem znaczenia prefiksów[362].

Perspektywy

Badania korpusowe w tej pracy wychodziły od automatycznej ekscerpcji danych, po czym następowała analiza statystyczna, semantyczna i składniowa. Część z zaplanowanych badań nie została jednak przeprowadzona z uwagi na niewystarczającą liczbę poświadczeń. Pojawia się tym samym pytanie, jak duża próbka danych mogłaby być uznana za wystarczającą; to problem często podnoszony w przypadku analiz korpusowych (Hebal-Jezierska, Kaczmarska i Rosen 2016). Powiększenie czesko-polskiej części korpusu InterCorp wpłynęłoby z pewnością pozytywnie na możliwość przeprowadania nań badań. Inną kwestią jest perspektywa oceniania potencjalnych odpowiedników[363]; warto

[360] Także we frazach odnoszących się do osoby trzeciej (*měl tě rád / měla ho rada*) często nie można ustalić, czy chodzi o miłość czy sympatię. Również kontekst w wielu sytuacjach nie pomaga w rozwikłaniu tej kwestii.

[361] *Parallel corpus as functional context of aspectual interpretation – the case of Czech, Croatian and Polish biaspectual verbs* (Jurkiewicz-Rohrbacher, Edyta; Kaczmarska, Elżbieta; Rosen, Alexandr)

[362] Badania nad słowackimi ekwiwalentami tłumaczeniowymi polskich czasowników przedrostkowych prowadziła Marianna Petrincová (Petrincová 2016).

[363] Publikacja dotycząca ustalania ekwiwalencji na podstawie danych z polskiego i amerykańskiego WordNetu wymienia trzy typy ekwiwalencji: *strong*, *regular* i *weak*, wyróżniane na podstawie kryteriów formalnych, semantycznych i przekładowych (Rudnicka, Piasecki, Bond, Grabowski, Piotrowski 2019: 3).

wziąć pod uwagę sytuacje, w jakich się pojawiają dane leksemy, ich konotację, prozodię semantyczną[364] oraz ograniczenia semantyczne. Gdyby InterCorp dysponował rozszerzoną anotacją, część z tych informacji mogłaby być dostępna automatycznie, co pozwoliłoby na uproszczenie procesu ustalania ekwiwalentów już na etapie wyszukiwania poświadczeń[365]. Istotny jest również wybór odpowiedniego *tertium comparationis*[366].

Obiecującym narzędziem wspomagającym ustalanie ekwiwalencji czesko-polskiej wydaje się Sketch Engine, jednak analiza z wykorzystaniem tego narzędzia nie była do tej pory w pełni możliwa; usługa generowania profili typu Word Sketch nie jest bowiem dostępna dla języka czeskiego poprzez InterCorp, a czeskie przykłady ekscerpowane są z dużo obszerniejszego CNK. Możliwość generowania paralelnych profili Word Sketch z zasobów InterCorpu jest więc kwestią przyszłości. Rozwiązaniem jest utworzenie własnego korpusu[367] i wygenerowanie paralelnych profili Word Sketch[368].

Innym – również przyszłościowym rozwiązaniem – mogłoby być zastosowanie narzędzi *WoSeDon* i *Słowosieci*. W rozdziale 11 opisany został eksperyment z czasownikiem *cierpieć*; jego rezultatem było połączenie *cierpieć*$_1$ z *trpět*$_2$. Dzięki ujednoznacznieniu czasowników polskich i czeskich, mogłoby dojść do wiązania ich konkretnych znaczeń; warto rozważyć również zintegrowanie polskiej *Słowosieci* i czeskiego WordNetu[369].

Analiza danych zebranych na podstawie korpusu równoległego ukazała również, iż mimo swoich ograniczeń InterCorp w przypadku każdego badania prezentował szereg możliwości rozumienia i przekładu analizowanych jednostek; dla językoznawców, a szczególnie leksykografów i tłumaczy stanowi zarówno wsparcie, jak i inspirację.

[364] Ta jednak, w zależności od kontekstu, może być różna.

[365] Obecnie, przeprowadzając badania, anotację semantyczną można przeprowadzić jedynie manualnie.

[366] Zbigniew Greń proponuje dwa możliwe punkty wyjścia w badaniach leksykalnych języków blisko spokrewnionych: 1) od formy – wyrazów słownikowych czy tekstowych, 2) od funkcji semantycznej – pojęć realizowanych przez określone leksemy (Greń 2004: 180).

[367] Od niedawna tworzenie własnego korpusu w języku polskim jest dużo prostsze dzięki narzędziu *Korpusomat* (Kieraś, Kobyliński i Ogrodniczuk 2018), służącym do samodzielnego tworzenia elektronicznych korpusów tekstów.

[368] Usługa dostępna na stronie https://www.sketchengine.eu/

[369] O ekwiwalencji pomiędzy poszczególnymi znaczeniami w polskim i angielskim WordNecie – (Rudnicka, Bond, Grabowski, Piasecki i Piotrowski 2017) (Rudnicka, Piasecki, Bond, Grabowski, Piotrowski 2019).

Załącznik 1

soucítit syn v7 freq = 3,986 (0.8 per million)

coord	257	-1.6
litovat	13	1.74
chápat	10	0.71
rozumět	6	0.13

post_prep	1,853	-7.5
s	1,815	1.06

post_s	906	-20.6
utrpení	16	3.32
outsider	4	1.98
oběť	47	1.86
lid	9	1.35
rodina	80	0.8
neštěstí	5	0.71
uprchlík	4	0.09

Obraz 30. Łączliwość czasownika *soucítit* wygenerowana przez Sketch Engine

soucit syn v7 freq = 19.381 (3.8 per million)

coord	3.343	-3.4
snášenlivost	59	8.22
milosrdenství	66	7.99
empatie	88	7.9
litost	185	7.66
slitování	27	7.3
něha	54	7.2
laskavost	45	6.99
jemnocit	13	6.77
pochopení	194	6.36
útrpnost	9	6.33
neohrožen ost	9	6.26
vcítění	11	6.24
altruismus	11	6.09
solidarita	131	6.06
bázeň	17	5.97
porozumění	87	5.94
lidskost	32	5.73
vlídnost	9	5.63
odpuštění	29	5.52
soustrast	14	5.45
vřelost	7	5.33
obětavost	18	5.32
sympatie	57	5.27
ohleduplnost	14	5.2
dojetí	20	5.12

prec_verb	1.455	-2.0
vzbuzovat	169	5.73
budit	138	5.04
zasluhovat	25	5.01
dovolávat	6	4.43
probouzet	22	4.28
vynucovat	6	4.15
zasloužit	70	3.18
vzbudit	29	3.05
vyvolávat	63	3.05
zneužívat	11	2.5
projevovat	30	2.34
pocítit	11	2.2
cítit	68	2.17
vyjadřovat	35	2.12
pociťovat	10	2.11
ubránit	6	1.71
projevit	23	1.3
probudit	6	1.14
vyvolat	20	1.1
mízet	7	0.85
potřebovat	63	0.66
postrádat	8	0.57
rozvíjet	7	0.43
znát	42	0.26
posilovat	5	0.15

post_verb	569	-0.7
usmrtit	14	3.88
zasluhovat	8	3.42
zasloužit	26	1.76
budit	9	1.12
vzbuzovat	6	0.93

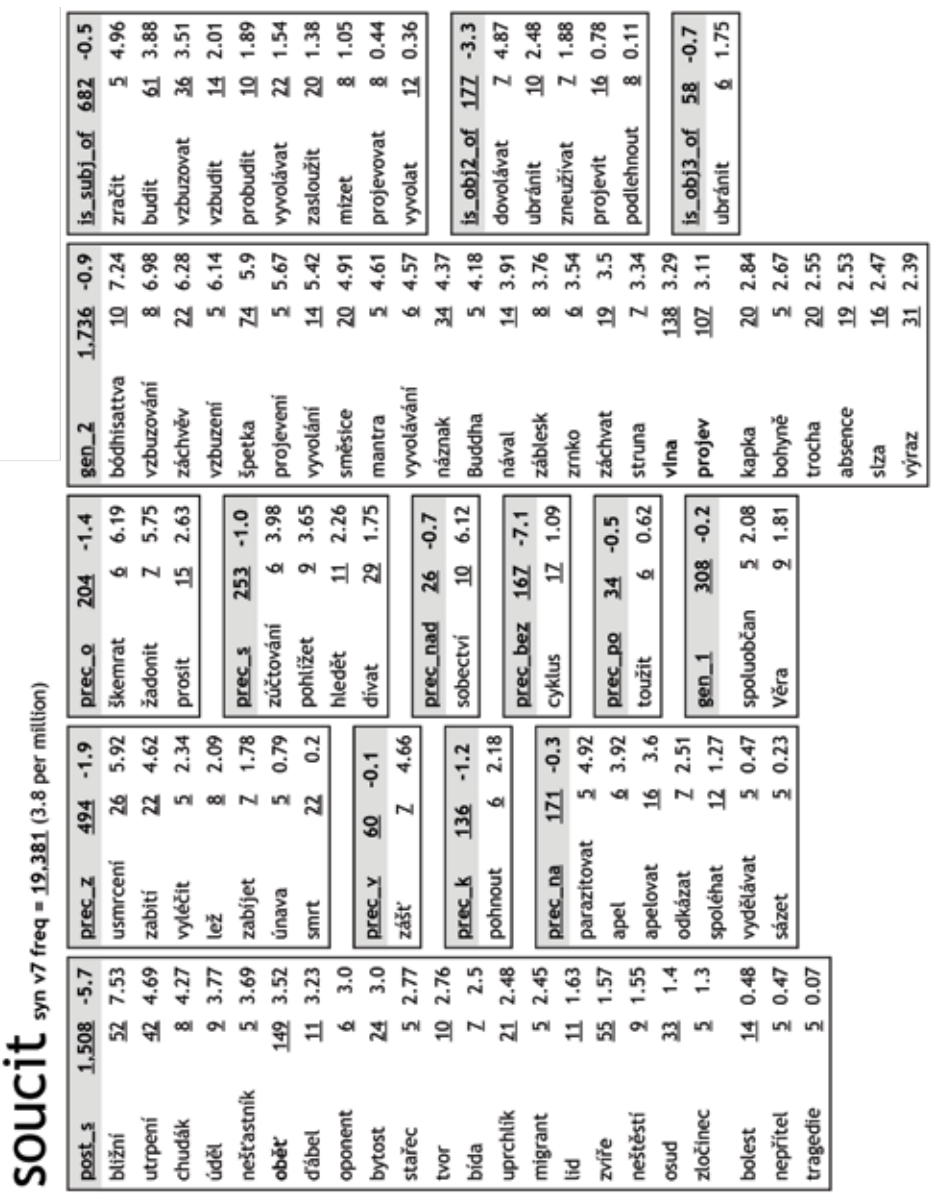
a_modifier	2.547	-0.5
vzbuzující	38	7.93
budící	24	7.27
dusivý	21	6.89
sentimentální	27	6.51
diváků v	12	6.09
projevený	10	5.85
předstíraný	13	5.82
projevovaný	5	5.68
chápavý	12	5.57
útrpný	9	5.51
pohrdavý	6	5.48
upřímný	50	5.33
nelíčený	5	5.29
bezmezný	8	5.22
nemístný	6	4.86
pobavený	5	4.79
falešný	92	4.75
shovívavý	5	4.54
přehnaný	27	4.53
okázalý	12	4.53
aristokratický	5	4.38
bratrský	16	4.36
láskyplný	7	4.35
vyvolávající	5	4.31
ponižující	6	4.29

prec_prep	3.000	-0.9
namísto	7	2.48
skrz	10	2.34
bez	347	1.78
z	1.181	0.54

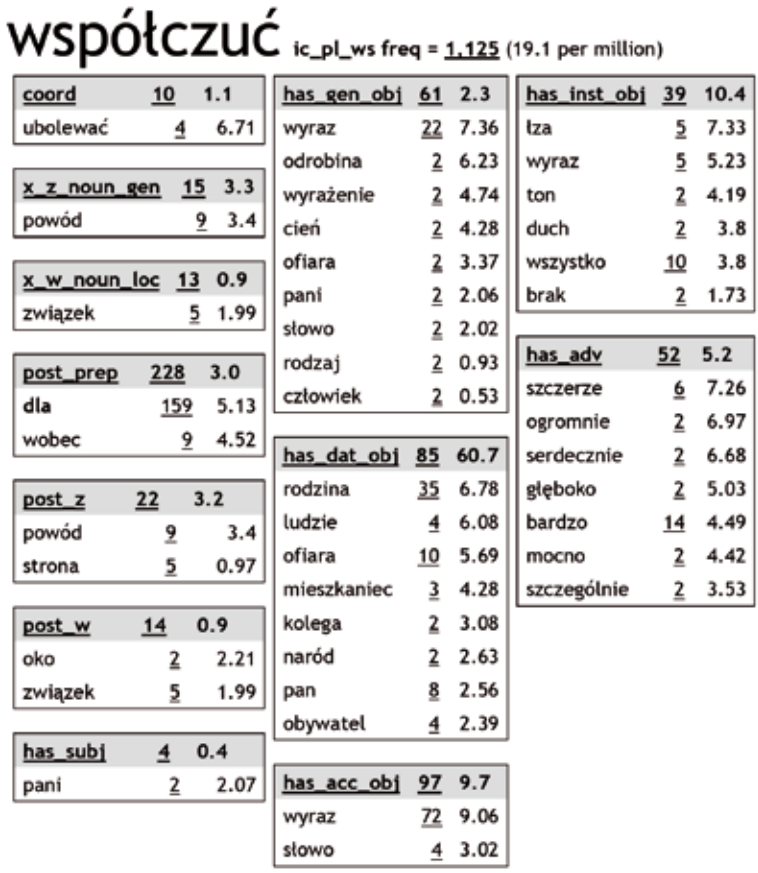
post_v	110	-0.1
zemědělství	9	0.07

post_k	231	-2.0
bližní	12	5.62
chlapec	77	2.81

Obraz 31. Łączliwość rzeczownika *soucit* wygenerowana przez Sketch Engine (1)



Obraz 32. Łączliwość rzeczownika *soucit* wygenerowana przez Sketch Engine (2)



Obraz 33. Łączliwość czasownika *współczuć* wygenerowana przez Sketch Engine

Załącznik 2

závidět syn v7 freq = 61,350 (12.0 per million)

coord	1,183	-1.8
pomlouvat	40	7.4
žárlit	26	6.45
nenávidět	22	4.47
žasnout	8	4.02
ubližovat	5	3.41
gratulovat	9	3.4
vyčítat	19	3.39
obdivovat	49	3.35
přát	88	3.08
tleskat	8	2.75
hádat	6	2.44
nadávat	9	2.42
litovat	14	1.83
škodit	8	1.64
házet	8	1.11
chtít	47	1.02
toužit	8	1.02
snášet	7	0.82
doufat	11	0.63
krást	5	0.3
chápat	7	0.19
stěžovat	6	0.14

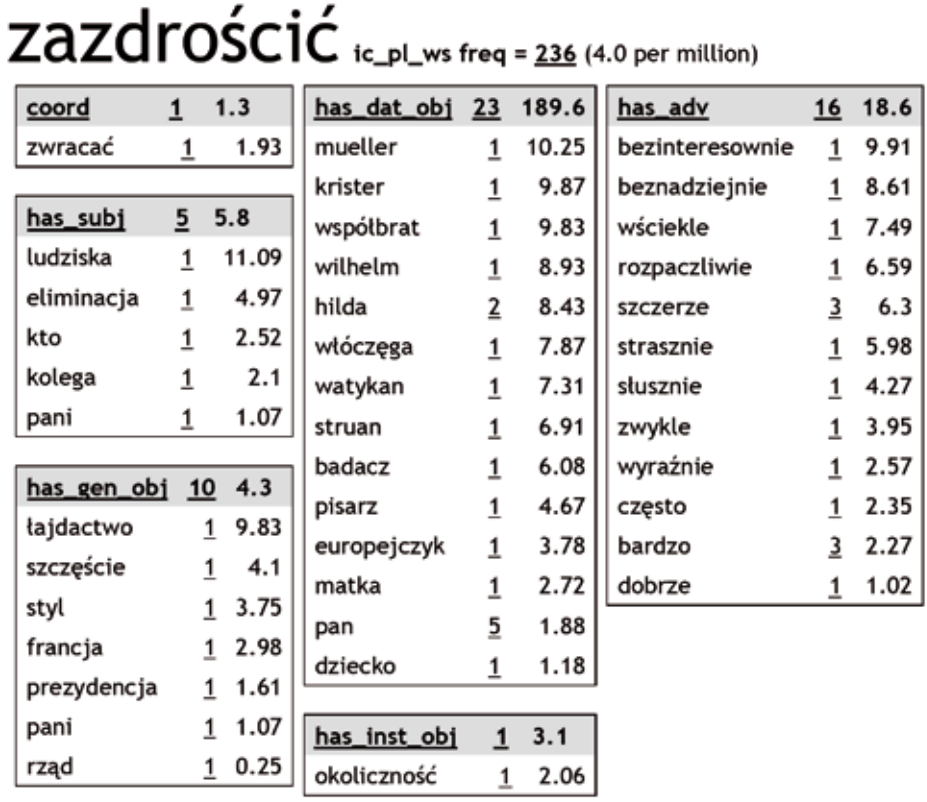
post_o	34	-0.3
generace	21	0.53

has_subj	6,139	-6.9
divačka	8	4.59
spolužák	52	4.4
vrstevník	22	4.31
třicátník	10	4.19
našinec	7	4.04
vrstevnice	5	4.02
soused	98	3.83
sólistka	7	3.73
pěvkyně	12	3.66
Pražák	8	3.59
chlap	43	3.56
padesátník	6	3.49
fanynka	10	3.45
kolega	200	3.44
čtyřicátník	5	3.32
kolegyně	25	3.22
Jim	7	3.17
kamarádka	27	3.15
spolužačka	5	2.98
konkurent	27	2.86
profesionál	21	2.77
sousedka	6	2.76
dvacítka	11	2.32
Rakušan	8	2.16
modelka	14	2.14

has_obj3	2,619	-46.2
spolužačka	13	4.62
Noro	5	3.26
Angličan	10	3.24
vrstevník	9	3.17
spolužák	21	3.16
soused	47	2.8
kamarádka	19	2.7
Maďar	7	2.63
kolega	105	2.52
kolegyně	14	2.45
Slovák	18	2.34
Rakušan	7	2.06
Francouz	12	2.01
Japonec	5	2.01
modelka	11	1.84
pták	19	1.8
Barcelona	9	1.8
Pražan	8	1.73
důchodce	25	1.68
kamarád	32	1.61
Brit	6	1.5
Ital	6	1.34
vinař	6	1.28
Polák	8	1.25
Němec	22	1.11

has_obj4
kubismus
vitalita
penis
kolegyně
luxus
popularita
dobrota
úspěch
sláva
bohatství
přízeň
figura
odvaha
nos
penze
nadšení
svět
kamarádka
talent
cestování
peníze
plat
návštěvnost
schopnost
zbytek

Obraz 34. Łączliwość czasownika *závidět* wygenerowana przez Sketch Engine



Obraz 35. Łączliwość czasownika *zazdrościć* wygenerowana przez Sketch Engine

žárlit syn v7 freq = 13,835 (2.7 per million)

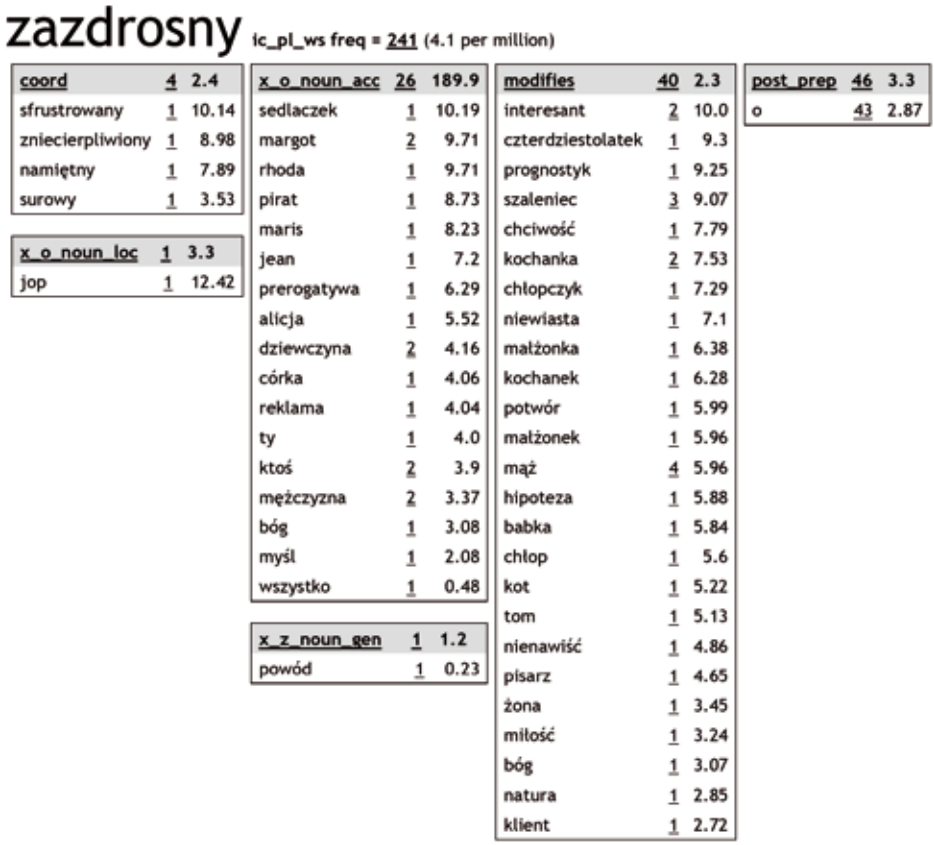
coord	641	-1.9
pomlouvat	5	4.5
závidět	26	4.1
odpouštět	4	3.03
nenávidět	7	2.84
zuřit	7	2.8
vyvádět	4	2.77
vyčítat	6	1.74
nadávat	5	1.59
bít	6	1.41
chtít	35	0.6
snášet	5	0.34
milovat	10	0.32
hlídat	8	0.28
toužit	4	0.03

post_prep	2,891	-5.7
na	2,782	0.37

post_na	1,104	-5.9
fanynka	12	3.97
Nicole	4	3.27
sourozenec	13	2.92
Romano	4	2.58
popularita	11	1.98
partnerka	8	1.97
milenka	4	1.83
milenec	5	1.7
přítelkyně	11	1.59
sok	4	1.5
kolegyně	6	1.25
úspěch	72	1.14
kamarádka	5	0.79
manžel	20	0.73
sláva	8	0.69
manželka	15	0.47
táta	4	0.44
modelka	4	0.4
oblíba	4	0.39
bratr	15	0.37
miminko	4	0.34
sestra	12	0.16

has_subj	2,183	-4.7
blázen	190	7.8
Moja	5	5.69
Krainová	5	4.94
Mandlová	4	4.86
Angelina	6	3.71
manžel	111	3.2
chlap	31	3.13
Emil	14	2.99
Pešek	4	2.89
Eliáš	5	2.86
přítelkyně	26	2.82
partnerka	14	2.75
Alice	6	2.64
Anička	5	2.63
manželka	63	2.54
koza	8	2.52
Diana	5	2.52
Felix	4	2.51
fanynka	4	2.33
Max	5	2.16
milenec	7	2.16
milenka	5	2.11
sourozenec	7	2.0
partner	51	1.36
přítel	36	1.08

Obraz 36. Łączliwość czasownika *žárlit* wygenerowana przez Sketch Engine



Obraz 37. Łączliwość czasownika *zazdrosny* wygenerowana przez Sketch Engine

Elżbieta Kaczmarska

Methods of determining equivalents of verbs expressing emotions in translation from Czech to Polish, based on texts from the InterCorp parallel corpus

This study focuses on the problem of determining Polish equivalents of Czech verbs. The aim of the analysis is to explore which method can best handle this task and what factors determine the choice of the equivalent. The object of the analysis are Czech lexical units referring to various mental or emotional states. Although Czech is a language closely related to Polish and false friends might be naively expected to pose the main challenge, finding an appropriate Polish equivalent for most lexemes of this type involves many issues common in translation from more distant languages. This concerns both the task of decoding the meaning of the lexeme in the source text according to the conceptual patterns of the source language as well as encoding the meaning in the target language according to its conceptual patterns.

To study factors responsible for an optimal choice of an equivalent, the inquiry is carried out on texts from a parallel corpus. Parallel texts offer the chance to examine many occurrences of the source and target lexemes in their morphosyntactic, structural, semantic and pragmatic contexts, including details about the situation in which the lexemes are used and about the text itself. Given the challenging choice of psych verbs, fiction texts or texts of a similar type are needed. InterCorp is a parallel corpus that includes a substantial share of such texts. On the basis of available lexicons and the corpus data, the task can then be approached from various perspectives, resulting in a sequence of steps, each invoking a method of searching for an equivalent.

The perspectives combine the approaches of translatology, theoretical linguistics and corpus linguistics. Through the prism of translation, the author shows

the possibilities of applying various linguistic theories (valency theory, Case Grammar, Pattern Grammar) and corpus tools (Treq, Sketch Engine), demonstrating the merits of using the parallel corpus, but also identifying its limitations. Building on these resources, an algorithm to determine Polish equivalents for the Czech verbs is proposed.

The agenda is reflected in the outline of the book, which consists of thirteen chapters, best seen as forming four parts – theoretical and data-related (Chapters 1–5), analytical (Chapters 6–10), proposal-related (Chapters 11–12), and a summary (Chapter 13). Bibliography and appendices conclude the volume. Whenever appropriate, the text is amply illustrated with examples and accompanied by tables, summarizing statistical data.

After a short introduction, Chapter 2 discusses theoretical issues concerning equivalence and translation, as well as presenting various definitions of equivalence and the relationship between traditional translation and automatic alignment. The next chapter includes comments on the correlation of translation and contrastive studies (Chapter 3). The InterCorp parallel corpus is presented in Chapter 4, together with specifics of its Czech-Polish section and details about Treq, the database of lexical equivalents extracted from the parallel corpus. Chapter 5 introduces the methodology used in the six following case studies, each concerned with a Czech verbal lexeme. The theoretical part proceeds with chapters presenting the case studies: *zdát se* 'seem' (Chapter 6), *trápit* 'bother, worry, torment' (Chapter 7), *soucítit* 'sympathize' (Chapter 8), *závidět* 'envy' and *žárlit* 'be jealous' (Chapter 9), *mít rád* 'like' and *milovat* 'love' (Chapter 10). The following chapter presents an algorithm that facilitates the determination of equivalents. The procedure consists of the following steps: dictionary lookup, automatic extraction of equivalent pairs from the parallel corpus, valency-based analysis, analysis in terms of Case Grammar (deep cases), analysis in terms of Pattern Grammar, and analysis employing other methods or resources: Word Sketch, WordNet, WoSeDon (Chapter 11). The algorithm is exemplified in the following chapter, a case study of the Czech verb *toužit* 'desire' (chapter 12). Each of the steps is described in detail and exemplified to show the extent of usability of the overall methodology and of each of the individual approaches. The concluding summary indicates further research perspectives (Chapter 13).

The monograph presents an interdisciplinary analysis combining the approaches of linguistics and translatology, together with elements of statistics and computational linguistics, based on data from a parallel corpus. The analysis results in the design of a procedure aiming at the choice of the most fitting lexical equivalent. Recent years have seen a number of research results published in the field combining corpus linguistics and translation. However, most of them concern English. The study is unique not only due to the languages studied, but also due to the methods employed, especially methods from the borderline of translation studies and comparative analysis.

Elżbieta Kaczmarska

Metody určování ekvivalentů sloves vyjadřujících emoce v překladu z češtiny do polštiny na základě textů z paralelního korpusu InterCorp

Tato studie se zaměřuje na problematiku určování polských ekvivalentů českých sloves. Cílem analýzy je prozkoumat, která metoda může tento úkol nejlépe zvládnout a jaké faktory volbu ekvivalentu určují. Předmětem analýzy jsou české lexikální jednotky vyjadřující různé duševní nebo emocionální stavy. I když je čeština jazyk blízký polštině a podle naivního očekávání by mohly být hlavním úskalím falešní přátelé, najít vhodný polský ekvivalent pro většinu lexémů tohoto typu s sebou nese řadu problémů běžných při překladech z typologicky vzdálenějších jazyků. To se týká jak úlohy dekódování významu lexému ve zdrojovém textu v pojmové struktuře zdrojového jazyka, tak i kódování významu v pojmové struktuře a výrazech jazyka cílového.

Aby našla faktory odpovědné za optimální volbu ekvivalentu, zabývá se autorka texty z paralelního korpusu. Paralelní texty nabízejí možnost prozkoumat mnoho výskytů zdrojových a cílových lexémů v kontextu, a to v kontextu analyzovaném morfosyntakticky, syntakticky, sémanticky i pragmaticky, včetně podrobností o situaci, ve které se lexém vyskytuje, a podrobností o samotném textu. Slovesa vyjadřující emoce nejsou dostatečně zastoupeny v každém typu textů. Potřeby studia chování takových sloves nejlépe splňuje beletrie a podobné žánry. V paralelním korpusu InterCorp je takových textů poměrně dost. Na základě dostupných slovníků a korpusových dat je pak možné úkol najít nejvhodnější ekvivalent nahlížet z různých úhlů. Výsledkem je postupnost kroků, z nichž každý představuje odlišnou metodu, jak ekvivalent najít.

Použité metody kombinují přístupy translatologie, teoretické lingvistiky a korpusové lingvistiky. Prismatem překladu autorka ukazuje možnosti uplatnění různých lingvistických teorií (valenční teorie, Case Grammar, Pattern

Grammar) a korpusových nástrojů (Treq, Sketch Engine). Ozřejmují se přitom výhody využití paralelního korpusu, ale také jeho omezení. V návaznosti na tyto zdroje je pak navržen algoritmus pro stanovení polských ekvivalentů českých sloves.

Půdorys studie se promítá do osnovy knihy, která se skládá z třinácti kapitol. Ty se logicky člení do čtyř částí – části věnující se teorii a datům (kapitoly 1–5), části analytické (kapitoly 6–10), části věnované návrhu algoritmu (kapitoly 11–12) a shrnutí (kapitola 13). V závěru je uvedena bibliografie a přílohy. Text je bohatě ilustrován příklady a doplněn tabulkami, které shrnují statistická data.

Po krátkém úvodu pojednává kapitola 2 o teoretických otázkách ekvivalence a překladu a prezentuje různé definice ekvivalence a vztahu mezi tradičním překladem a automatickým zarovnáváním. Další kapitola se věnuje vztahu překladatelských a kontrastivních studií (kapitola 3). Paralelní korpus InterCorp je stručně předveden v kapitole 4 spolu se specifiky česko-polské sekce InterCorpu a podrobnostmi o databázi lexikálních ekvivalentů extrahovaných z paralelního korpusu Treq. Kapitola 5 představuje metodiku použitou v šesti následujících případových studiích, z nichž každá se zabývá konkrétním českým slovesným lexémem. Teoretická část pokračuje kapitolami tyto případové studie obsahujícími: *zdát se* (kapitola 6), *trápit* (*se*) (kapitola 7), *soucítit* (kapitola 8), *závidět* a *žárlit* (kapitola 9), *mít rád* a *milovat* (kapitola 10). Další kapitola představuje algoritmus, který má za cíl nalezení ekvivalentů usnadnit. Procedura se skládá z následujících kroků: vyhledávání ve slovníku, automatická extrakce ekvivalentních párů z paralelního korpusu, valenční analýza, analýza z hlediska Case Grammar (na základě hloubkových pádů), analýza z hlediska Pattern Grammar a analýza s využitím jiných metod nebo zdrojů: Word Sketch, WordNet, WoSeDon (kapitola 11). Využití algoritmu je předvedeno v následující kapitole, případové studii českého slovesa *toužit* (kapitola 12). Každý z kroků je podrobně popsán a doložen příklady, které ukazují rozsah použitelnosti celkové metodiky a jednotlivých přístupů, včetně případů, kdy daný krok žádoucí výsledek nepřináší. Závěrečný souhrn uvádí další výzkumné perspektivy (kapitola 13).

Monografie představuje interdisciplinární analýzu, kombinující přístupy lingvistiky a translatologie spolu s některými postupy statistiky a informatiky, a to na základě dat z paralelního korpusu. Výsledkem analýzy je návrh procedury, jejímž cílem je výběr nejvhodnějšího lexikálního ekvivalentu. Prací, které kombinují korpusovou lingvistiku a překlad, se v posledních letech objevilo hodně. Většina z nich se však týká angličtiny. Tato studie je ovšem jedinečná nejen díky studovaným jazykům, ale také díky použitým metodám, zejména metodám na pomezí translatologie a kontrastivní analýzy.

Źródła

Przykłady zamieszczone w pracy pochodzą z następujących źródeł:

InterCorp

AKKOS1970pl	Kuśniewicz, A. (1970). *Król obojga Sycylii.* Warszawa: Państwowy Instytut Wydawniczy.
AKKOS1975cz	Kuśniewicz, A. (1975). *Král obojí Sicilie.* Praha: Odeon. Tłumaczenie: A. Balajková.
ALPDP2010cz	Lindgren, A. (2010). *Pipi Dlouhá punčocha.* Praha: Albatros. Tłumaczenie: D. Hartlová i J. Vohryzek.
ALPL1945se	Lindgren, A. (1945). *Pippi Långstrump.* Stockholm: Rabén & Sjögren.
ALPL1954en	Lindgren, A. (1954). *Pippi Longstocking.* London: Puffin. Tłumaczenie: E. Hurup.
ALPP1961pl	Lindgren, A. (1961). *Pippi Pończoszanka.* Warszawa: Nasza Księgarnia. Tłumaczenie: I. Szuch-Wyszomirska.
ASHP2001cz	Stasiuk, A. (2001). *Haličské povídky.* Olomouc: Periplum. Tłumaczenie: J. Kamińska.
ASMO1993cz	Sapkowski, A. (1993). *Meč osudu.* Praha: Winston Smith. Tłumaczenie: J. Pilch, D. Chmelař i R. Sochor.
ASMP1992pl	Sapkowski, A. (1992). *Miecz przeznaczenia.* Warszawa: Supernowa.
ASOG2006pl	Stasiuk, A. (2006). *Opowieści galicyjskie.* Wołowiec: Wydawnictwo Czarne.
BHANM1992cz	Hrabal, B. (1992). *Aurora na mělčině.* Praha: Pražská imaginace.
BHANM2005pl	Hrabal, B. (2005). *Aurora na mieliźnie.* Warszawa: Świat Literacki. Tłumaczenie: M. Falski.
BHPHS1994cz	Hrabal, B. (1994). *Příliš hlučná samota.* Praha: Pražská imaginace.
BHZGS1993pl	Hrabal, B. (1993). *Zbyt głośna samotność.* Kraków: Wydawnictwo Literackie. Tłumaczenie: P. Godlewski.
CMDI1981pl	Miłosz, C. (1981). *Dolina Issy.* Kraków: Wydawnictwo Literackie.
CMRE1990pl	Miłosz, C. (1990). *Rodzinna Europa.* Warszawa: Czytelnik.
CMRE1997cz	Miłosz, C. (1997). *Rodná Evropa.* Olomouc: Votobia. Tłumaczenie: H. Stachová.

CMUI1993cz	Miłosz, C. (1993). *Údolí Issy.* Praha: Mladá fronta. Tłumaczenie: H. Stachová.
DMCAB2004cz	Masłowska, D. (2004). *Červená a bilá.* Praha: Euromedia Group. Tłumaczenie: B. Gregorová.
DMWPR2002pl	Masłowska, D. (2002). *Wojna polsko-ruska pod flagą biało-czerwoną.* Warszawa: Lampa i Iskra Boża.
ERA1973pl	Redliński, E. (1973). *Awans.* Warszawa: Czytelnik.
ERV1977cz	Redliński, E. (1977). *Vzestup.* Praha: Odeon. Tłumaczenie: V. Dvořáčková.
ESS1971pl	Stachura, E. (1971). *Siekierezada albo Zima leśnych ludzi.* Warszawa: Czytelnik.
ESS1980cz	Stachura, E. (1980). *Sekerezáda aneb Zima lesních lidí.* Praha: Odeon. Tłumaczenie: H. Stachová.
GHGDP1995cz	Herling-Grudziński, G. (1995). *Deník psaný v noci.* Praha: Torst. Tłumaczenie: H. Stachová.
GHGDPN1993pl	Herling-Grudziński, G. (1993). *Dziennik pisany nocą.* Kraków: Wydawnictwo Literackie.
HCJS2008pl	Coben, H. (2008). *Jedyna szansa.* Warszawa: Albatros. Tłumaczenie: Z. A. Królicki.
HCTNN2003cz	Coben, H. (2003). *Teď, nebo nikdy.* Praha: Knižní klub. Tłumaczenie: P. Kůsová.
HGZV1996cz	Grynberg, H. (1996). *Židovská válka.* Praha: Aurora. Tłumaczenie: O. Hostovská.
HGZW1965pl	Grynberg, H. (1965). *Żydowska wojna.* Warszawa: Czytelnik.
IBSLAV1997cz	Singer, I. B. (1997). *Láska a vyhnanství.* Praha: Argo. Tłumaczenie: M. Bidlová.
IBSMIW1991pl	Singer, I. B. (1991). *Miłość i wygnanie.* Wrocław: Wydawnictwo Dolnośląskie. Tłumaczenie: L. Czyżewski.
IDB2003pl	Dousková, I. (2003). *Bądźżesz.* Poznań: Prószyński i S-ka. Tłumaczenie: J. Derdowaska.
IDHB1998cz	Dousková, I. (1998). *Hrdý Budžes.* Praha: Hynek.
JCVC2005cz	Chmielewska, J. (2005). *Všude červená.* Tłumaczenie: A. Penčeva.
JCWC2001pl	Chmielewska, J. (2001). *Wszystko czerwone.* Praha: Euromedia Group. Tłumaczenie: I. Daňková.
JDB1993cz	Durych, J. (1993). *Bloudění.* Brno: Atlantis.
JDZ1959pl	Durych, J. (1959). *Zbłąkani.* Kraków: Instytut Wydawniczy „PAX". Tłumaczenie: M. Erhartowa.
JGD1990cz	Gruša, J. (1990). *Dotazník.* Brno: Atlantis.
JGK2003pl	Gruša, J. (2003). *Kwestionariusz, czyli modlitwa za pewne miasto i przyjaciela.* Warszawa: Twój Styl. Tłumaczenie: P. Godlewski.
JHODV1996cz	Hašek, J. (1996). *Osudy dobrého vojáka Švejka za světové války.* Praha: Baronet.

JHPDW1957pl	Hašek, J. (1957). *Przygody dobrego wojaka Szwejka podczas wojny światowej.* Warszawa: Państwowy Instytut Wydawniczy. Tłumaczenie: P. Hulka-Laskowski.
JPMPS2006pl	Pilch, J. (2006). *Moje pierwsze samobójstwo.* Warszawa: Świat Książki.
JPMPS2009cz	Pilch, J. (2009). *Moje první sebevražda.* Zlín: Kniha Zlín. Tłumaczenie: N. Vrbovcová.
JPPMA2007pl	Pilch, J. (2007). *Pod Mocnym Aniołem.* Kraków: Wydawnictwo Literackie.
JPUSA2007cz	Pilch, J. (2007). *U strážnýho anděla.* Praha: Agite/Fra. Tłumaczenie: B. Gregorová.
JSBC2004pl	Škvorecký, J. (2004). *Batalion czołgów.* Łódź: Fundacja Pogranicze. Tłumaczenie: A. S. Jagodziński.
JSFS1999pl	Škvorecký, J. (1999). *Fajny sezon.* Warszawa: Twój Styl. Tłumaczenie: P. Godlewski.
JSLE1965pl	Škvorecký, J. (1965). *Legenda Emöke.* Praha: Ivo Železný.
JSLE1996cz	Škvorecký, J. (1996). *Legenda Emöke.* Warszawa: Czytelnik. Tłumaczenie: A. Piotrowski.
JSPILD1992cz	Škvorecký, J. (1992). *Příběh inženýra lidských duší.* Brno: Atlantis.
JSPILD2008pl	Škvorecký, J. (2008). *Przypadki inżyniera ludzkich dusz.* Warszawa: Dowody na Istnienie. Tłumaczenie: A. S. Jagodziński.
JSPS1991cz	Škvorecký, J. (1991). *Prima sezóna.* Praha: Odeon.
JST1970pl	Škvorecký, J. (1970). *Tchórze.* Katowice: Śląsk. Tłumaczenie: E. Witwicka.
JSTP1990cz	Škvorecký, J. (1990). *Tankový prapor.* Praha: Galaxie.
JSZ1964cz	Škvorecký, J. (1964). *Zbabělci.* Katowice: Československý spisovatel.
JUCZE1993cz	Updike, J. (1993). *Čarodějky z Eastwicku.* Praha: Svoboda. Tłumaczenie: J. Jirák.
JUCZE2008pl	Updike, J. (2008). *Czarownice z Eastwick.* Poznań: Dom Wydawniczy REBIS. Tłumaczenie: K. Bogucka-Krenz.
KCFA1971pl	Čapek, K. (1971). *Fabryka absolutu.* Katowice: Śląsk. Tłumaczenie: P. Hulka-Laskowski.
KCTNA1982cz	Čapek, K. (1982). *Továrna na absolutno.* Praha: Československý spisovatel.
LFOMB1980cz	Fuks, L. (1980). *Obraz Martina Blaskowitze.* Praha: Melantrich.
LFOMB1988pl	Fuks, L. (1988). *Obraz Marcina Blaskowitza.* Warszawa: Czytelnik. Tłumaczenie: E. Madany.
LFPZ1979pl	Fuks, L. (1979). *Palacz zwłok.* Kraków: Wydawnictwo Literackie. Tłumaczenie: J. Anderman i T. Lis.
LFSM1983cz	Fuks, L. (1983). *Spalovač mrtvol.* Praha: Mladá fronta.
LFVPTS2003cz	Fuks, L. (2003). *Variace pro temnou strunu.* Praha: Odeon.

LFWNNS1970pl	Fuks, L. (1970). *Wariacje na najniższej strunie.* Warszawa: Państwowy Instytut Wydawniczy. Tłumaczenie: M. Erhardt-Gronowska.
MADM1993cz	Ajvaz, M. (1993). *Druhé město.* Praha: Mladá fronta.
MAIN2005pl	Ajvaz, M. (2005). *Inne miasto.* Sejny: Pogranicze. Tłumaczenie: L. Engelking.
MKGO1961pl	Kuncewiczowa, M. (1961). *Gaj oliwny.* Warszawa: PAX.
MKN1993cz	Kundera, M. (1993). *Nesmrtelnost.* Brno: Atlantis.
MKN2008pl	Kundera, M. (2008). *Nieśmiertelność.* Warszawa: Państwowy Instytut Wydawniczy. Tłumaczenie: M. Bieńczyk.
MKNLB1985cz	Kundera, M. (1985). *Nesnesitelná lehkost bytí.* Toronto: Sixty-Eight Publishers.
MKNLB1992pl	Kundera, M. (1992). *Nieznośna lekkość bytu.* Warszawa: Państwowy Instytut Wydawniczy. Tłumaczenie: A. Holland.
MKOH1971cz	Kuncewiczowa, M. (1971). *Olivový háj.* Praha: Lidové nakladatelství. Tłumaczenie: V. Dvořáčková.
MKVNR1997cz	Kundera, M. (1997). *Valčík na rozloučenou.* Brno: Atlantis.
MKWP2001pl	Kundera, M. (2001). *Walc pożegnalny.* Warszawa: Państwowy Instytut Wydawniczy. Tłumaczenie: P. Godlewski.
MKZ1991cz	Kundera, M. (1991). *Žert.* Brno: Atlantis.
MKZ1999pl	Kundera, M. (1999). *Żart.* Warszawa: Państwowy Instytut Wydawniczy. Tłumaczenie: E. Witwicka.
MŁR2007pl	Łoziński, M. (2007). *Reisefieber.* Praha: Dybbuk. Tłumaczenie: B. Gregorová.
MŁR2008cz	Łoziński, M. (2008). *Reisefieber.* Kraków: Znak.
MVPDK2005pl	Viewegh, M. (2005). *Powieść dla kobiet.* Poznań: Zysk i S-ka. Tłumaczenie: J. Boratyńska.
MVPNK2003cz	Viewegh, M. (2003). *Případ nevěrné Kláry.* Brno: Petrov.
MVRPZ2001cz	Viewegh, M. (2001). *Román pro ženy.* Brno: Petrov.
MVSNK2007pl	Viewegh, M. (2007). *Sprawa niewiernej Klary.* Poznań: Zysk i S-ka. Tłumaczenie: M. Lemańczyk.
MVUZ2001cz	Viewegh, M. (2001). *Účastníci zájezdu.* Brno: Petrov.
MVVDVC1994cz	Viewegh, M. (1994). *Výchova dívek v Čechách.* Praha: Český spisovatel.
MVW2008pl	Viewegh, M. (2008). *Wycieczkowicze.* Warszawa: Agave. Tłumaczenie: J. Illg.
MVWDWC2005pl	Viewegh, M. (2005). *Wychowanie dziewczat w Czechach.* Warszawa: Świat Literacki. Tłumaczenie: J. Stachowski.
MVZOL2003cz	Viewegh, M. (2003). *Zapisovatelé otcovský lásky.* Brno: Petrov.
MVZOM2007pl	Viewegh, M. (2007). *Zapisywacze ojcowskiej miłości.* Poznań: Zysk i S-ka. Tłumaczenie: J. Derdowaska.
MWC2007cz	Witkowski, M. (2007). *Chlípnice.* Praha: Agite/Fra. Tłumaczenie: J. Jeništa.

MWL2005pl	Witkowski, M. (2005). *Lubiewo.* Kraków: Korporacja Ha!art.
OTB2007pl	Tokarczuk, O. (2007). *Bieguni.* Karków: Wydawnictwo Literackie.
OTDD1998pl	Tokarczuk, O. (1998). *Dom dzienny, dom nocny.* Wałbrzych: Wydawnictwo Ruta.
OTDD2002cz	Tokarczuk, O. (2002). *Denní dům, noční dům.* Brno: Host. Tłumaczenie: P. Vidlák.
OTPAJC1999cz	Tokarczuk, O. (1999). *Pravěk a jiné časy.* Brno: Host. Tłumaczenie: P. Vidlák.
PHCCG2007pl	Hůlová, P. (2007). *Czas Czerwonych Gór.* Warszawa: Wydawnictwo W.A.B. Tłumaczenie: D. Dobrew.
PHPMP2002cz	Hůlová, P. (2002). *Paměť mojí babičce.* Praha: Torst.
RDMM2009cz	Denemarková, R. (2009). Mladý muž 1989 aneb pomohu revoluci osobním vydrátkováním parket. W: *Mein 1989*, Praha: Goethe Institut. Pozyskano z http://www.goethe.de.
RDMM2009pl	Denemarková, R. (2009). Młody mężczyzna 1989 albo Wesprę rewolucję własnoręcznym cyklinowaniem parkietów. W: *Mein 1989*, Praha: Goethe Institut. Pozyskano z http://www.goethe.de. Tłumaczenie: M. Kalinowska.
SCHH1997pl	Chwin, S. (1997). *Hanemann.* Brno: Host. Tłumaczenie: P. Vidlák.
SCHH2005cz	Chwin, S. (2005). *Hanemann.* Gdańsk: Słowo Obraz Terytoria.
SLC2002pl	Lem, S. (2002). *Cyberiada.* Kraków: Wydawnictwo Literackie.
SLGP1984pl	Lem, S. (1984). *Głos Pana.* Kraków: Wydawnictwo Literackie.
SLK1983cz	Lem, S. (1983). *Kyberiáda.* Praha: Československý spisovatel. Tłumaczenie: F. Jungwirth.
SLPH1981cz	Lem, S. (1981). *Pánův hlas.* Praha: Svoboda. Tłumaczenie: F. Jungwirth i P. Weigel.
TNABK1984cz	Nowak, T. (1984). *...až budeš králem, až budeš katem...* Praha: Naše vojsko. Tłumaczenie: J. Vlášek.
TNAJK1968pl	Nowak, T. (1968). *A jak królem a jak katem będziesz.* Warszawa: Ludowa Spółdzielnia Wydawnicza.
VPTSZ1968pl	Páral, V. (1968). *Targowisko spełnionych życzeń.* Warszawa: Czytelnik. Tłumaczenie: E. Madany.
VPVSP2001cz	Páral, V. (2001). *Veletrh splněných přání.* Praha – Litomyšl: Paseka.
VRK1954pl	Řezáč, V. (1954). *Krawędź.* Warszawa: Czytelnik. Tłumaczenie: M. Erhartowa.
VRR1944cz	Řezáč, V. (1944). *Rozhraní.* Praha: František Borový.
WGD1988pl	Gombrowicz, W. (1988). *Dziennik 1953–1956.* Kraków: Wydawnictwo Literackie.
WGD1994cz	Gombrowicz, W. (1994). *Deník.* Praha: Torst. Tłumaczenie: H. Stachová.
WGTA2000pl	Gombrowicz, W. (2000). *Trans-Atlantyk.* Kraków: Wydawnictwo Literackie.

WGTA2007cz	Gombrowicz, W. (2007). *Trans-Atlantik.* Praha: Společnost pro Revolver Revue. Tłumaczenie: H. Stachová.
WKG2003pl	Kuczok, W. (2003). *Gnój.* Warszawa: Wydawnictwo W.A.B.
WKS2009cz	Kuczok, W. (2009). *Smrad.* Praha: Havran. Tłumaczenie: B. Gregorová.
WRZO1957pl	Reymont, W. (1957). *Ziemia obiecana.* Kraków.
WRZZ1967cz	Reymont, W. (1967). *Zaslíbená země.* Praha: Svoboda. Tłumaczenie: B. Křemenák.
WTCR1981cz	Terlecki, W. (1981). *Černý román.* Praha: Odeon. Tłumaczenie: V. Dvořáčková.
WTCR1981pl	Terlecki, W. (1981). *Czarny romans.* Warszawa: Czytelnik.
WTOPB1975pl	Terlecki, W. (1975). *Odpocznij po biegu.* Warszawa: PAX.
WTOPB1981cz	Terlecki, W. (1981). *Odpočni po běhu.* Praha: Odeon. Tłumaczenie: V. Dvořáčková.

Zapis obrad Parlamentu Europejskiego (InterCorp)

EPcz	*European Parliament Proceedings Parallel Corpus 1996–2011: Czech.* Pozyskano z http://www.statmt.org/europarl/.
EPpl	*European Parliament Proceedings Parallel Corpus 1996–2011: Polish.* Pozyskano z http://www.statmt.org/europarl/.

ParaSol

SLPZWW1961pl	Lem, S. (1961). *Pamiętnik znaleziony w wannie.* Kraków: Wydawnictwo Literackie.
SLDNVV1999cz	Lem, S. (1999). *Deník nalezený ve vaně.* Praha: Baronet. Tłumaczenie: P. Weigel.

Bibliografia

Adamec, P. (1973). Tři roviny modálnosti a jejich vztah k aktuálnímu členění. W: *Otázky slovanské syntaxe III. Sborník symposia „Modální výstavba výpovědi v slovanských jazycích"* (141–147). Brno: Universita J.E. Purkyně.

Ágel, V. (2000). *Valenztheorie*. Tübingen: Gunter Narr Verlag.

Ajdukiewicz, K. (1935). Die syntaktische Konnexität, *Studia Philosophica*, 1, 1–27.

Ajdukiewicz, K. (1967). Syntactic Connexion. W: S. McCall (red.), *Polish Logic 1920–1939* (207–231), Oxford: Oxford University Press.

Altenberg, B. (1999). Adverbial connectors in English and Swedish: Semantic and lexical correspondences. W: H. Hasselgård i S. Oksefjell (red.), *Out of Corpora. Studies in Honour of Stig Johansson* (249–268). Amsterdam: Rodopi.

Altenberg, B. i Granger, S. (red.). (2002). *Lexis in Contrast: Corpus-based Approaches*. Amsterdam: John Benjamins Publishing.

Anderman, G. M. i Rogers, M. (red.). (2008). *Incorporating Corpora: The Linguist and the Translator*. Clevendon: Multilingual Matters.

Baker, C.F., Fillmore, C.J. i Lowe, J.B. (1998). The Berkeley FrameNet Project. W: *Proceedings of the 17th International Conference on Computational Linguistics – COLING 1998* (Tom 1, 86–90). Montreal, Quebec, Canada: Association for Computational Linguistics.

Baker, M. (1992). *In Other Words: A Coursebook on Translation*. London & New York: Routledge.

Baker, P. (2010). *Sociolinguistics and corpus linguistics*. Edinburgh: Edinburgh University Press.

Balcerzan, E. (1998). Poetyka przekładu artystycznego. W: E. Balcerzan (red.), *Literatura z literatury (strategie tłumaczy)* (17–31). Katowice: Wydawnictwo Naukowe Śląsk.

Baños, R., Bruti, S. i Zanotti, S. (2013). Corpus linguistics and Audiovisual Translation: in search of an integrated approach, *Perspectives: Studies in Translatology*, 21(4), 483–490.

Bańczyk, Ł. (2010). Polské konstrukce s neosobními tvary slovesa na -no/-to a jejich české překladové ekvivalenty – analýza úrovně ekvivalence. W: F. Čermák i J. Kocek (red.), *Mnohojazyčný korpus InterCorp: Možnosti studia* (204–212). Praha: Nakladatelství Lidové noviny.

Bańczyk, Ł., Dybalska, R. i Vavřín, M. (2017). *Korpus InterCorp – polština*, wersja 10, 01.12.2017. Praha: Ústav Českého národního korpusu Filozofické fakulty Univerzity Karlovy.

Bańko, M. (2008). O tzw. prozodii semantycznej i jej opisie w słownikach. W: P. Żmigrodzki i R. Przybylska (red.), *Nowe studia leksykograficzne* (2) (151–161). Kraków.

Barańczak, S. (2004). *Ocalone w tłumaczeniu: szkice o warsztacie tłumacza poezji z dodatkiem małej antologii przekładów-problemów* (wyd. 3). Kraków: Wydawnictwo a5.

Bauer, J. i Grepl, M. (1970). *Skladba spisovné čestiny*. Praha: Státní pedagogické nakladatelství.

Belletti, A. i Rizzi, L. (1988). Psych-Verbs and θ-Theory, *Natural Language & Linguistic Theory*, 6(3), 291–352.

Benešová, V. i Bojar, O. (2006). Czech verbs of communication and the extraction of their frames. W: *Text, Speech and Dialogue: 9th International Conference TSD 2006* (29–36). Berlin: Springer.

Benko, V. (2014). Compatible sketch grammars for comparable corpora. W: A. Abel, C. Vettori i N. Ralli (red.), *Proceedings of the 16th EURALEX International Congress* (417–430). Bolzano: EURAC research.

Będkowska-Kopczyk, A. (2014). Verbs of emotions with *se* in Slovene: between middle and reflexive semantics. A cognitive analysis, *Cognitive Studies | Études cognitives*, 14, 203–218.

Biały, A. (2005). *Polish Psychological Verbs at the Lexicon-Syntax Interface in Cross- linguistic Perspective*. Frankfurt am Main: Peter Lang.

Bierwiaczonek, B. (2002). *A Cognitive Study of the Concept of Love in English*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Blakemore D. (1999). Evidence and modality. W: K. Brown i J. Miller (red.), *Encyclopedia of Grammatical Categories* (141–145), Amsterdam: Elsevier.

Bogucki, Ł. (2009). *Tłumaczenie wspomagane komputerowo*. Warszawa: Wydawnictwo Naukowe PWN.

Bogusławski, A. (1978). Towards an Operational Grammar, *Studia Semiotyczne*, VII, 29–90.

Bojar, O. (2003). Towards automatic extraction of verb frames, *Prague Bulletin of Mathematical Linguistics*, 79–80, 101–120.

Bojar, O. (2009). *Exploiting Linguistic Data in Machine Translation*. Praha: Ústav formální a aplikované lingvistiky Matematicko-fyzikální fakulty Univerzity Karlovy.

Bojar, O. (2012). *Čeština a strojový překlad. Strojový překlad našincům, našinci strojovému překladu.* Studies in computational and theoretical linguistics. Praha: Ústav formální a aplikované lingvistiky Matematicko-fyzikální fakulty Univerzity Karlovy.

Bojar, O. i Prokopová, M. (2006). Czech-English word alignment. W: N. Calzolari i in. (red.), *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC 2006)* (1236–1239). Genoa, Italy: European Language Resources Association (ELRA).

Boniecka, B. (1976). O pojęciu modalności (przegląd problemów badawczych), *Język Polski* LVI(2).

Bowker, L. (2002). *Computer-aided Translation Technology: A Practical Introduction*. Ottawa: University of Ottawa Press.

Brankačkec, K. (2017). Valenční vazby ekvivalentů německého „sich (an etw. gewöhnen)" v korpusech Diakorp, Hotko a Dotko (ČNK), *Prace Filologiczne*, LXX, 105–127.

Bratman, M.E. (1987). *Intentions, Plans, and Practical Reason*. Massachusetts: Harvard University Press.

Bresnan, J. (2001). *Lexical-Functional Syntax*. Oxford: Blackwell.

Brownlee, J. (2016). *Supervised and Unsupervised Machine Learning Algorithms*. Pozyskano z: https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/.

Bühler, K. (1934). *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Jena: Gustav Fischer Verlag.

Bühler, K. (2004). *Teoria języka*. Warszawa: Universitas.

Buttler, D. (1976). *Innowacje składniowe współczesnej polszczyzny*. Warszawa: Wydawnictwo Naukowe PWN.

Čermák, F. i Kocek, J. (red.). (2010). *Mnohojazyčný korpus InterCorp: Možnosti studia.* Praha: Nakladatelství Lidové noviny.

Čermák, F. i Rosen, A. (2012). The case of InterCorp, a multilingual parallel corpus, *International Journal of Corpus Linguistics*, 13(3), 411–427.

Čermák, F., Corness, P. i Klégr, A. (red.). (2010). *InterCorp: Exploring a Multi-lingual Corpus*. Praha.

Čermák, P. i Nádvorníková, O. (2015). *Románské jazyky a čeština ve světle paralelních korpusů*. Praha: Karolinum.

Čermáková, A. i Kopřivová, M. (2018). Korpusový výzkum mluveného jazyka na příkladu češtiny a angličtiny: současný stav, *Slovo a slovesnost*, 79(3), 217–240.

Čermáková, A. (2009). *Valence českých substantiv*. Praha: Nakladatelství Lidové.

Čermáková, A., Chlumská, L. i Malá, M. (red.). (2016). *Jazykové paralely*. Praha: Nakladatelství Lidové noviny.

Charciarek, A. (2010). *Polskie wyrażenia metatekstowe o funkcji fatycznej i ich odpowiedniki czeskie i rosyjskie*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Charciarek, A. (2014). Tendencje rozwojowe we współczesnym polskim i czeskim systemie adresatywnym. W: A. Piotrowicz, M. Witaszek-Samborska i K. Skibski (red.), *Kultura Komunikacji Językowej 3. Kultura Języka w Komunikacji Zawodowej – Poznańskie Spotkania Językoznawcze*, 28 (7–22). Poznań: Instytut Filologii Polskiej, Uniwersytet im. Adama Mickiewicza.

Charciarek, A. (2018). Korpus równoległy InterCorp w leksykografii przekładowej polsko-rosyjskiej. W: A. Pstyga, T. Kananowicz i M. Buchowska (red.), *Frazeologia z perspektywy językoznawcy i tłumacza – Słowo z perspektywy językoznawcy i tłumacza* VII (54–66). Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego.

Charciarek, A. (2018). Možnosti využití korpusu InterCorp v česko-polské překladové lexikografii, *Časopis pro moderní filologii*, 100(2), 206–222.

Chesterman, A. (1998). *Contrastive Functional Analysis*. Amsterdam: John Benjamins Publishing Company.

Chesterman, A. (2007). Similarity analysis and the translation profile. *Belgian Journal of Linguistics* (W. Vandeweghe, S. Vandepitte i M. van de Velde, red.), 21(1), 53–66.

Chlumská, L. (2017). *Překladová čeština a její charakteristiky*. Praha: Nakladatelství Lidové noviny.

Cieślikowa, A. (1996). Jak „ocalić" w tłumaczeniu nazwy własne? W: M. Filipowicz-Rudek i J. Konieczna-Twardzikowa (red.), *Między oryginałem a przekładem* II (311–320), Kraków: Universitas.

Cieślikowa, A. (1998). Nazwy własne w przekładzie literackim. W: E. Rzetelska-Feleszko (red.), *Polskie nazwy własne. Encyklopedia* (392). Warszawa–Kraków: Instytut Języka Polskiego Polskiej Akademii Nauk.

Cosma, R., Cristea, D., Kupietz, M., Tufiş, D. i Witt, A. (2016). DRuKoLA – Towards Contrastive German-Romanian Research based on Comparable Corpora. W: N. Calzolari i in. (red.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016), Portorož, Slovenia* (28–32). Paris: European Language Resources Association (ELRA).

Croft, W. (1993). Case Marking and the Semantics of Mental Verbs. W: J. Pustejovsky (red.), *Semantics and the Lexicon* (55–72). Dordrecht: Springer.

Daneš, F. (1971). Větné členy obligatorní, potenciální a fakultativní. W: M. Komárek (red.) *Miscellanea Linguistica* (131–138). Ostrava: Profil.

Daneš, F. i Hlavsa, Z. (1987). *Větné vzorce v češtině*. Praha: Academia.

Daneš, F., Hlavsa, Z. i Kořenský, J. (1973). Postavení slovesa v struktuře české věty. W: B. Havránek (red.) *Československé přednášky pro VII. Mezinárodní sjezd slavistů ve Varšavě* (129–139). Praha: Academia.

Dąmbska-Prokop, U. (2000). *Mała encyklopedia przekładoznawstwa*. Częstochowa: Wydawnictwo Wyższej Szkoły Języków Obcych i Ekonomii Educator.

De Baets, P., Vandevoorde, L., De Sutter, G. (2018). The Usefulness Of Parallel Corpora For Contrastive Linguistics. A Behavioral-Profile Analysis Of The Semantic Stability Hypothesis, conference. Referat wygłoszony na konferencji OLINCO, Olomouc.

de Groot, A. W. (1949). *Structurele Syntaxis*. Den Haag: Servire.

Deckert, M. (red.). (2018). *Audiovisual Translation – Research and Use*. Bern, Switzerland: Peter Lang.

Demski, T. (2011). *Od pojedynczych drzew do losowego lasu.* Pozyskano z: https://www.statsoft.pl/Czytelnia/Data-mining/.

Dębski, A. (1982). Semantyczna walencja czasownika w aspekcie konfrontatywnym, *Biuletyn Polskiego Towarzystwa Językoznawczego*, XXXIX, 79–90.

Di Pietro, R.J. (1971). *Language Structures in Contrast*. Rowley, Massachusetts: Newbury House.

Didakowski, J. i Geyken, A. (2013). From DWDS corpora to a German Word Profile – methodological problems and solutions. W: A. Abel i L. Lemnitzer (red.) *Network Strategies, Access Structures and Automatic Extraction of Lexicographical Information OPAL – Online publizierte Arbeiten zur Linguistik* X (43–52). Mannheim: Institut für Deutsche Sprache.

Dobrotová, I. (2005). Netykieta i dyskusje polityczne w czeskim i polskim internecie, *Bohemistyka*, 5(3), 199–207.

Dobrotová, I. i Chlebda, W. (2015). Trojjazyčný tematický frazeologický slovník: koncepce, metodologické předpoklady, realizace. W: *Tematický česko-polsko-ruský a polsko-česko-ruský slovník pohraničí*. Opole: Uniwersytet Opolski.

Dubois, J. (red.). (1973). *Dictionnaire de Linguistique*. Paris: Larousse.

Dyvik, H. (1998). Translations as semantic mirrors. W: *Proceedings of Workshop W*, 13, (24–44).

Dyvik, H. (2002). Translations as semantic mirrors: from parallel corpora to wordnet. W: K. Aijmer i B. Altenberg (red.), *23rd International Conference on English Language Research on Computerized Corpora of Modern and Medieval English (ICAME 23)*, (308–326). Pozyskano z: https://epdf.tips/queue/advances-in-corpus-linguistics-papers-from-the-23rd-international-conference-on-.html

Dyvik, H. (2004). Translations as semantic mirrors: from parallel corpus to wordnet. W: *Advances in Corpus Linguistics* (309–326). Brill Rodopi. Pozyskano z: https://doi.org/10.1163/9789004333710_019

Dyvik, H. (2005). Translations as a semantic knowledge source. W: *Proceedings of the second baltic conference on human language technologies* (27–38). Tallinn: Tallinn University of Technology & Institute of the Estonian Language.

Dyvik, H., Meurer, P., Rosén, V. i De Smedt, K. (2009). Linguistically Motivated Parallel Parsebanks. W: M. Passarotti, A. Przepiórkowski, S. Raynaud i F. Van Eynde (red.), *Proceedings of the Eighth International Workshop on Treebanks and Linguistic Theories* (71–82). Milano: EDUCatt.

Dziob, A. i Piasecki, M. (2018). Implementation of the Verb Model in plWordNet 4.0. W: F. Bond, T. Kuribayashi, C. Fellbaum i P. Vossen (red.), *Proceedings of the 9th Global Wordnet Conference* (114–123). Singapore: The Global WordNet Association.

Dziwirek, K. i Lewandowska-Tomaszczyk, B. (2009). Love and Hate: Unique transitive emotions in Polish and English. W: K. Dziwirek i B. Lewandowska-Tomaszczyk (red.), *Studies in Cognitive Corpus Linguistics* (297–317). Frankfurt am Main: Peter Lang.

Dziwirek, K. i Lewandowska-Tomaszczyk, B. (2010). *Complex Emotions and Grammatical Mismatches: A Contrastive Corpus-Based Study*. Berlin: De Gruyter Mouton.

Ebeling, J. i Ebeling, S. O. (2013). *Patterns in Contrast*. Amsterdam: John Benjamins.

Ebeling, S.O. (2013). Semantic prosody in a cross-linguistic perspective. W: M. Huber i J. Mukherjee (red.), *Corpus Linguistics and Variation in English: Focus on Non-Native Englishes – Studies in Variation, Contacts and Change in English 13*. Helsinki:

Research Unit for Variation, Contacts, and Change in English. Pozyskano z: http://www.helsinki.fi/varieng/series/volumes/13/ebeling/.

Eldén, L., Merkel, M., Ahrenberg, L. i Fagerlund, M. (2013). Computing Semantic Clusters by Semantic Mirroring and Spectral Graph Partitioning, *Mathematics in Computer Science*, 7 (3), 293–313.

Ellis, J. (1966). *Towards a general Comparative Linguistics*. The Hague: Mounton.

Engelberg, S. (2018). The argument structure of psych-verbs: A quantitative corpus study on cognitive entrenchment. W: H.C. Boas i A. Ziem (red.), *Constructional Approaches to Syntactic Structures in German* (47–84). Berlin–Boston: Walter de Gruyter.

Fellbaum, C. (red.). (1998). *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.

Filipec, J., Daneš, F., Machač, J. i Mejstřík, V. (red.). (1994). *Slovník spisovné češtiny pro školu a veřejnost* (wyd. 2). Praha: Academia.

Fillmore, C.J. (1968). The Case for Case. W: Bach, E. i Harms, R.T. (red.), *Universals in Linguistic Theory* (1–88). New York: Holt, Rinehart and Winston.

Fillmore, C.J. (1977). The case for case reopened. W: P. Cole (red.), *Syntax and Semantics 8: Grammatical Relations* (59–81). New York: Academic Press.

Fišer, D. i Ljubešić, N. (2013). Best friends or just faking it? Corpus-based extraction of Slovene-Croatian translation equivalents and false friends, *Slovenščina 2.0: empirical, applied and interdisciplinary research*, 1(1), 50–77.

Fisiak, J. (red.). (1980). *Theoretical Issues in Contrastive Linguistics*. John Benjamins Publishing Company.

Fisiak, J. (red.). (1984). *Contrastive linguistics. Prospects and problems*. Berlin: Mouton de Gruyter.

Francis, G., Hunston, S. i Manning, E. (1996). *Collins COBUILD Grammar Patterns 1: Verbs*. London: Harper Collins.

Gast, V. (2012A). Contrastive analysis. W: M. Byram i A. Hu (red.), *Routledge Encyclopedia of Language Teaching and Language Learning*. London: Routledge Taylor & Francis Group.

Gast, V. (2012B). Contrastive linguistics: Theories and methods. W: B. Kortmann i J. Kabatek (red.), *Dictionaries of Linguistics and Communication Science: Linguistic theory and methodology*. Berlin: Mouton de Gruyter.

Głąbska, M. (2014). *Obraz miłości we współczesnym języku polskim (na podstawie analizy znaczeń leksemów kochać i miłość)*. Warszawa: Uniwersytet Warszawski.

Gołąb, Z. (1967). Próba syntaktycznej klasyfikacji czasowników polskich (na zasadzie konotacji), *Biuletyn Polskiego Towarzystwa Językoznawczego*, XXV, 3–43.

Gottlieb, H. (1998). Subtitling. W: M. Baker (red.), *Routledge Encyclopedia of Translation Studies* (244–248). London–New York: Routledge.

Górski, R.L. (2012). Zastosowanie korpusów w badaniu gramatyki. W: A. Przepiórkowski, M. Bańko, R.L. Górski i B. Lewandowska-Tomaszczyk (red.), *Narodowy Korpus Języka Polskiego* (291–300). Warszawa: Wydawnictwo Naukowe PWN.

Górski, R.L. i Łaziński, M. (2012). Reprezentatywność i zrównoważenie korpusu. W: A. Przepiórkowski, M. Bańko, R.L. Górski i B. Lewandowska-Tomaszczyk (red.), *Narodowy Korpus Języka Polskiego* (25–36). Warszawa: Wydawnictwo Naukowe PWN.

Grabowski, Ł. (2012). *A Corpus-Driven Study of Translational and Non-Translational Texts: the Case of Nabokov's 'Lolita'*. Opole: Wydawnictwo Uniwersytetu Opolskiego.

Grabowski, Ł. (2015). O frazeologii z perspektywy językoznawstwa korpusowego. Przegląd głównych nurtów badawczych ostatniego dwudziestolecia w Wielkiej Brytanii i USA, *Problemy frazeologii europejskiej*, X, 23–48.

Grabowski, Ł. i Hebal-Jezierska, M. (2016). O różnych korpusowych metodach badawczych – próba krytycznej reflekcji, *Komunikacja Specjalistyczna*, 11, 65–83.

Granger, S. (2003). The corpus approach: A common way forward for Contrastive Linguistics and Translation Studies? W: S. Granger, J. Lerot i S. Petch-Tyson (red.), *Corpus-based approaches to Contrastive Linguistics and Translation Studies* (17–30). Amsterdam: Rodopi.

Granger, S. (2010). Comparable and translation corpora in cross-linguistic research Design, analysis and applications, *Journal of Shanghai Jiastong University*, 2, 14–21.

Greń, Z. (1992). Z problematyki konfrontatywnego badania czasownika w językach blisko spokrewnionych (na materiale polskim i czeskim). W: *Z polskich studiów slawistycznych*, VIII (63–68).

Greń, Z. (1994). *Semantyka i składnia czasowników oznaczających akty mowy w języku polskim i czeskim*. Warszawa: Slawistyczny Ośrodek Wydawniczy.

Greń, Z. (2003). Czeskie czasowniki *vzkázat : vyřídit* i ich polskie ekwiwalenty. W: W. Banyś, L. Bednarczuk i K. Polański (red.), *Etudes linguistiques romano-slaves offertes a Stanisław Karolak* (187–193). Kraków: Oficyna Wydawnicza Edukacja.

Greń, Z. (2015). Wariantywność formalna w procesie adaptacji anglicyzmów w języku słowackim, Semantyka a konfrontacja językowa (D. Roszko i J. Satoła-Staśkowiak, red.), 5, 125–137.

Greń, Z. (2016). Pokus o zavedení valenční teorie v Polsku. W: K. Skwarska i E. Kaczmarska (red.), *Výzkum slovesné valence ve slovanských zemích* (27–36). Praha: Slovanský ústav AV ČR.

Greń, Z. i Rytel-Kuc, D. (1991). Wykorzystanie przekładów literackich w pracy nad dwujęzycznym słownikiem walencyjnym, *Prace Slawistyczne. Problemy teoretyczno-metodologiczne badań konfrontatywnych języków słowiańskich* (H. Běličova, G. Nieszczimienko i Z. Rudnik-Karwatowa, red.), 89, 69–78.

Gries, S.T. (red.). (2007). *Corpora in Cognitive Linguistics. Corpus-Based Approaches to Syntax and Lexis*. Berlin–Boston: De Gruyter Mouton.

Gruszczyńska, E. i Leńko-Szymańska, A. (red.). (2016). *Polskojęzyczne korpusy równoległe. Polish-language Parallel Corpora*. Warszawa: Instytut Lingwistyki Stosowanej Uniwersytetu Warszawskiego.

Guentchéva, Z. (red.). (2018). *Epistemic Modalities and Evidentiality in Cross-Linguistic Perspective*. Berlin/Boston: Walter de Gruyter.

Hajnicz, E. i Nitoń, B. (2017). *Instrukcja dostępu do słownika walencyjnego WALENTY za pośrednictwem programu Slowal.* Pozyskano z: http://clarin-pl.eu/wp-content/uploads/2017/05/instrukcja_uzytkownika_Walentego.pdf

Hajnicz, E., Patejuk, A., Przepiórkowski, A. i Woliński, M. (2016). Walenty – słownik walencyjny języka polskiego z bogatym komponentem frazeologicznym. W: K. Skwarska i E. Kaczmarska (red.), *Výzkum slovesné valence ve slovanských zemích* (71–102). Praha: Slovanský ústav Akademie věd České republiky.

Halliday, M.A. (1985). *An Introduction to Functional Grammar*. London: Arnold.

Han, A.L.-F., Lu, Y., Wong, D.F., Chao, L.S., He, L. i Xing, J. (2013). Quality Estimation for Machine Translation Using the Joint Method of Evaluation Criteria and Statistical Modeling. W: O. Bojar i in. (red.), *Proceedings of the Eighth Workshop on Statistical Machine Translation* (365–372). Sofia, Bulgaria: The Association for Computational Linguistics.

Harkins, J. i Wierzbicka, A. (red.). (2001). *Emotions in crosslinguistic perspective*. Berlin: Mouton de Gruyter.

Havránek, B., Bělič, J., Helcl, M., Jedlička, A., Křístek, V. i Trávníček, F. (red.). (1989). *Slovník spisovného jazyka českého* (wyd. 2). Praha: Academia.

Hebal-Jezierska, M. (2008). *Wariantywność końcówek fleksyjnych rzeczowników męskich żywotnych w języku czeskim.* Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.

Hebal-Jezierska, M. (2010). Jak se překládají české univerbizáty do polštiny. W: F. Čermák i J. Kocek (red.), *Mnohojazyčný korpus InterCorp: Možnosti studia* (261–268). Praha: Nakladatelství Lidové noviny.

Hebal-Jezierska, M. (2013). Podstawowe zasady korzystania z korpusów przy badaniu języka. W: W. Chlebda (red.), *Tropem korpusów. W poszukiwaniu optymalnych zbiorów tekstów* (17–30). Opole: Uniwersytet Opolski.

Hebal-Jezierska, M. (2014). Praktyczny przewodnik po korpusie języka czeskiego. W: M. Hebal-Jezierska (red.), *Praktyczny przewodnik po korpusach języków słowiańskich* (29–54). Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.

Hebal-Jezierska, M., Kaczmarska, E. i Rosen, A. (2016). Between the devil and the deep blue sea or between users' needs and the compilers' powers: An analysis of the Czech-Polish part of the parallel corpus InterCorp. W: E. Gruszczyńska i A. Leńko-Szymańska, *Polskojęzyczne korpusy równoległe / Polish-language Parallel Corpora* (41–65). Warszawa: Instytut Lingwistyki Stosowanej Uniwersytetu Warszawskiego.

Hejwowski, K. (2009). *Kognitywno-komunikacyjna teoria przekładu*. Warszawa: Wydawnictwo Naukowe PWN.

Hirschová, M. (2017). Evidenciálnost. In: P. Karlík, M. Nekula, J. Pleskalová (red.), *CzechEncy – Nový encyklopedický slovník češtiny*. Praha: Nakladatelství Lidové noviny. Pozyskano z: https://www.czechency.org/slovnik/EVIDENCIÁLNOST.

Hladká, B., Holub, M. i Kríž, V. (2013). Feature engineering in the NLI Shared Task 2013: Charles University submission report. W: *Proceedings of the Eighth Workshop*

on Innovative Use of NLP for Building Educational Applications (232–241). Atlanta, Georgia: The Association for Computational Linguistics.

Hockett, C.F. (1958). *A Course in Modern Linguistics*. New York: Macmillan.

Holvoet, A. (2011). O leksykalnych wykładnikach użycia interpretatywnego, *Linguistica Copernicana*, 5(1), 77–91. DOI: 10.12775/LinCop.2011.005.

Homola, P. (2009). *Syntactic Analysis in Machine Translation*. Praha: Ústav formální a aplikované lingvistiky Matematicko-fyzikální fakulty Univerzity Karlovy.

House, J. (1977). *A Model for Translation Quality Assessment*. Tübingen: TBL Verlag Gunter Narr.

House, J. (1997). *Translation Quality Assessment: A Model Revisited*. Tübingen: TBL Verlag Gunter Narr.

House, J. (2001). Translation Quality Assessment: Linguistic Description vs Social Evaluation, *META – journal des traducteurs / Translators' Journal*, 2(46), 243–257.

Hunston, S. (2007). Semantic prosody revisited, *International Journal of Corpus Linguistics*, 12(2), 249–268.

Hunston, S. i Francis, G. (2000). *Pattern Grammar: A Corpus-driven Approach to the Lexical Grammar of English*. Amsterdam: John Benjamins.

Ivanová, M. (2011). Epistemické funkcie evidenčných operátorov v slovenčine. W: M. Ološtiak, M. Ivanová i D. Slančová (red.), *Vidy jazyka a jazykovedy. Na počesť Miloslavy Sokolovej* (145–154). Prešov: Filozofická fakulta Prešovskej univerzity.

James, C. (1980). *Contrastive Analysis*. London: Longman.

Janz, A., Kędzia, P. i Kaszewski, D. (2018). *Word Sense Disambiguation tool WoSeDon*. CLARIN-PL digital repository, http://hdl.handle.net/11321/540.

Jaszczolt, K.M. (2011). Contrastive analysis. W: J.-O. Östman i J. Verschueren (red.), *Pragmatics in Practice* (111–117). Amsterdam: John Benjamins Publishing Company.

Jelínek, T. (2014). Improvements to dependency parsing using automatic simplification of data. W: N. Calzolari i in. (red.), *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)* (73–77). Reykjavik, Iceland: European Language Resources Association (ELRA).

Jirásek, K. (2011). Využití paralelního korpusu InterCorp k získávání ekvivalentů pro chorvatsko-český slovník. W: Čermák, F. (red.), Korpusová lingvistika Praha 2011: 1 – InterCorp (45–55). Nakladatelství Lidové noviny, Praha.

Johansson, S. (1998). On the role of corpora in cross-linguistic research. W: S. Johansson i S. Oksefjell (red.), *Corpora and cross-linguistic research: Theory, method, and case studies* (3–24). Amsterdam–Atlanta: Rodopi.

Johansson, S. (2006). How well can WELL be translated? On the English discourse particle WELL and its correspondences in Norwegian and German. W: K. Aijmer i A.-M. Simon-Vandenbergen (red.), *Pragmatic Markers in Contras*t (115–137). Amsterdam: Elsevier.

Johansson, S. (2007A). Seeing through Multilingual Corpora. *On the use of corpora in contrastive studies*. Amsterdam: John Benjamins Publishing Company.

Johansson, S. (2007B). The English verb *seem* and its correspondences in Norwegian. W: S. Johansson, *Seeing through multilingual corpora: On the use of corpora in contrastive studies* (117–138). Amsterdam: John Benjamin Publishing Company.

Johansson, S. (2009). Which way? On English *way* and its translations. *International Journal of Translation*, 21(1–2), 15–40.

Jurko, P. (2017). Pragmatic meaning in contrast: semantic prosodies of Slovene and English, *Perspectives: Studies in Translation Theory and Practice*, 25(1), 157–176.

Kacnel'son, S.D. (1948). O grammatičeskoj kategorii, *Vestnik Leningradskogo Gosudarstvennogo Universiteta*, 2, 14–134.

Kaczmarska, E. (2001). Badanie struktury walencyjnej czeskich i polskich predykatów posiadających pozycję Experiencera, *Studia z Filologii Polskiej i Słowiańskiej*, 37, 177–187.

Kaczmarska, E. (2002). Nominalizacje odczasownikowe w języku polskim i czeskim (wybrane problemy), *Studia z Filologii Polskiej i Słowiańskiej*, 38, 87–99.

Kaczmarska, E. (2003). Związki frazeologiczne o tematyce marynistycznej w języku polskim na tle porównawczym innych języków słowiańskich, *Prace Filologiczne*, XLVIII, 251–257.

Kaczmarska, E. (2006A). (Nie)typowe anglicyzmy w czeskiej i polskiej prasie dla mężczyzn. W: J. Królak i J. Molas (red.), *Slavica Leguntur: Aktualne problemy badawcze slawistyki* (108–116). Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.

Kaczmarska, E. (2006B). Nominalizacje w czeskiej i polskiej prasie dla rodziców małych dzieci. W: M. Olšiak (red.) *Varia XIV – Zborník materiálov zo XIV. kolokvia mladých jazykovedcov* (345–353). Bratislava: Slovenská jazykovedná spoločnosť pri Slovenskej akadémii vied.

Kaczmarska, E. (2007). Polski, czeski i angielski Harry Potter, czyli o formacjach nominalizowanych w oryginale i przekładach. W: A. Kątny (red.), *Słowiańsko-niesłowiańskie kontakty językowe w perspektywie dia- i synchronicznej* (165–170). Olecko: Wszechnica Mazurska.

Kaczmarska, E. (2008A). Funkcja stylistyczna konstrukcji z nominalizacjami w języku polskim i czeskim. W: A. Kątny (red.), *Kontakty językowe i kulturowe w Europie. Sprach- und Kulturkontakte in Europa* (129–136). Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego.

Kaczmarska, E. (2008B). Polskie i czeskie artykuły sponsorowane. W: K. Michalewski (red.), *Język w marketingu* (75–82). Łódź: Uniwersytet Lódzki.

Kaczmarska, E. (2009A). Kilka uwag o tytułach artykułów jako skrytych komunikatach reklamowych (na podstawie polskiej i czeskiej prasy dla mężczyzn). W: I. Generowicz, E. Kaczmarska i I.M. Doliński (red.), *Świat ukryty w słowach, czyli o znaczeniu gramatycznym, leksykalnym i etymologicznym* (175–183). Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.

Kaczmarska, E. (2009B). Relacje temporalne w konstrukcjach z nominalizacjami w języku polskim i czeskim. W: J. Mindak-Zawadzka i I. M. Doliński (red.), *Językowy świat Słowian. Zjawiska, interpretacje, znaki zapytania* (115–127). Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.

Kaczmarska, E. (2010). Analiza zdolności konotacyjnych polskich i czeskich predykatów odnoszących się do strachu, złości i wstydu. W: J. Goszczyńska i Z. Greń (red.), *Res slavisticae* (135–153). Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.

Kaczmarska, E. (2012A). Czeski czasownik *zdát se* w przekładzie na język polski (na podstawie badań z wykorzystaniem czesko-polskiego korpusu równoległego InterCorp), *Studia z Filologii Polskiej i Słowiańskiej*, 47, 247–261.

Kaczmarska, E. (2012B). Searching for equivalents on the basis of Czech-Polish parallel corpus (the case of the verb *zdát se*). W: P. Karagiozov, K. Bahneva, V. Geshev, I. Hristova i M. Mladenova (red.), *Време и история в славянските езици, литератури и култури. Езикознание* (238–245). Sofia: Universitetsko Izdatelstvo Sv. Kliment Ohridski.

Kaczmarska, E. (2014A). Czy na pewno się (nie)rozumiemy? O problemach, uproszczeniach i stratach w przekładzie (na podstawie czesko-polskiej części korpusu równoległego InterCorp). W: M. Benešová, R. Rusin Dybalska i L. Zakopalová (red.), *Proměny polonistiky. Tradice a výzvy polonistických studií* (192–199). Praha: Karolinum.

Kaczmarska, E. (2014B). Czeskie czasowniki oznaczające stany psychiczne – sposoby ustalania polskich ekwiwalentów na podstawie korpusu równoległego InterCorp. W: A. Stolarczyk-Gembiak i M. Woźnicka (red.), *Zbliżenia: językoznawstwo – translatoryka – literaturoznawstwo* (45–56). Konin: Państwowa Wyższa Szkoła Zawodowa w Koninie.

Kaczmarska, E. (2015A). W poszukiwaniu znaczenia czasowników wyrażających stany psychiczne. Analiza czeskich czasowników i ich polskich ekwiwalentów – próba implementacji wybranych teorii lingwistycznych (walencja, gramatyka przypadków głębokich, Pattern Grammar, lingwistyka kognitiwna), *Prace Filologiczne*, LXVII, 131–150.

Kaczmarska, E. (2015B). Mít rád czy milovat? O czeskiej miłości po polsku. W: K. Waszakowa i M. Falkowska (red.), *Pojęcia zapisane w języku* (139–156). Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.

Kaczmarska, E. (2015C). Rekonceptualizacje w przekładzie czesko-polskim (na przykładzie czasowników wyrażających uczucia i emocje). Referat wygłoszony na konferencji „Imago Mundi", Warszawa.

Kaczmarska, E. (2016). O dwóch jednostkach leksykalnych będących wykładnikami negatywnych stanów emocjonalnych i ich polskich ekwiwalentach. Analiza na materiale z korpusu równoległego InterCorp. W: *Polskojęzyczne korpusy równoległe* (227–248). Warszawa: Wydział Lingwistyki Stosowanej Uniwersytetu Warszawskiego.

Kaczmarska, E. (2017A). Corpus-based Analysis of Czech Units Expressing Mental States and Their Polish Equivalents. Identification of Meaning and Establishing Polish Equivalents Referring to Different Theories. W: P. Pęzik i J.T. Waliński (red.),

Language, Corpora and Cognition, Łódź Studies in Language 51 (181–200). Frankfurt am Main: Peter Lang.

Kaczmarska, E. (2017B). Czeskie jednostki wyrażające współczucie oraz ich polskie ekwiwalenty. Analiza na podstawie korpusu paralelnego InterCorp, *Prace Filologiczne*, LXX, 225–249.

Kaczmarska, E. (2018). Walencja w przekładzie – stracona czy zachowana? *Prace Filologiczne* (Z. Zaron i K. Skwarska, red.), LXXII, 115–132.

Kaczmarska, E. i Kłos, Z. (2012). Grupy tematyczne zapożyczeń niemieckich w języku polskim, czeskim, słowackim i łużyckim. Problem analizy pożyczek pochodzenia niemieckiego w językach słowiańskich, *Zeszyty Łużyckie*, 46, 43–63.

Kaczmarska, E. i Rosen, A. (2013). Między znaczeniem leksykalnym a walencją – próba opracowania metody ekstrakcji ekwiwalentów na podstawie korpusu równoległego, *Studia z Filologii Polskiej i Słowiańskiej*, 48, 103–121.

Kaczmarska, E. i Rosen, A. (2014A). Czego nie można wyrazić w języku polskim, czyli o leksykalnych w nim brakach, *Polonica*, 34, 53–66.

Kaczmarska, E. i Rosen, A. (2014B). Praktyczny przewodnik po korpusie równoległym InterCorp. W: M. Hebal-Jezierska (red.), *Praktyczny przewodnik po korpusach języków słowiańskich* (207–231). Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.

Kaczmarska, E. i Rosen, A. (2016). Niedosłowności w dialogu czesko-polskim. W: M. Odelski, A. Knapik, P. Chruszczewski i W. Chłopicki (red.), *Niedosłowności w języku. Język a komunikacja*, 37 (45–57). Kraków: Tertium.

Kaczmarska, E. i Rosen, A. (2016A). Syntakticko-sémantický popis vybraných skupin sloves vyjadřujících emoce a pocity – metody kontrastivního zkoumání valence na základě paralelního korpusu. W: K. Skwarska i E. Kaczmarska (red.), *Výzkum slovesné valence ve slovanských zemích* (319–350). Praha: Slovanský ústav Akademie věd České republiky.

Kaczmarska, E. i Rosen, A. (2016B). Jak najít optimální překlad polysémních sloves – porovnání metod automatické analýzy paralelních textů, *Časopis pro moderní filologii*, 157–168.

Kaczmarska, E. i Rosen, A. (2018). When a verb does not "mean" a verb – a corpus-based analysis of Czech verbs and their non-verbal equivalents in Polish. Referat wygłoszony na konferencji OLINCO, Olomouc.

Kaczmarska, E. i Rosen, A. (2019). Analiza korpusowa czeskiego czasownika *postrádat* i jego polskich ekwiwalentów. W: D.K. Rembiszewska i M. Hebal-Jezierska (red.), *Tom jubileuszowy prof. Janusza Siatkowskiego*. W druku.

Kaczmarska, E., Rosen, A., Hana, J. i Hladká, B. (2015). Syntactico-semantic analysis of arguments as a method for establishing equivalents of Czech and Polish verbs expressing mental states, *Prace Filologiczne*, LXVII, 151–174.

Kaczmarska, E., Stefaniak, M. i Susfał, J. (2011). O „bezpiecznej" prawdzie w tabloidach polskich i czeskich, *Oblicza Komunikacji, Tabloidy – język, wartości, obraz świata*, 4, 217–229.

Kalisz, R. (1981). *The pragmatics, semantics, and syntax of the English sentences with indicative* that *complements and the Polish sentences with* że *complements. A Contrastive Study*. Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego.

Kamasa, V. (2015). Prozodia semantyczna. Pojęcie, problemy, przydatność, *Socjolingwistyka*, XXIX, 105–120.

Karlík, P., Nekula, M. i Rusínová, Z. (1996). *Příruční mluvnice češtiny* (wyd. 2). Praha: Nakladatelství Lidové noviny.

Karolak, S. (1972). *Zagadnienia składni ogólnej*. Warszawa: PWN.

Karolak, S. (1984). Składnia wyrażeń predykatywnych. W: Z. Topolińska (red.), *Gramatyka współczesnego języka polskiego. Składnia* (11–211). Warszawa: PWN.

Kasprzak, K. (2017). Próba ustalenia polskiego ekwiwalentu dla czeskiego czasownika *hrotit*. Niepublikowana praca dyplomowa.

Kettnerová, V. (2014). *Lexikálně-sémantické konverze ve valenčním slovníku*. Praha: Karolinum.

Kettnerová, V., Lopatková, M. i Bejček, E. (2012). The Syntax-Semantics Interface of Czech Verbs in the Valency Lexicon. W: R. V. Fjeld i R. M. Torjusen (red.) *Proceedings of the 15th EURALEX International Congress* (434–443). Oslo: Department of Linguistics and Scandinavian Studies, University of Oslo.

Kędzia, P., Piasecki, M. i Orlińska, M. (2016). *WoSeDon*. CLARIN-PL digital repository, http://hdl.handle.net/11321/290.

Kędzia, P., Piasecki, M., Kocoń, J. i Indyka-Piasecka, A. (2014). Distributionally Extended Network-Based Word Sense Disambiguation in Semantic Clustering of Polish Texts, *IERI Procedia – International Conference on Future Information Engineering* (FIE 2014) (G. Lee, red.), 10, 38–44. DOI: 10.1016/j.ieri.2014.09.073.

Kielar, B.Z. (1988). *Tłumaczenie i koncepcje translatoryczne*. Wrocław: Zakład Narodowy im. Ossolińskich.

Kieraś, W., Kobyliński, Ł. i Ogrodniczuk, M. (2018). Korpusomat – a tool for creating searchable morphosyntactically tagged corpora, *Computational Methods in Science and Technology*, 24(1), 21–27.

Kilgarriff, A., Baisa, V.B., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P. i Suchomel, V. (2014). The Sketch Engine: ten years on, *Lexicography*, 1, 7–36.

Klégr, A., Malá, M. i Šaldová, P. (2012). *Anglické ekvivalenty nejfrekventovanějších českých předložek*. Praha: Karolinum.

Klemensiewicz, Z. (1937). *Składnia opisowa współczesnej polszczyzny kulturalnej*. Kraków: Polska Akademii Umiejętności.

Kolberová, U. (2018). O obrazie i stereotypie rzemieślnika w polskich i czeskich mediach. W: U. Kolberová (red.), *Rzemiosło i rzemieślnik w językach i kulturze Słowian / Řemeslo a řemeslník v jazycích a kultuře Slovanů* (140–145). Ostrava: Filozofická fakulta, Ostravská univerzita.

Koller, W. (1995). The Concept of Equivalence and the Object of Translation Studies. *Target*, 7, 191–222.

Komissarov, V.N. (1973). *Slovo o perevode*. Moskva: Mezhdunarodnye otnosheniia.

Komissarov, V.N. (1980). *Lingvistika perevoda*. Moskwa: Mezhdunarodnye otnosheniia.

Kondas, A. (2016A). *Wstęp do Machine Learning*. Pozyskano z: http://itcraftsman.pl/wstep-do-machine-learning/.

Kondas, A. (2016B). *Walidacja krzyżowa – RandomSplit – PHP-ML*. Pozyskano z: http://itcraftsman.pl/walidacja-krzyzowa-randomsplit-php-ml/.

Konvalinková, R. (2011). Vyjadřování prostorových vztahů v češtině (valenčně-kognitivní analýza). Dysertacja. Brno: Filozofická fakulta Masarykovy university.

Kornacki, M. (2018). *Computer-assisted Translation (CAT) Tools in the Translator Training Process*. Bern, Switzerland: Peter Lang.

Korytkowska, M. (1984). Kategoria przypadka semantycznego (na materiale języka polskiego, bułgarskiego i serbsko-chorwackiego). W: V. Koseska-Toszewa i M. Korytkowska (red.), *Studia konfrontatywne polsko-południowosłowiańskie* (11–38). Wrocław: Zakład Narodowy im. Ossolińskich.

Korytkowska, M. (1992). *Typy pozycji predykatowo–argumentowych. Gramatyka konfrontatywna bułgarsko-polska*. Warszawa: Slawistyczny Ośrodek Wydawniczy.

Korytkowska, M. (1993). O konfrontatywnym opisie predykatorów bułgarskich i polskich (na przykładzie jednostek otwierających miejsce dla argumentu o wartości Experiencer). W: V. Koseska-Toszewa i M. Korytkowska (red.), *Studia gramatyczne bułgarsko-polskie* V–VI (121–150). Warszawa: Slawistyczny Ośrodek Wydawniczy.

Korytkowska, M. i Mazurkiewicz-Sułkowska, J. (2014). O konfrontatywnym badaniu polskich i bułgarskich czasowników mentalnych, *Rocznik Slawistyczny*, LXIII, 47–76.

Koseska-Toszewa, V. i Roszko, R. (2016). Języki słowiańskie i litewski w korpusach równoległych Clarin-PL, *Studia z Filologii Polskiej i Słowiańskiej*, 51, 191–217.

Kozłowska, Z. (2007). *O przekładzie tekstu naukowego (na przykładzie tekstów językoznawczych)*. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.

Kruszewski, G., Paperno, D. i Baroni, M. (2015). Deriving Boolean structures from distributional vectors, *Transactions of the Association for Computational Linguistics*, 3, 375–388.

Krzeszowski, T. (1990). *Contrasting Languages: The Scope of Contrastive Linguistics*. Berlin–New York: Walter de Gruyter.

Krzeszowski, T.P. (1984). Tertium comparationis. W: J. Fisiak (red.), *Contrastive Linguistics: Prospects and Problems* (301–312). Berlin: Mouton de Gruyter.

Krzeszowski, T.P. (2016). *The Translation Equivalence Delusion. Meaning and Translation*. Frankfurt am Main: Peter Lang.

Krzyżanowska, A. (2013). Nić Ariadny / Le fill d'Ariane – czyli o międzyjęzykowej ekwiwalencji mitologizmów frazeologicznych, *Rocznik Przekładoznawczy. Studia nad teorią, praktyką i dydaktyką przekładu*, 8, 35–48.

Künstler, I. (2009). Napisy dla niesłyszących – problemy i wyzwania, *Przekładaniec,* 20, 115–124.

Kyseľová, M., Ivanová, M. (2013). *Sloveso vo svetle kognitívnej gramatiky*. Prešov: Filozofická fakulta Prešovskej univerzity.

Lado, R. (1957). *Linguistics Across Cultures*. Ann Arbor: University of Michigan Press.

Langacker, R. (1987). *Foundations of Cognitive Grammar, vol. 1, Theoretical Prerequisites*. Stanford: Stanford University Press.

Langacker, R. (1991). *Foundations of Cognitive Grammar, vol. 2, Descriptive Application*. Stanford: Stanford University Press.

Langacker, R. (2008). *Cognitive Grammar: A Basic Introduction*. New York: Oxford University Press.

Lee, D.Y. (2008). Corpora and discourse analysis: New ways of doing old things. W: V.K. Bhatia, J. Flowerdew i R.H. Jones (red.), *Advances in Discourse Studies* (86–99). London: Routledge.

Legeżyńska, A. (1985). Tłumacz, czyli „drugi autor". W: E. Balcerzan i S. Wysłouch (red.), *Miejsca wspólne. Szkice o komunikacji literackiej i artystycznej* (160–181). Warszawa: Państwowe Wydawnictwo Naukowe.

Leonardi, V. (2007). *Gender and Ideology in Translation: Do Women and Men Translate Differently? A Contrastive Analysis from Italian into English*. Frankfurt am Main: Peter Lang.

Levin, B. (1993). *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.

Levý, J. (2013). *Umění překladu*. Praha: Apostrof.

Lewandowska-Tomaszczyk, B. (1984). *Conceptual analysis, linguistic meaning, and verbal interaction*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

Lewandowska-Tomaszczyk, B. (1996A). *Depth of negation – a cognitive semantic study*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

Lewandowska-Tomaszczyk, B. (1996B). Cross-linguistic and language-specific aspects of semantic prosody, *Language Sciences*, 18, 153–178.

Lewandowska-Tomaszczyk, B. (1999). A cognitive-interactional model of cross-linguistic analysis: New perspectives on 'tertium comparationis' and the concept of equivalence. W: B. Lewandowska-Tomaszczyk (red.), *Cognitive Perspectives on Language* (53–76). Frankfurt am Main: Peter Lang.

Lewandowska-Tomaszczyk, B. (red.) (2005). *Podstawy językoznawstwa korpusowego*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

Lewandowska-Tomaszczyk, B. (2010A). Nowe spojrzenie na przekład: podobieństwo, granice ekwiwalencji i rekonceptualizacja, *Lingwistyka Stosowana*, 3, 9–31.

Lewandowska-Tomaszczyk, B. (2010B). Re-conceptualization and the emergence of discourse meaning as a theory of translation. W: B. Lewandowska-Tomaszczyk i M. Thelen (red.), *Meaning in Translation* (105–147). Frankfurt am Main: Peter Lang.

Lewandowska-Tomaszczyk, B. (2012). Cognitive Corpus Studies: A New Qualitative & Quantitative Agenda for Contrasting Languages, *Connexion – A Journal of Humanities and Social Sciences*, 1(1), 26–64.

Lewandowska-Tomaszczyk, B. (2013). Komunikacja i konstruowanie znaczeń w przekładzie. Referat wygłoszony na konferencji „Zbliżenia: językoznawstwo". Konin.

Lewandowska-Tomaszczyk, B. i Pęzik, P. (2018). Parallel and comparable language corpora, cluster equivalence and translator education. W: *Society and Languages in the Third Millenium – Communication.Education.Translation* (131–142). Moscow: Peoples' Friendship University of Russia (RUDN University).

Lewicki, R. (1993). *Konotacja obcości w przekładzie*. Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej.

Lewicki, R. (2000). *Obcość w odbiorze przekładu*. Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej.

Linde-Usiekniewicz, J., Derwojedowa, M. i Zawisławska, M. (2008). Verbal Aspect and the Frame Elements in the FrameNet for Polish. W: E. Bernal i J. DeCesaris (red.) *Proceedings of the XIII EURALEX International Congress* (1511–1518). Barcelona: Universitat Pompeu Fabra.

Lipińska-Grzegorek, M. (1977). *Some problems of contrastive analysis: Sentences with nouns and verbs of sensual perception in English and Polish*. Edmonton: Linguistic Research.

Lison, P. i Tiedemann, J. (2016). OpenSubtitles2015: Extracting Large Parallel Corpora from Movie and TV Subtitles. W: N. Calzolari i in. (red.), *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC-2016), Portorož, Slovenia* (923–929). Paris: European Language Resources Association (ELRA).

Lopatková, M., Kettnerová, V., Bejček, E., Vernerová, A. i Žabokrtský, Z. (2016). *Valenční slovník českých sloves VALLEX*. Praha: Karolinum.

Lotko, E. (1992). *Zrádná slova v polštině a češtině (lexikologický pohled a slovník)*. Olomouc: Votobia.

Lotko, E. (1997). *Synchronní konfrontace češtiny a polštiny (soubor statí)*. Olomouc: Vydavatelství Univerzity Palackého.

Louw, B. (1993). Irony in the Text or Insincerity in the Writer? The Diagnostic Potential of Semantic Prosodies. W: M. Baker, G. Francis, E. Tognini-Bonelli i J. Sinclair (red.), *Text and technology. In honour of John Sinclair* (157–176). Philadelphia: John Benjamins Publishing Company.

Lüdtke, U. M. (red.). (2015). *Emotion in Language*. Amsterdam: John Benjamins Publishing Company.

Łaziński, M. (2014). Praktyczny przewodnik po korpusach równoległych. Wiadomości wstępne. Korpus ParaSol i Korpus polsko-rosyjski UW. W: M. Hebal-Jezierska (red.), *Praktyczny przewodnik po korpusach języków słowiańskich* (199–206). Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.

Łaziński, M., Kuratczyk, M., Orekhov, B. i Słobodjan, E. (2012). The Polish-Russian Parallel Corpus and Its Application in the Linguistic Analysis, *Prace Filologiczne*, LXIII, 209–218.

Malá, M. i Bruhová, G. (2018). English presentative semantic patterns as seen through a parallel translation corpus, *Languages in Contrast – International Journal for Contrastive Linguistics*. Pozyskano z: https://benjamins.com/catalog/lic.17007.mal.

Malmkjær, K. (2004). Censorship: Marry Howitt and a Problem in Descriptive TS. W: G. Hansen i D. Gile (red.), *Translation Studies: Claims, Changes and Challenges. Selected papers from EST Congress, Copenhagen 2001*. Amsterdam: John Benjamins.

Malmkjær, K. (2005). Translation and Linguistics. Perspectives, *Studies in Translation Theory and Practice*, 13(1), 5–20.

Maziarz, M., Piasecki, M. i Rudnicka, E. (2014). Słowosieć – polski wordnet. Proces tworzenia tezaurusa, *Polonica*, XXXIV, 79–97.

McArthur, T. (red.). (1992). *The Oxford Companion to the English Language*. New York: Oxford University Press.

McCall, S. M. (red.). (1967). *Polish Logic 1920–1939*. Oxford: Oxford University Press.

McEnery, T. i Xiao, R. (2008). Parallel and Comparable Corpora: What is Happening? W: G.M. Anderman i M. Rogers (red.), *Incorporating Corpora: The Linguist and the Translator* (18–31). Clevedon: Multilingual Matters.

McEnery, T., Xiao, R. i Tono, Y. (2006). *Corpus-Based Language Studies. An Advanced Resource Book*. London: Routledge.

Mielczuk, I. A. (1974). *Opyt teorii lingvističeskich modelej «Smysl ⇔ Tekst»*. Moskva: Nauka.

Mielczuk, I.A. (1988). Semantic Description of Lexical Units in an Explanatory Combinatorial Dictionary: Basic Principles and Heuristic Criteria, *International Journal of Lexicography*, 1(3), 165–188.

Mietła, J. (1998). *Multiwerbizacja w języku czeskim i polskim*. Toruń: Wydawnictwo Uniwersytetu Mikołaja Kopernika.

Mikołajczuk, A. (1997). Pole semantyczne 'gniewu' w polszczyźnie (Analiza leksemów: *gniew, oburzenie, złość, irytacja*). W: R. Grzegorczykowa i Z. Zaron (red.), *Semantyczna struktura słownictwa i wypowiedzi* (149–171). Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.

Mikołajczuk, A. (1999). *Gniew we współczesnym języku polskim. Analiza semantyczna*. Warszawa: Wydawnictwo Energeia.

Miller, G.A. (1995). WordNet: A Lexical Database for English, *Communications of the ACM*, 38(11), 39–41.

Młodzki, R., Kopeć, M. i Przepiórkowski, A. (2012). Word Sense Disambiguation in the National Corpus of Polish, *Prace Filologiczne*, LXIII, 155–166.

Morley, J. i Partington, A. (2009). A few Frequently Asked Questions about semantic – or evaluative – prosody, *International Journal of Corpus Linguistics*, 14(2), 139–158.

Muryc, J. (2015). K teorii mezijazykové interference v lexikálním plánu češtiny a polštiny, *Bohemistyka*, 15(4), 347–354.

Muryc, J. (2017). Polsko-czeska homonimia jako problem translatorski. W: M. Majewska (red.), *Wokół homonimii międzyjęzykowej* (97–105). Warszawa: Wydawnictwo Naukowe Uniwersytetu Kardynała Stefana Wyszyńskiego w Warszawie.

Nádvorníková, O. (2010). The French gérondif and its Czech equivalents. W: F. Čermák, P. Corness i A. Klégr (red.), *InterCorp: Exploring a Multilingual Corpus* (83–95). Praha: Nakladatelství Lidové noviny.

Navigli, R. (2009). Word Sense Disambiguation: A Survey, *ACM Computing Surveys*, 41(2).

Newmark, P. (1981). *Approaches to Translation*. New York: State University of New York Press.

Nida, E.A. (1964). *Toward a Science of Translating*. Leiden: Brill.

Nida, E.A. i Taber, C.R. (1974). *The Theory and Practice of Translation* (wyd. 2). Leiden: Brill.

Nivre, J. i Hall, J. (2005). MaltParser: A language-independent system for data-driven dependency parsing. W: *Proceedings of the Fourth Workshop on Treebanks and Linguistic Theories* (13–95). Barcelona, Spain.

Nowakowska-Kempna, I. (1986). *Konstrukcje zdaniowe z leksykalnymi wykładnikami predykatów uczuć.* Katowice: Uniwersytet Śląski.

Nowakowska-Kempna, I. (1995). *Konceptualizacja uczuc w jezyku polskim. Prolegomena.* Warszawa: Wyższa Szkoła Pedagogiczna – Towarzystwo Wiedzy Powszechnej.

Nowakowska-Kempna, I. (2000). *Konceptualizacja uczuc w jezyku polskim. Data.* Warszawa: Wyższa Szkoła Pedagogiczna – Towarzystwo Wiedzy Powszechnej.

Och, F.J. i Ney, H. (2003). A Systematic Comparison of Various Statistical Alignment Models, *Computational Linguistics*, 29(1), 19–51.

Oliva, K. (1994). *Polsko-český slovník. A–Ó* (Tom 1). Praha: Academia.

Oliva, K. (1995). *Polsko-český slovník. P–Ž* (Tom 2). Praha: Academia.

Orłoś, Z.T. (1980). *Polsko-czeskie związki językowe*. Kraków: Polska Akademia Nauk.

Orłoś, Z.T. (2003). *Czesko-polski słownik zdradliwych wyrazów i pułapek frazeologicznych*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.

Orłoś, Z.T. (2004). *Czesko-polska pozorna ekwiwalencja językowa*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.

Oster, U. (2010). Using corpus methodology for semantic and pragmatic analyses: What can corpora tell us about the linguistic expression of emotions? *Cognitive Linguistics*, 21(4), 727–763.

Padó, S. i Lapata, M. (2009). Cross-lingual annotation projection for semantic roles, *Journal of Artificial Intelligence Research*, 36, 307–340.

Pala, K. i Šachová, Z. (1986). Popis významů českých sloves označujících emoce. W: *Sborník prací Filozofické fakulty brněnské univerzity. A, Řada jazykovědná*, 35(A34), 99–119. Brno: Univerzita J.E. Purkyně.

Pala, K., Čapek, T., Zajíčková, B., Bartůšková, D., Kulková, K., Hoffmannová, P., Hajič, J. (2011). *Czech WordNet 1.9 PDT*. LINDAT/CLARIN digital library, http://hdl.handle.net/11858/00-097C-0000-0001-4880-3.

Palion-Musioł, A. (2012). Przekład audiowizualny jako wyzwanie dla współczesnego tłumacza. Narzędzia wykorzystywane w procesie translatorycznym, *Rocznik Przekładoznawczy*, 7, 95–107.

Panevová, J. (1975). On verbal Frames in Functional Generative Description. *Prague Bulletin of Mathematical Linguistics*, 23, 17–52.

Panevová, J. (1980). *Formy a funkce ve stavbě české věty*. Praha: Academia.

Panevová, J. (1994). Valency frames and the meaning of the sentence. W: P. Luelsdorff (red.), *The Prague School of Structural and Functional Linguistics. A Short Introduction* (223–243). Amsterdam–Philadelphia: John Benjamins.

Panevová, J. (1998). Ještě k teorii valence. *Slovo a slovesnost*, 59(1), 1–14.

Papierz, M. (1982). *Nominalizacje we współczesnym języku słowackim.* Zeszyty naukowe Uniwersytetu Jagiellońskiego, Prace językoznawcze 72. Kraków: Nakł. Uniwersytetu Jagiellońskiego.

Papierz, M. (2009). Realizacja kategorii określoności/nieokreśloności w tekstach słowackich i polskich. W: M. Cichońska (red.), *Kategorie gramatyczne a kategorie semantyczne i pragmatyczne w językach słowiańskich* (61–69). Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Partington, A.S. (2004). "Utterly content in each other's company": Semantic prosody and semantic preference, *International Journal of Corpus Linguistics*, 9(1), 131–156.

Partington, A.S. (2007). Semantic prosody and semantic preference. W: P. Hanks (red.) *Lexicology: Critical Concepts in Linguistics* (144–164). London: Routledge.

Petrincová, M. (2016). Wyszukiwanie ekwiwalentów w Polsko-Słowackim Korpusie Równoległym. W: E. Gruszczyńska i A. Leńko-Szymańska (red.), *Polskojęzyczne korpusy równoległe. Polish-language Parallel Corpora* (143–155). Warszawa: Wydział Lingwistyki Stosowanej Uniwersytetu Warszawskiego.

Pęzik, P. (2012). Wyszukiwarka PELCRA dla danych NKJP. W: A. Przepiórkowski, M. Bańko, R. L. Górski i B. Lewandowska-Tomaszczyk (red.), *Narodowy Korpus Języka Polskiego* (253–279). Warszawa: Wydawnictwo Naukowe PWN.

Piasecki, M., Kędzia, P. i Orlińska, M. (2016). plWordNet in Word Sense Disambiguation task. W: V. B. Mititelu, C. Forăscu, C. Fellbaum i P. Vossen (red.), *Proceedings of the 8th Global Wordnet Conference* (280–289). Bucharest, Romania: The Global WordNet Association.

Pieńkos, J. (1993). *Przekład i tłumacz we współczesnym świecie: aspekty lingwistyczne i pozalingwistyczne*. Warszawa: Wydawnictwa Naukowe PWN.

Pieńkos, J. (2003). *Podstawy przekładoznawstwa. Od teorii do praktyki*. Kraków: Zakamycze.

Piotrowski, T. (2011). Ekwiwalencja w słownikach dwujęzycznych. W: W. Chlebda (red.), *Na tropach translatów. W poszukiwaniu odpowiedników przekładowych* (45–69). Opole: Wydawnictwo Uniwersytetu Opolskiego.

Pisarska, A. i Tomaszkiewicz, T. (1996). *Współczesne tendencje przekładoznawcze.* Poznań: Wydawnictwo Naukowe UAM.

Polański, K. (red.). (1999). *Encyklopedia językoznawstwa ogólnego*. Wrocław.

Pollard, C. i Sag, I.A. (1994). *Head-Driven Phrase Structure Grammar*. Chicago: The University of Chicago Press.

Przepiórkowski, A. (2009). Towards the Automatic Acquisition of a Valence Dictionary for Polish. W: M. Marciniak i A. Mykowiecka (red.), *Aspects of Natural Language Processing. Essays dedicated to Leonard Bolc on the Occasion of His 75th Birthday.* Lecture Notes in Computer Science (5070) (191–210). Berlin: Springer.

Przepiórkowski, A. (2017). *Argumenty i modyfikatory w gramatyce i słowniku.* Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.

Przepiórkowski, A., Bańko, M., Górski, R.L. i Lewandowska-Tomaszczyk, B. (red.). (2012). *Narodowy Korpus Języka Polskiego*. Warszawa: Wydawnictwo Naukowe PWN.

Przepiórkowski, A., Hajnicz, E., Andrzejczuk, A., Patejuk, A.I Woliński, M. (2017). Walenty: gruntowny składniowo-semantyczny słownik walencyjny języka polskiego, *Język Polski*, XCVII(1), 29–46.

Ramón García, N. (2002). Contrastive Linguistics and Translation Studies Interconnected: The Corpus-based Approach, *Linguistica Antverpiensia, New Series – Themes in Translation Studies*, 1, 393–406.

Rosen, A. (2016). InterCorp – a look behind the façade of a parallel corpus. W: E. Gruszczyńska i A. Leńko-Szymańska (red.), *Polskojęzyczne korpusy równoległe. Polish-language Parallel Corpora* (21–40). Warszawa: Wydział Lingwistyki Stosowanej Uniwersytetu Warszawskiego.

Rosen, A. i Vavřín, M. (2014). *Korpus InterCorp – čeština,* verze 7 z 19.12.2014. http://www.korpus.cz. Praha: Ústav Českého národního korpusu Filozofické fakulty Univerzity Karlovy.

Rosen, A. i Vavřín, M. (2015). *Treq*. http://treq.korpus.cz. Praha: Ústav Českého národního korpusu Filozofické fakulty Univerzity Karlovy.

Rosen, A., Kaczmarska, E. i Škodová, S. (2014). Zdrobnienia jako element kultury i pułapka glottodydaktyczna. Czeskie i polskie deminutiva w ujęciu konfrontatywnym na podstawie badań korpusowych. W: E. Kaczmarska i A. Zieniewicz (red.), *Glottodydaktyka wobec wielokulturowości* (51–63). Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.

Rosen, A., Vavřín, M. i Zasina, A.J. (2017). *Korpus InterCorp*, verze 10 z 1.12.2017. http://www.korpus.cz. Praha: Ústav Českého národního korpusu Filozofické fakulty Univerzity Karlovy.

Rosen, A., Vavřín, M. i Zasina, A.J. (2018). *Korpus InterCorp*, verze 11 z 19.10.2018. http://www.korpus.cz. Praha: Ústav Českého národního korpusu Filozofické fakulty Univerzity Karlovy.

Roszko, R. (1993). *Wykładniki modalności imperceptywnej w języku polskim i litewskim*. Warszawa: Slawistyczny Ośrodek Wydawniczy.

Roszko, R. (2016). Korpusy wielojęzyczne. Rzutowanie anotacji z jednego języka na drugi. Referat wygłoszony na konferencji „CLARIN-PL". Wrocław. Pozyskano z: http://clarin-pl.eu/wp-content/uploads/2016/04/konferencja/Roszko_Wroclaw_Clarin_2016.pdf.

Rozwadowska, B. (2012). On the onset of psych eventualities. W: E. Cyran, H. Kardela i B. Szymanek (red.), *Sound, structure and sense. Studies in memory of Edmund Gussmann* (533–554). Lublin: Wydawnictwo Katolickiego Uniwersytetu Lubelskiego Jana Pawła II.

Rudnicka, E., Bond, F., Grabowski, Ł., Piasecki, M. i Piotrowski, T. (2017). Towards equivalence links between senses in pl WordNet and Princeton WordNet, *Lodz Papers in Pragmatics*, 13(1), 3–24.

Rudnicka, E., Piasecki, M., Bond, F., Grabowski, Ł. i Piotrowski, T. (2019). Sense Equivalence in plWordNet to Princeton WordNet Mapping, *International Journal of Lexicography*. Pozyskano z: https://doi.org/10.1093/ijl/ecz004

Ružička, J. (1968). Valencia slovies a intencia slovesného deja, *Jazykovedný časopis*, 19, 50–56.

Rytel, D. (1982). *Leksykalne środki wyrażania modalności w języku czeskim i polskim*. Wrocław: Zakład Narodowy im. Ossolińskich.

Rytel-Kuc, D. (1989). Wybrane problemy opisu walencyjnego języka, *Studia z Filologii Polskiej i Słowiańskiej*, XXVI, 237–247.

Rytel-Kuc, D. (1990). *Niemieckie passivum i man-Sätze a ich przeklad w języku czeskim i polskim*. Wrocław: Polska Akademia Nauk.

Rytel-Kuc, D. (red.). (1991). *Walencja czasownika a problemy leksykografii dwujęzycznej*. Wrocław: Zakład Narodowy im. Ossolińskich.

Saloni, Z. i Świdziński, M. (1985). *Składnia współczesnego języka polskiego*. Warszawa: Państwowe Wydawnictwo Naukowe.

Sgall, P. (1998). Teorie valence a její formální zpracování, *Slovo a slovesnost*, 59(1), 15–29.

Sharp, B., Zock, M., Carl, M. i Jakobsen, A.L. (red.). (2011). *Proceedings of the 8th International NLPCS Workshop. Human-Machine Interaction in Translation.* Copenhagen Studies in Language, 41. Frederiksberg, Denmark: Samfundslitteratur.

Siatkowski, J. (2006). *Obce nazwy geograficzne w języku czeskim i polskim*. Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.

Siatkowski, J. i Basaj, M. (2002). *Słownik czesko-polski*. Warszawa: Wiedza Powszechna.

Šimandl, J. (red.) (2016): *Slovník afixů užívaných v češtině*. Praha: Karolinum.

Simon, S. (1996). *Gender in Translation. Cultural identity and the politics*. London: Routledge.

Sinclair, J. (1996). The search for units of meaning, *Textus*, IX, 75–106.

Sinclair, J. (2003). *Reading concordances: An introduction*. London – New York: Pearson / Longman.

Sinclair, J. (2004). *Trust the text: Language, corpus and discourse*. New York: Routledge.

Šindlerová, J. i Bojar, O. (2009). Towards English-Czech parallel valency lexicon via treebank examples. W: M. Passarotti, A. Przepiórkowski, S. Raynaud i F. Van Eynde (red.), *Proceedings of the Eighth International Workshop on Treebanks and Linguistic Theories* (185–195). Milano: EDUCatt.

Šindlerová, J. i Bojar, O. (2010). Building a bilingual ValLex using treebank token alignment: First observations. W: N. Calzolari i in. (red.), *Proceedings of the Seventh International Language Resources and Evaluation (LREC'10)* (304–309). Valletta: European Language Resources Association (ELRA).

Šindlerová, J., Veselovská, K. i Hajič jr., J. (2014). Tracing sentiments: Syntactic and semantic features in a subjectivity lexicon. W: A. Abel, C. Vettori i N. Ralli (red.), *Proceedings of the XVI EURALEX International Congress* (405–413). Bolzano: EURAC research.

Skandera, P. i Burleigh, P. (2005). *A Manual of English Phonetics and Phonology: Twelve Lessons with an Integrated Course in Phonetic Transcription*. Tübingen: Gunter Narr Verlag .

Škodová, S., Kaczmarska, E. i Rosen, A. (2017). Deminutiva v češtině a polštině s důrazem na popis jejich užití v tzv. žákovském mezijazyce. W: W. Wysoczański i B. Gaska (red.), *Slavica Wratislaviensia CLXV. Acta Universitatis Wratislaviensis*, 357–369. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.

Škodová, S., Rosen, A. i Kaczmarska, E. (2016). Česká deminutiva a úskalí jejich užití pro nerodilé mluvčí. W: M. Švrčinová i Z. Vlasáková (red.), *Metody výuky a testování cizích jazyků (včetně češtiny pro cizince)* (207–215). Praha: Ústav jazykové a odborné přípravy Univerzity Karlovy.

Skoumalová, H. (2008). Extracting dictionaries from parallel corpora. W: Čermák, F., Marcinkevičienė, R., Rimkute, E. i Zabarskaite, J. (red.), *Proceedings of The Third Baltic Conference on Human Language Technologies* (297–301). Kaunas: Vytautas Magnus University.

Škrabal, M. i Vavřín, M. (2017). Databáze překladových ekvivalentů Treq, *Časopis pro moderní filologii*, 99(2), 245–260.

Skwarska, K. (2001). Tykání a jeho zdvořilejší protějšek v češtině a polštině. W: *Obraz světa v jazyce* (137–145). Praha: Univerzita Karlova.

Skwarska, K. i Kaczmarska, E. (red.). (2016). *Výzkum slovesné valence ve slovanských zemích*. Praha: Slovanský ústav Akademie věd České republiky.

Slovník českých synonym a antonym. (2007). Praha: Lingea.

Smith, K.A. i Nordquist, D. (2012). A critical and historical investigation into semantic prosody, *Journal of Historical Pragmatics*, 13(2), 291–312.

Smułczyński, M. (2013). Walencja semantyczna polskich i duńskich czasowników ruchu w ujęciu kontrastywnym. W: *Studia Linguistica XXXII. Acta Universitatis Wratislaviensis 3551* (173–188). Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.

Spagińska-Pruszak, A. (1994). *Język emocji. Studium leksykalno-semantyczne rzeczownika w języku polskim, rosyjskim i serbsko-chorwackim*. Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego.

Spalatin, L. (1969). Formal Correspondence and Translation Equivalence in Contrastive Analysis. W: *The Yugoslav Serbo-Croatian Contrastive Project B. Studies 1* (26–35). Zagreb.

Stankovska, P. (2018). Infinitiv jako pravovalenční doplnění sloves v češtině a ve slovinštině, *Prace Filologiczne*, LXXII, 345–355.

StatSoft. (2006). *Elektroniczny Podręcznik Statystyki*. Pozyskano z: http://www.statsoft.pl/textbook/stathome.html.

Štěpánek, J. (2006). Závislostní zachycení větné struktury v anotovaném syntaktickém korpusu (nástroje pro zajištění konsistence dat). Dysertacja. Praha: Matematicko-fyzikální fakulta Univerzity Karlovy.

Stubbs, M. (1995). Collocations and semantic profiles: On the cause of trouble with quantitative studies, *Functions of Language*, 2(1), 23–55.

Stubbs, M. (2001). *Words and phrases: Corpus studies of Lexical Semantics*. Oxford: Blackwell Publishers.

Švejcer, A. D. (1973). *Perevod i lingvistika: o gazetno-informacionnom i voenno-publicističeskom perevode*. Moskva: Voennoe Izdateľstvo Ministerstva Oborony SSSR.

Svoboda, K. (1973). O gramatické modálnosti se zřetelem k souvětí. W: *Otázky slovanské syntaxe III, Sborník symposia „Modální výstavba výpovědi v slovanských jazycích"* (233–241). Brno: Universita J. E. Purkyně.

Szarkowska, A. (2009). Przekład audiowizualny w Polsce – perspektywy i wyzwania, *Przekładaniec*, 20, 8–25.

Szczepańska, E. (1994). *Uniwerbizacja w języku czeskim a polskim*. Kraków: Universitas.

Szczepańska, E. (2004). Jeszcze o czesko-polskich pułapkach językowych, *Bohemistyka*, 3, 212–218.

Szwedowa, N.Y. (red.). (1970). *Grammatika sovremennogo russkogo literaturnogo jazyka*. Moskva: Nauka.

Szymczak, M. (red.). (1995). *Słownik języka polskiego*. Warszawa: Państwowe Wydawnictwo Naukowe.

Tabakowska, E. (1991). Przekład i obrazowanie. *Arka*, 34, 52–61.

Tabakowska, E. (1993). *Cognitive Linguistics and Poetics of Translation*. Tübingen: Gunter Narr Verlag.

Tabakowska, E. (1995). *Gramatyka i obrazowanie. Wprowadzenie do językoznawstwa kognitywnego*. Kraków: PAN.

Taber, C.R. i Nida, E.A. (1971). *La traduction: théorie et méthode*. London: Alliance Biblique Universelle.

Taboada, M. (2016). Sentiment Analysis: An Overview from Linguistics, *Annual Review of Linguistics*, 2.

Tantos, A. (2006). Rhetorical Relations in Verbal Eventual Representations? Evidence from Psych Verbs. W: P. Denis, E. McCready, A. Palmer i B. Reese (red.), *Proceedings of the 2004 Texas Linguistics Society Conference: Issues at the Semantics-Pragmatics Interface* (123–136). Somerville, MA: Cascadilla Proceedings Project.

Tarajło-Lipowska, Z. (2000). *Kapoan Naopak. O czeskim dla Polaków, być może mało zaawansowanych, ale mocno zainteresowanych*. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.

Taylor, J.R. (2002). *Cognitive Grammar*. Oxford: Oxford University Press.

Tesnière, L. (1953). *Esquisse d'une syntaxe structurale*. Paris: Klincksieck.

Tesnière, L. (1959). *Éléments de syntaxe structurale*. Paris: Klincksieck.

Tesnière, L. (2015). *Elements of Structural Syntax*. Amsterdam: John Benjamins.

Teubert, W. (2001). Corpus Linguistics and Lexicography, *International Journal of Corpus Linguistics*, 6(3), 125–153.

Tian, L., Wong, D.F., Chao, L.S. i Oliveira, F. (2014). A Relationship: Word Alignment, Phrase Table, and Translation Quality, *The Scientific World Journal,* 2014(438106). DOI: 10.1155/2014/438106.

Tian, L., Wong, F. i Chao, S. (2010). An Improvement of Translation Quality with Adding Key--Words in Parallel Corpus, *Machine Learning and Cybernetics (ICMLC)*, 3, 1273–1278.

Tiedemann, J. (2003). Combining Clues for Word Alignment. W: *Proceedings of the Tenth Conference of the European Chapter of the Association for Computational Linguistics* (339–346). Budapest, Hungary: The Association for Computational Linguistics.

Tiedemann, J. (2004). Word to word alignment strategies. W: *Proceedings of COLING 2004* (212–218). Geneva: COLING.

Tiedemann, J. (2007). Improved Sentence Alignment for Movie Subtitles. W: *Proceedings of the Conference on Recent Advances in Natural Language Processing (RANLP'07)*. Borovec, Bulgaria.

Tiedemann, J. (2008). Synchronizing Translated Movie Subtitles. W: *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC'2008)*. Marrakech, Marocco: European Language Resources Association (ELRA).

Tkaczewski, D. (2017A). Symbolika v českých a polských somatických frazémech s lexémy słuch/sluch a ucho a jejich vzájemná ekvivalence. *Bohemica litteraria*. *Pocta Jiřímu Levému – Ad translationem*, 20(2), 21–34.

Tkaczewski, D. (2017B). Symbolika w czeskich i polskich frazeologizmach somatycznych z leksemami zrak/wzrok i oko oraz ich ekwiwalencja wzajemna. W: L. Janovec (red.), *Svět v obrazech a ve frazeologii / World in Pictures and in Phraseology* (65–80). Praha: Pedagogická fakulta Univerzity Karlovy.

Tkaczewski, D. (2018). Mentalnościowy i językowy „obraz sąsiada". Polsko-czeskie i czesko-polskie stereotypy językowe i aproksymaty. W: M. Cichońska i I. Genew-Puhalewa (red.), *Tożsamość Słowian zachodnich i południowych w świetle dwudziestowiecznych dyskusji i polemik. Konteksty filologiczne i kulturoznawcze* (Tom 2, 217–236). Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Tognini-Bonelli, E. (2001). *Corpus Linguistics at Work*. Studies in Corpus Linguistics 6. Amsterdam: John Benjamins.

Tognini-Bonelli, E. (2002). Functionally complete units of meaning across English and Italian. Towards a corpus-driven approach. W: B. Altenberg i S. Granger (red.), *Lexis in Contrast*, Studies in Corpus Linguistics 7 (73–95). Amsterdam: John Benjamins.

Tokarz, B. (2010). *Spotkania. Czasoprzestrzeń przekładu artystycznego.* Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Tomaszkiewicz, T. (2006). *Przekład audiowizualny.* Warszawa: Wydawnictwo Naukowe PWN.

Tomaszkiewicz, T. (2010). Przekład audiowizualny, werbalno-wizualny czy intersemiotyczny: różne wymiary tej samej rzeczywistości? *Lingwistyka Stosowana*, 3, 33–44.

Toury, G. (1980A). *In Search of Theory of Translation.* Tel Aviv: The Porter Institute for Poetics and Semiotics, Tel Aviv University.

Toury, G. (1980B). Contrastive Linguistics and Translation Studies: Towards a tripartite model. W: G. Toury (red.), *In Search of a Theory of Translation* (19–34). Tel Aviv: The Porter Institute for Poetics and Semiotics, Tel Aviv University.

Tryuk, M. (2006). *Przekład ustny środowiskowy*. Warszawa: Wydawnictwo Naukowe PWN .

Tryuk, M. (2007). *Przekład ustny konferencyjny*. Warszawa: Wydawnictwo Naukowe PWN .

Tryuk, M. (2009). Co to jest tłumaczenie audiowizualne? *Przekładaniec*, 20, 26–39.

Tufiş, D.I., Koeva, S., Erjavec, T., Gavrilidou, M. i Krstev, C. (2008). Building Language Resources and Translation Models for Machine Translation Focused on South Slavic and Balkan Languages. W: M. Tadić, M. Dimitrova-Vulchanova i S. Koeva (red.), *Proceedings of the Sixth International Conference Formal Approaches to South Slavic and Balkan Languages (FASSBL 2008)* (145–152). Zagreb: Croatian Language Technologies Society, Faculty of Humanities and Social Science.

Tuong, L.D. (2016). *Contrastive Linguistics: An Introduction*. Nghe An: Vinh University.

Tutak, K. (2003). *Leksykalne nieczasownikowe wykładniki modalności epistemicznej w autobiografiach*. Kraków: Księgarnia Akademicka.

Understanding Machine Learning #1 – How machines learn? (2016, 8 3). Pozyskano z: http://e-abm.com/understanding-machine-learning-1-how-machines-learn/.

Understanding machine learning #2: Do we need machine learning at all? (2016, 8 23). Pozyskano z: http://e-abm.com/understanding-machine-learning-2-do-we-need-machine-learning-at-all/.

Understanding machine learning #3: Confusion matrix – not all errors are equal. (2016, 10 11). Pozyskano z: http://e-abm.com/understanding-machine-learning-3-confusion-matrix-not-all-errors-are-equal/.

Urbańczyk-Adach, N. (2011). *Wariantywność walencji czeskiego czasownika*. Warszawa: Slawistyczny Ośrodek Wydawniczy.

Urešová, Z., Fučíková, L., Hajičová, E. i Hajič, J. (2018). A Cross-lingual Synonym Classes Lexicon, *Prace Filologiczne*, LXXII, 403–416.

Urešová, Z., Štěpánek, J., Hajič, J., Panevová, J. i Mikulová, M. (2014). *PDT-Vallex: Czech Valency lexicon linked to treebanks*. Praha: LINDAT/CLARIN digital library. http://hdl.handle.net/11858/00-097C-0000-0023-4338-F.

Vandevoorde, L. (2016). *On semantic differences: a multivariate corpus-based study of the semantic field of inchoativity in translated and non-translated Dutch*. Ghent University. Faculty of Arts and Philosophy, Ghent, Belgium.

Voellnagel, A. (2014). *Jak nie tłumaczyć tekstów technicznych* (wyd. 5). Warszawa: Wydawnictwo Translegis.

von Waldenfels, R. (2006). Compiling a parallel corpus of Slavic languages. Text strategies, tools and the question of lemmatization in alignment. W: B. Brehmer, V. Zdanova i R. Zimny (red.), *Beiträge der Europäischen Slavistischen Linguistik (POLYSLAV)*, 9 (123–138). München: Otto Sagner.

von Waldenfels, R. (2010). How good is parallel corpus data? Contrasting Polish and Czech *must* and *can* in InterCorp. W: F. Čermák, P. Corness i A. Klégr (red.), *InterCorp: Exploring a Multilingual Corpus* (96–116). Praha: Nakladatelství Lidové noviny.

von Waldenfels, R. (2011). Recent developments in ParaSol: Breadth for depth and XSLT based web concordancing with CWB. W: D. M. i R. Garabík (red.), *Natural Language Processing, Multilinguality. Proceedings of Slovko 2011* (156–162). Bratislava: Jazykovedný ústav Ľudovíta Štúra, Slovenská akadémia vied.

von Waldenfels, R. (2012). ParaSol: Introduction to a Slavic Parallel Corpus, *Prace Filologiczne*, LXIII, 293–301.

von Waldenfels, R. i Meyer, R. (2006). *ParaSol, a Corpus of Slavic and Other Languages.* http://www.parasolcorpus.org.

Walczak, J. (2013). *Teoria i praktyka polskiej translatoryki na przykładzie nowopolskich tłumaczeń wybranych utworów Wiliama Shakespeare'a i Johna Miltona*. Dyzertacja. Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.

Warchał, M. (2011A). O leksyce bezekwiwalentnej w tłumaczeniu. Szkic psycholingwistyczny, *Przekłady Literatur Słowiańskich*, 2(1), 138–143.

Warchał, M. (2011B). Przekład artystyczny wobec funkcji dyskursu aksjologicznego w tekście źródłowym, *Przekłady Literatur Słowiańskich*, 2(1), 97–106.

Wiemer, B. (2018). Evidentials and Epistemic Modality. W: A.Y. Aikhenvald (red.), *The Oxford Handbook of Evidentiality* (85–108). Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780198759515.001.0001.

Wierzbicka, A. (1971). *Kocha – lubi – szanuje. Medytacje semantyczne*. Warszawa: Wiedza Powszechna.

Wierzbicka, A. (1980). *Lingua mentalis: The semantics of natural language*. Sydney–New York: Academic Press.

Wierzbicka, A. (1991). *Cross-cultural pragmatics: the semantics of human interaction*. Berlin–New York: Mouton de Gruyter.

Wierzbicka, A. (1997). *Understanding cultures through their key words: English, Russian, Polish, German, and Japanese*. New York–Oxford: Oxford University Press.

Wierzbicka, A. (2001). *What Did Jesus Mean? Explaining the Sermon on the Mount and the parables in simple and universal human concepts*. New York–Oxford: Oxford University Press.

Wiking, M. (2017). *Hygge. Klucz do szczęścia*. Warszawa: Wydawnictwo Czarna Owca.

Wojtasiewicz, O. (1957). *Wstęp do teorii tłumaczenia*. Wrocław: Zakład im. Ossolińskich.

Wróbel, H. (1991). O modalności, *Język Polski* LXXI, 260–270.

Wu, L., Xia, Y., Zhao, L., Tian, F., Qin, T., Lai, J. i Liu, T.-Y. (2018). *Adversarial Neural Machine Translation*. Pozyskano z: https://arxiv.org/pdf/1704.06933.pdf.

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W. i in. (2016). *Google's neural machine translation system: Bridging the gap between human and machine translation.* Pozyskano z: https://arxiv.org/pdf/1609.08144.pdf.

Xiao, R. i McEnry, T. (2006). Collocation, Semantic Prosody, and Near Synonymy: A Cross-Linguistic Perspective, *Applied Linguistics*, 27(1), 103–129.

Yarowsky, D. i Ngai, G. (2001). Inducing Multilingual POS Taggers and NP Bracketers via Robust Projection Across Aligned Corpora. W: *NAACL '01 Proceedings of the second meeting of the North American Chapter of the Association for Computational Linguistics on Language technologies* (1–8). Stroudsburg, PA: Association for Computational Linguistics.

Yaumen, N. (2013). *Właściwości dystrybucyjne predykatów czasownikowych o znaczeniu emotywnym: analiza konfrontatywna wybranych czasowników polskich i białoruskich*. Dyzertacja. Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.

Zakharkevich, A. (2009). Ocalone czy utracone nazwy własne w tłumaczeniach baśni Marii Konopnickiej „O krasnoludkach i sierotce Marysi"? *Między oryginałem a przekładem*, XV, 303–318.

Zaliwska-Okrutna, U. (2002). Czy tłumaczenia mogą być kobiece lub męskie? W: W. Chłopicki (red.), *Język a komunikacja 4.2: Język trzeciego tysiąclecia II Tom 2 – Polszczyzna i języki obce: przekład i dydaktyka* (81–92). Kraków: Tertium.

Zaron, Z. (2012). Konotacja nie jedno ma imię. Wymagania składniowe nazw osobowych. W: J.D. Apresjan, I. Boguslavsky, M.-C. L'Homme, L. Iomdin, J. Milicevic, A. Polguère i L. Wanner (red.), *Meaning, Texts and other Exciting Things. A Festschrift to Commemorate the 80th Anniversary of Professor Igor Alexandrovič Mel'čuk* (672–681). Moskva: Languages of Slavic Culture.

Zaron, Z. i Skwarska, K. (red.). (2017). *Prace Filologiczne* (Tom LXX).

Zaron, Z. i Skwarska, K. (red.). (2018). *Prace Filologiczne* (Tom LXXII).

Zasina, A.J. (2016A). Korespondence kolokací s lexémem *ženský/mužský* v překladu z češtiny do polštiny, *Studia Slavica*, 20(2), 123–132.

Zasina, A.J. (2016B). Překlad a Korpus. Textová ekvivalence vybraných českých adjektiv, *Studia Filologiczne*, 5, 25–58.

Zawisławska , M. (1998). The meaning of English verb see and its Polish equivalent widzieć in a perspective of frame semantics. W: W. Teubert, E. Tognini Bonelli i N. Volz (red.), *TELRI. Trans-European Language Resources Infrastructure. Proceedings of the Third European Seminar "Translation Equivalence", Montecatini Terme, Italy October 16–18, 1997*. Mannheim: TELRI Association e.V., Institut fur deutsche Sprache, The Tuscan Word Centre.

Zennaki, O., Semmar, N. i Besacier, L. (2019). A neural approach for inducing multilingual resources and natural language processing tools for low-resource languages, *Natural Language Engineering*, 25(1), 43–67.