

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/299358533>

O dwóch czeskich jednostkach leksykalnych będących wykładnikami negatywnych stanów emocjonalnych i ich polskich ekwiwalentach. Analiza na materiale z korpusu paralelnego InterCorp

Chapter · March 2016

DOI: 10.13140/RG.2.1.3486.6325

CITATION

1

READS

153

1 author:



[Elzbieta Kaczmarska-Zglejszewska](#)

University of Warsaw

69 PUBLICATIONS 169 CITATIONS

SEE PROFILE

O dwóch czeskich jednostkach leksykalnych będących wykładnikami negatywnych stanów emocjonalnych i ich polskich ekwiwalentach. Analiza na materiale z korpusu paralelnego InterCorp

On two Czech lexical items representing negative emotional states and their Polish equivalents: An analysis based on the parallel corpus InterCorp

Abstract

The analysis is focused on the Czech verbs *žárlit* ‘to be jealous’ and *závidět* ‘to envy’. The goal is to establish their closest equivalents in Polish. We use dictionary definitions to find the correct meaning of the analyzed verbs and link them with equivalents proposed by a traditional Czech-Polish dictionary. Equivalents automatically extracted from the corpus help us to find translations available in *InterCorp*. Although the results are consistent with those proposed by the bilingual dictionary, the number of equivalents found in *InterCorp* is larger.

Next, we apply a method developed in our pilot studies, including automatic excerption of given words with aligned segments from *InterCorp*. The segments are analysed manually. In each segment we check how a given word was translated and we examine its collocations and arguments. The study was supposed to determine if valence requirements could influence the choice of an equivalent in Polish. A pilot study concerning the ambiguous Czech verb *toužit* ‘to miss, to want, to desire’ (Kaczmarska, Rosen, 2013) was supposed to reveal if valence can influence the choice of an equivalent in Polish. It was assumed that for some senses the equivalent can be established based on the convergence of the valence requirements (Levin, 1993).

Unfortunately, for the analysed Czech verbs *žárlit* and *závidět*, the number of occurrences is insufficient, so collocation profiling (using the Word Sketch tool available in the Sketch Engine) cannot be applied to analyze syntactic contexts. We conduct a corpus-based research instead. The data from *InterCorp* confirm our assumptions based on the dictionary definitions. The equivalent-searching

algorithm, based also on a syntactico-semantic analysis (automatic extraction of pairs of equivalents, valence analysis, Case Grammar, Pattern Grammar, Cognitive Grammar) and described in last part of the paper, cannot be applied to the two verbs. We found Word Sketch as a promising tool for our research and we hope it to be the turning point for building our algorithm (Word Sketch for the Czech part of *InterCorp* is in the phase of preparation). We hope that our algorithm will be able to cooperate with machine translation tools. This is why, in addition to a manual analysis, we also try to conduct experimental trials of stochastic modelling of the choice of an equivalent on the basis of the context (Kaczmarska et al., 2015).

Słowa kluczowe: korpus równoległy, język polski, język czeski, ekwiwalent, analiza syntaktyczno-semantyczna

Keywords: parallel corpus, Polish, Czech, equivalent, syntactico-semantic analysis

1. Wprowadzenie

Celem tego artykułu jest próba ustalenia ekwiwalentów dwóch czeskich jednostek będących językowymi wykładnikami emocji negatywnych. Analizie zostaną poddane struktury semantyczno-składniowe tych jednostek oraz ich polskich odpowiedników, szczegółowo opisane zostaną czasowniki *závidět* i *žárlit* oraz ich polskie ekwiwalenty¹. Badanie przeprowadzone zostanie na materiale z korpusu równoległego InterCorp (Čermák, Rosen, 2012; Kaczmarska, Rosen, 2014b; Rosen, Vavřín, 2014).

Wskazywanie trafnych ekwiwalentów często bywa wyzwaniem nie tylko w przypadku konfrontacji języków odległych geograficznie, kulturowo i dla siebie wzajemnie egzotycznych. Nieoczekiwane trudności pojawiają się również na styku języków blisko spokrewnionych². Niemożność oddania w języku docelowym dokładnie tego samego, co zostało zakomunikowane w języku wyjściowym, wiąże się często z problemem dokładnego zrozumienia przekładanej jednostki³ (Kaczmarska, Rosen, 2014a; w druku). Szczególnie problemowymi w tym kontekście jawią się czasowniki będące wykładnikami różnych stanów psychicznych, które same w sobie mogą być źródłem niejasności i nieporozumień. Pisała

1 Badanie to jest częścią większego projektu omawiającego sposoby ustalania ekwiwalentów czeskich jednostek wyrażających różne stany emocjonalne.

2 Oczekiwany trudności można nazwać pojawienie się zjawiska „fałszywych przyjaciół” czy leksyki bezekwiwalentnej (Kaczmarska, 2014a).

3 Przekład jest procesem złożonym z kilku etapów, z których podstawowym wydaje się odkodowanie znaczenia jednostki w języku wyjściowym. Już ten etap może nastroczać wiele trudności (Kaczmarska, Rosen, 2015b).

o tym m.in. A. Wierzbicka, wspominając o relacji między przeżywaniem uczuć i mówieniem o nich (Wierzbicka, 1971: 30)⁴:

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 niewyraźne.

Myśl zawarta w tym cytacie sugeruje, że w języku docelowym niemożliwe jest wyrażenie tego, czego nie da się oddać w języku wyjściowym. I to opracowanie nie ma na celu tego, co niemożliwe. W procesie ustalania odpowiednich ekwiwalentów czasowników oznaczających emocje, skoncentrujemy się na tym, co już zostało wyrażone, czyli na znaczeniu konkretnej jednostki. Zadanie to jest utrudnione, ponieważ analizowane czasowniki są wieloznaczne; często trudno ustalić, co w danym kontekście znaczą, w związku z tym skomplikowany jest ich przekład, podczas którego możemy stracić część znaczenia (Kaczmarska, 2015a, 2015b; Kaczmarska, Rosen, w druku).

W poszukiwaniu trafnego ekwiwalentu mógłby pomóc słownik dwujęzyczny, jednak ze względu na swoje ograniczenia najczęściej nie podaje on odpowiedników wraz z przykładami użycia (Siatkowski, Basaj, 2002). Większymi możliwościami teoretycznie dysponują słowniki i translatory elektroniczne oraz słowniki umieszczone w Internecie w postaci plików⁵, jednakże dla języków czeskiego i polskiego nie są one całkowicie wiarygodne. Na przykład w przypadku analizy jednostki *mít rád*, tłumaczonej na język polski jako *lubić* i *kochać* (Siatkowski, Basaj, 2002), popularny portal Glosbe⁶ proponuje dwa odpowiedniki – *lubić* i *podobac się*, czyniąc ją tylko pozornie bardziej jednoznaczna, 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. Translatory automatyczne obsługujące język polski i czeski również nie są na razie wiarygodnym źródłem ekwiwalentów. Działający na zasadzie translatora (a nazwany słownikiem) portal Tri-Lite⁷ tłumaczy *milovat* jako *miłość*, a *mít rád* jako *jak*. Podobnie funkcjonuje portal Etranslator⁸, który zarówno dla *milovat*, jak

4 Ta klasyczna lektura jest wciąż aktualna. Od tego czasu powstało jednak szereg nowszych opracowań dotyczących komunikowania emocji w ujęciu konfrontatywnym (Wierzbicka 1991, 1997; Harkins, Wierzbicka, 2001; Lüdtke, 2015).

5 Na przykład: *Czesko-polski słownik tematyczny dla uczniów i studentów* autorstwa Dariusza Sieczkowskiego i Aleny Wolfovej – <http://stodolni.org/wp-content/uploads/2013/09/slovníktematyczny.pdf>

6 Dostęp online: <https://pl.glosbe.com/cs/pl/>

7 Dostęp online: <http://www.cz-pl.tri-lite.pl/>

8 Dostęp online: <http://www.etranslator.ro/pl/czeski-polski-online-tlumacz.php>

i *mít rád* proponuje ekwiwalent *miłość* (Kaczmarska, 2015b). Także popularny Google Tłumacz nie jest dobrym źródłem ekwiwalentów tego typu jednostek⁹.

W badaniu tym chcielibyśmy ustalić, czy możliwe jest znalezienie trafnych ekwiwalentów dla analizowanych jednostek, opierając się na materiale wyekscerpowanym z korpusu równoległego. W centrum zainteresowania będą dwa czeskie czasowniki *závidět* i *žárlit*. Czasowniki te nie są synonimami, a jednak oba odnoszą się w języku polskim do zazdrości.

Schemat analizy jest zbliżony do metody zastosowanej we wcześniejszych badaniach dotyczących czasowników odsyłających do różnych emocji pozytywnych np. *toužit*, *mít rád*, *milovat* (Kaczmarska, 2014a, 2014b, 2015b; Kaczmarska, Rosen, 2013, 2014a, w druku; Kaczmarska et al., 2015) czy np. *zdat se* (Kaczmarska, 2012)¹⁰. Obejmuje on wygenerowanie słownika czesko-polskiego¹¹, automatyczną ekscerpcję przykładów z czesko-polskiej części korpusu InterCorp oraz analizę manualną.

Słownik czesko-polski generujemy dzięki narzędziu Treq¹² dostępnemu na stronach Czeskiego Korpusu Narodowego (dalej – CNK):



Rysunek 1. Interfejs serwisu Treq.

Serwis Treq umożliwia wygenerowanie listy najczęstszych ekwiwalentów badanych czasowników. Należy jednak pamiętać, iż nie jest to idealny słownik,

9 Czeska fraza *Mám Tě rád*. jest przełożona na język polski jako *Kocham cię* (bez alternatywy *Lubię cię*.) – dostęp 25.05.2015.

10 W artykułach tych okazjonalnie pojawiały się też odniesienia do czasowników odsyłających do emocji negatywnych, np. *mrzet*, *být lito*.

11 Poprzednie badania poprzedzone były samodzielnym generowaniem słownika (Kaczmarska, Rosen, 2013). Obecnie CNK oferuje usługę automatycznego generowania różnych słowników – dostęp: <http://treq.korpus.cz/>

12 <http://treq.korpus.cz/> – dostęp 26.05.2015

zawierający wyłącznie trafione ekwiwalenty. Ponieważ przy tej ilości danych narzędzie stosuje metodę automatyczną, wśród odpowiedników znaleźć się mogły sporadycznie nieekwiwalentne lub przypadkowe wyrazy, a nawet znaki interpunkcyjne (Kaczmarek, Rosen, 2013).

2. Analiza możliwości rozumienia i przekładu czasownika *závidět*

Tradycyjny słownik czesko-polski (Siatkowski, Basaj, 2002: 999) podaje jedno znaczenie tego wyrazu: *zazdrościć komu czego*¹³.

Odpowiednik zaproponowany przez słownik tradycyjny odpowiada pod względem wymagań składniowych czeskiej jednostce. Większą różnorodność ekwiwalentów przynosi korpusowe narzędzie Treq.

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

Tabela 1. Najczęstsze polskie ekwiwalenty czasownika *závidět* na podstawie serwisu Treq

Kilka z pojawiających się w tabeli ekwiwalentów tłumaczeniowych 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 możemy interpretować jako luźny synonim jednostki *závidět* w jednym z jej znaczeń. Znaczenia czasownika *závidět* oraz jego struktury składniowe możemy odnaleźć w słowniku literackiego języka czeskiego [dalej – SSJC (Havránek, 1989)]¹⁴.

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

14 Tłumaczenia definicji i przykładów ze słownika – autorka artykułu. Dostęp 15.03.2015: <http://ssjc.ujc.cas.cz/search.php?hledaj=Hledat&heslo=z%C3%A1vid%C4%Bti&sti=EMPTY&where=hesla&hsubstr=no>.

záviděti

- komu; komu co, koho, †čeho¹⁵
- pocítovat závist k někomu nepřát (komu co, koho)¹⁶
- *záviděti* bohatcům, šťastným lidem; lidé si navzájem závidí;
z. mladé dívky krásu, ženicha; mnozí mu té slávy záviděli; tu
funkci ti nezávidím nechtěl bych ji¹⁷

Definicja i charakterystyka składniowa zamieszczona w SSJC (Havránek, 1989) jest zgodna z opisem w słowniku walencyjnym języka czeskiego VALLEX¹⁸:

závidět^{impf}, závidívat^{iter}

[1] ≈ pocítovat závist; nepřát

-frame: **ACT**₁^{obl} **ADDR**₃^{obl} **PAT**₄^{obl}_{4,že,cont}

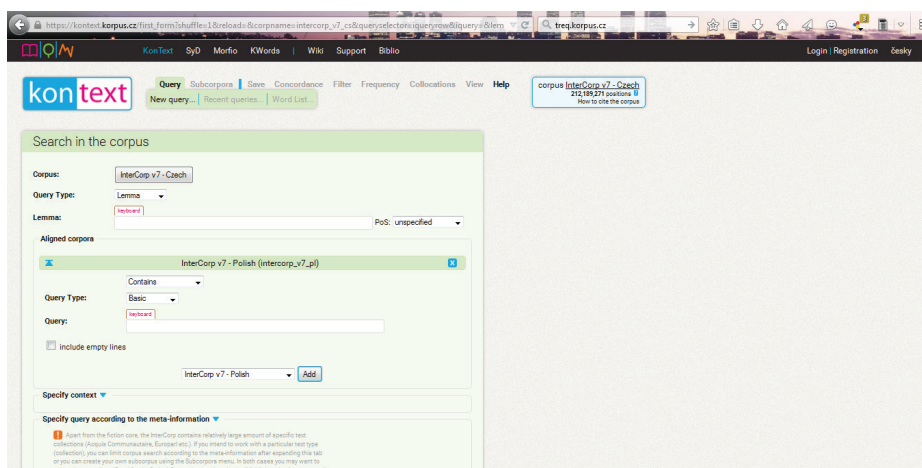
-example: záviděl jí její vzdělání / rozhled / že je chytrá; závidí i nos mezi očima; závidí mu, co všechno má

-diat: deagent: v naší vesnici se závidí i nos mezi očima

-rcp: ACT-ADDR: Jan a Marie si záviděli

-class: mental action

Rysunek 2. Hasło słownika walencyjnego VALLEX¹⁹



Rysunek 3. Interfejs wyszukiwarki KonText

¹⁵ komu? komu co? / kogo? (biernik), czego? (przestarzała struktura dopełniaczowa)

¹⁶ czuć zawiść wobec kogoś, nie życzyć komu, co / kogo (biernik)

¹⁷ *záviděti* bogaczom, szczęśliwym ludziom; ludzie sobie nawzajem zazdroścżą; młodej dziewczynie urody, narzeczonego; wielu mu tej slawy zazdrości; tej funkcji ci nie zazdroścżą, nie chciałbym jej

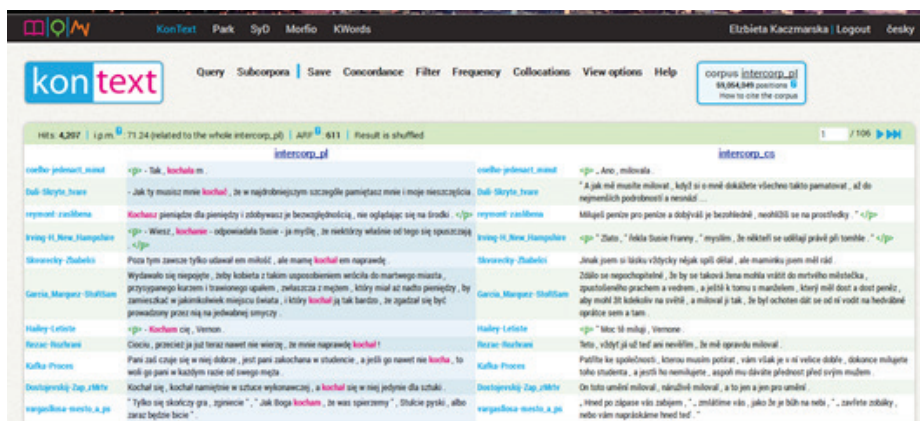
¹⁸ Dostęp online: <http://ufal.mff.cuni.cz/vallex/2.6.3/data/html/generated/alphabet/index.html>

¹⁹ czuć zawiść, nie życzyć; zazdrościć 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.

Definicje z obu słowników oraz z listy odpowiedników wygenerowanej z InterCorpu przybliżają nas do poznania znaczenia tego czasownika, a tym samym do potwierdzenia trafności ekwiwalentu ze słownika. Pełny obraz uzyskujemy dzięki przykładom wyekscerpowanym z InterCorpu²⁰.

2.1 Analiza korpusowa czasownika *závidět*

Materiał do badań ekscerpujemy poprzez wyszukiwarkę KonText.²¹



Rysunek 4. Interfejs wyszukiwarki KonText po wyszukaniu paralelnych konkordancji

cz. Tu jejich zjevnou pohodu a bezstarostnost **jim** upřímně *záviděl*.²²
 pl. Szczere **im** zazdrościł pogody ducha i beztroski.
 [Viewegh-UcastniciZajez]

cz. Agnes *záviděla* **Paulovi**, že žije, aniž si musí stále uvědomovat, že má tělo.
 pl. Agnes zazdrościła **Paulowi**, że może żyć bez ciągłej świadomości własnego ciała. [Kundera-Nesmrtelnost]

Wyszukany materiał eksportujemy do dokumentu Excel²³, dzięki czemu możemy go dowolnie filtrować i sortować, a po dodaniu kolumn – także tagować.

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

21 Od kwietnia 2015 jest to jedyna wyszukiwarka obsługująca CNK.

22 Tekst podkreślony to przedmiot zazdrości – *zazdrościć kogo czego* (w przypadku czasownika *žárlit* będzie to fraza przyimkowa *na* + obiekt, podobnie jak *być zazdrośny o* + obiekt), natomiast tekst podkreślony podwójną linią oznacza obiekt – *powód zazdrości* wyrażony frazą zdaniową. Frazy pogrubione to obiekt, któremu się zazdrości – *zazdrościć komu*.

23 Wyszukiwarka Excel umożliwia eksport konkordancji do różnego typu plików (Kaczmarek, Rosen, 2014b: 216).

KonText wskazał 50 poświadczeń leksemu *závidět*²⁴. Przy każdej konkordancji zaznaczyliśmy, jaki odpowiednik polski został wybrany. Po otagowaniu okazało się, iż w 90% przypadków w polskim tłumaczeniu występuje czasownik *zazdrościć* (ewentualnie wariant *pozazdrościć*).

Polskie odpowiedniki czasownika <i>závidět</i>	50
zazdrościć	45
pozazdrościć	3
być zazdrosny	1
Inne	1

Tabela 2. Polskie odpowiedniki czasownika *závidět* na podstawie materiału z czesko-polskiej części korpusu InterCorp

Jedynie tu poświadczenie z ekwiwalentem *zazdrosny*, przedstawia przykład z elipsą.

cz. (...) spokojen , že **mu** nemá co *závidět*...
 pl. (...) zadowolony, że nie potrzebuje *być zazdrosny*...
 [Paral-VeletrhSplnenych]

Moglibyśmy więc uznać, iż optymalnym ekwiwalentem jest dla tej jednostki czasownik *zazdrościć*. Ten jednak jest wieloznaczny (Szymczak, 1995: 916):

zazdrościć

- odczuwać zazdrość²⁵
- odczuwać żal z powodu tego, że komuś dobrze się powodzi, że ktoś coś ma, pragnąc tego dla siebie

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 się też w jednym ze znaczeń rzeczownika *zazdrość* (Szymczak, 1995: 916):

zazdrość

- uczucie przykrości, żalu spowodowane czyimś powodzeniem, szczęściem, stanem posiadania itp. i chęć posiadania tego samego
- uczucie niepokoju co do wierności osoby kochanej, podejrzliwość i dążenie do wyłączności w tym zakresie, chęć przeciwdziałania ewentualnemu naruszeniu tej wyłączności

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

25 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)

Odczytując znaczenie czasownika *zazdrościć* zgodnie z wcześniej przytoczoną definicją jako ‘odczuwać zazdrość’, możemy owo znaczenie rozszerzyć do postaci: ‘odczuwać uczucie przykrości, żalu spowodowane czymś powodzeniem, szczęściem, stanem posiadania itp. i chęć posiadania tego samego’.

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

pl. Stałem więc nadal, ufając w szczęśliwy zbieg okoliczności i *zazdroszcząc Wilhelmowi Habie, że tak łatwo znalazł wyjście z trudnej sytuacji*. [Rezac-Rozhrani]

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

3. Analiza możliwości rozumienia i przekładu czasownika *žárlit*

Drugą jednostką związaną z pojęciem zazdrości jest w języku českim wspomniany już czasownik *žárlit*. Analiza przebiega w ten sam sposób jak w przypadku poprzedniej jednostki.

Słownik česko-polski (Siatkowski, Basaj, 2002:1049) podaje dwa znaczenia tego czasownika: *być zazdrośnym (o kogo, o co)* i *zazdrościć (komu)*. Serwis Treq dostarcza szerszej gamy ekwiwalentów:

Najczęstsze polskie ekwiwalenty czasownika <i>žárlit</i> na podstawie serwisu Treq	
Zazdrośny	141
Zazdrość	25
zazdrościć	14
Być	2
Osilek	1
darzyć	1
owszem	1
rywalka	1
zawiść	1

Tabela 3. Najczęstsze polskie ekwiwalenty czasownika *žárlit* na podstawie serwisu Treq

Podobnie jak w sytuacji poprzedniego czasownika, również tutaj wśród oferowanych ekwiwalentów znajdują się jednostki przypadkowo związane (*być, osilek, darzyć, owszem*). Słowa *rywalka* i *zawiść* możemy traktować jako elementy struktur synonimicznych wobec słowa oznaczającego pojęcie ‘zazdrość’. Ekwiwalenty te warto porównać ze znaczeniem omawianego czasownika proponowanym przez słownik języka českiego (Havránek, 1989)²⁶.

²⁶ Na podstawie definicji (dostęp online 15.03.2015): <http://ssjc.ujc.cas.cz/search.php?hledej=Hledat&heslo=%C5%BE%C3%A1rliti&sti=EMPTY&where=hesla&substr=no>

žárliti

- na koho, co; ~; řs kým²⁷; pociťovat nelibost, bolest, zlost vůči někomu milovanému, protože má rád někoho jiného; pociťovat nelibost, bolest, zlost vůči někomu, koho má ráda milovaná osoba²⁸
- (na koho); pociťovat řevnivost, nevraživost na někoho pro jeho úspěchy; nevražít, řevnit 1; (nač) nepřát, závidět (komu co)²⁹

Z tymi definicjami zgodny jest opis w słowniku VALLEX:

žárlit^{impf}, žárlívat^{iter}

PDT-Vallex:
žárlit

1 ≈ pociťovat nelibost / bolest / vztek / zlost k něčemu / na někoho; závidět

-frame: ACT^{obl} PAT^{obl} na-4,že,cont

-example: Davídek žárlil na svou mladší sestřičku; žárlil na jeho úspěchy; nežárlím na jeho novou přítelkyni; strašně žárlil, že nevyhrál místo tebe; žárlil, jak mu to jde

-diat: deagent0: Ilesněším, když se žárlí.

-rcp: ACT-PAT: žárlil na sebe

-class: mental action

Rysunek 5. Hasło słownika walencyjnego VALLEX³⁰

3.1 Analiza korpusowa czasownika *žárlit*

Za pomocą wyszukiwarki KonText odnajdujemy poświadczenia, które analizujemy i sortujemy pod względem polskich ekwiwalentów. Przykłady wyekscerpowane z InterCorpu i w tym przypadku są ilustracją zgromadzonych definicji słownikowych:

cz. Zdálo se mi, že na pani učitelku žárlí a přišlo mi jí líto.
 pl. Wydawało mi się, że jest o panią nauczycielkę zazdrosna, i było mi jej trochę żal. [Douskova-Hrdy_Budzes]

cz. ...že nemusím na jejího muže žárlit...
 pl. ...že nie muszę być zazdrosny o jej męża... [kundera-zert]

cz. Navíc jsem žárlil na Alici, že mě nechala jen tak rozbaleného, že si mě přestala všimát, nemazlila se se mnou, nýbrž s Edvinem...
 pl. Ponadto byłem zazdrosny o Alicję, bo zostawiła mnie odwiniełego z pieluszek, bo przestała zwracać na mnie uwagę, pieściła się już nie ze mną, tylko z Edwinem... [Grusa-Dotaznik]

²⁷ na kogo, na co; z kim (przestarzała struktura narzędnikowa)

²⁸ czuć antypatię, ból, złość wobec kogoś ukochanego, ponieważ lubi kogoś innego; czuć antypatię, ból, złość wobec kogoś, kogo lubi ukochana osoba

²⁹ na kogo; czuć zawiść, wrogość wobec kogoś z powodu jego sukcesów; nienawidzić, zazdrościć komuś czegoś; nie życzyć komuś czegoś

³⁰ 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

KonText wskazał 59 poświadczeń leksemu *žárlit*. Podobnie jak w poprzednim badaniu, również w tym przypadku zaznaczyliśmy, jaki odpowiednik polski został wybrany. Po otagowaniu okazało się, iż w niemal 78% przypadków w polskim tłumaczeniu występuje fraza *być zazdrosny*, wprowadzająca odmienną strukturę składniową.

Polskie odpowiedniki czasownika <i>žárlit</i>	59
być zazdrosny	46
zazdrościć	7
zazdrość	4
Zawiść	1
Błąd	1

Tabela 4. Polskie odpowiedniki czasownika *žárlit* na podstawie materiału z czesko-polskiej części korpusu InterCorp

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:

cz. Povídám, jako vždýcky, von na mě žárlí, že jsem mladší než von.
 pl. Powiadam, jak zawsze, on **mi** zazdrości, že jestem młodszy niż on
 [Hrabal-Prilis_hl_samot]

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

cz. Právě proto, že už nechce *žárlit*, bere vážně a bez podezření jeho tvrzení!
 pl. Właśnie dlatego, że nie chce już *zazdrościć*, przyjmuje jego słowa poważnie i bez podejrzeń! [Kundera-Valcik_na_rozl]

Również w przypadku czasownika *žárlit* możemy wskazać trafny ekwiwalent w języku polskim.

być zazdrosny

- pragnący tego, co ma ktoś inny, odczuwający żal, że komuś powodzi się lepiej niż jemu
- bojący się o swoje dobro, podejrzliwie strzegący swego; zwłaszcza: podejrzliwy wobec współmałżonka, osoby kochanej

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 językowej³¹:

31 <http://sjp.pwn.pl/slowniki/zazdrosny.html> - dostęp online 15.03.2015 (pisownia oryginalna)

Mój dylemat dotyczy wyrażania zazdrości... Jestem mężatką. Wyobraźmy sobie, że 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 przedstawiona³².

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:

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

pl. Mama była Pažoutowi przez cały czas wierna, ale on mimo to *był o nią* nieustannie zazdrosny (...) [Viewegh-Roman_pro_zeny]

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

pl. Gdy się pobierali, *był zazdrosny o każdego mężczyznę*, który się do niej odezwał. [Viewegh-UcastniciZajej]

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

cs. Zdálo se mi, že na pani učitelku žárlí a přišlo mi jí líto.

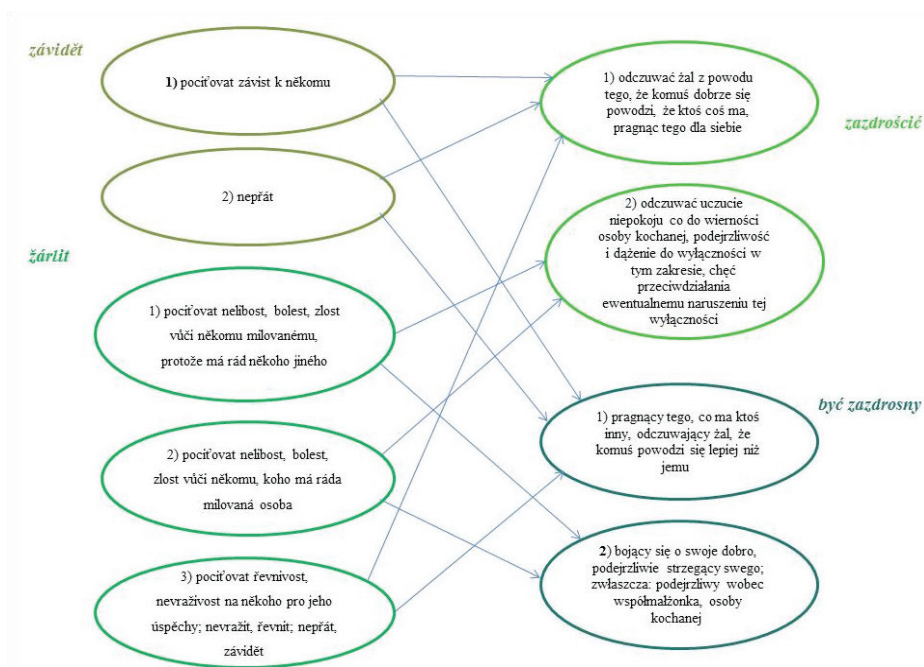
pl. Wydawało mi się, że *jest o panią nauczycielkę zazdrosna*, i było mi jej trochę żal. [Douskova-Hrды_Budzes]

4. Podsumowanie

Analiza przykładów wyekscerpowanych z InterCorpu i ich polskich odpowiedników pozwala na stworzenie sieci³³ znaczeń. 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; zdajemy sobie jednak sprawę z tego, iż na kompleksowej mapie znaczeń powinny się znaleźć również takie potencjalne

32 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) – <http://sjp.pwn.pl/slowniki/jest-zazdrosny.html> (15-03-2015).

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



Rysunek 6. Sieć powiązań znaczeń czeskich jednostek *závidět* i *žárlit* oraz polskich – *zazdrościć* i *być zazdrośny*

ekwiwalenty jak np. *zawiść*. Równoważności znaczeń zostały ustalone na podstawie manualnej analizy poświadczeń korpusowych³⁴.

Podczas analizy manualnej szczególną uwagę zwracają schematy frazy *zazdrośny o* + O (obiekt) pojawiające się jako odpowiednik wyłącznie frazy *žárlit* na + O. Ze znalezionych w korpusie 59 poświadczeń czasownika *žárlit*³⁵ 45 przekładanych jest jako *być zazdrośnym*, przy czym tylko 12 realizowanych było poprzez pełny schemat *być zazdrośnym o* + O; wśród nich 11 dotyczyło zazdrości, której obiektem była istota ludzka³⁶, a tylko w jednym przypadku obiekt był inny:

cs. 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*.
pl. 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 zazdrośna o jego tajemne ścieżki*.
[Kundera-Nesnesit_lehko]

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

35 Teksty oryginalnie czeskie w przekładzie na język polski.

36 7 razy obiektem zazdrości była kobieta (przyznał se jí konečně doma, że na ni žárlil / przyznał się wreszcie w domu, że *był o nią zazdrośny*), 4 razy mężczyzna (nemusím na jejího muže žárlit / nie muszę *być zazdrośny o jej męża*).

Rysunek 8. Łączliwość czasownika *zazdrościć* wygenerowana przez Word Sketch

Podobnie rzecz się ma z parą *žárlit* (rysunek 10.) – ponad sześć tysięcy poświad-
czeń i *zazdroсны* (rysunek 9.) – zaledwie 241.

Większa liczba wiązanych segmentów pozwoliłaby na typowe badanie
corpus driven (Hebal-Jezierska, 2008: 11-12), podczas którego na podstawie

danych korpusowych moglibyśmy określić zależność znaczenia od schematu składniowego danej jednostki (Levin, 1993; Kaczmarzka, 2015a; Kaczmarzka et al., 2015) i jeszcze dokładniej wskazać ekwiwalent dla konkretnego czasownika w danym kontekście. Dysponując tak małą liczbą poświadczeń, musieliśmy się w przypadku tych czasowników ograniczyć do badania typu corpus based (Hebal-Jezierska, 2008:12), wykorzystującego korpus w celu zilustrowania założenia przyjętego wcześniej na podstawie definicji słownikowych.



Rysunek 9. Łączliwość czasownika *zazdrosny* wygenerowana przez Word Sketch

žárlit () syn freq = **6493** (2.4 per million) Click on collocates in boldface to get multi word sketches.

post	prep	1379	10.7
na		1322	0.26

has	subj	1025	5.8
blázen		92	7.2
Krainová		5	6.15
emil		5	6.13
Alonso		5	4.26
koza		10	3.93
Pešek		4	3.51
chlap	16	25	3.14
máma	4	táta 5	
manžel	52	124	2.99
manželka	32	otec 9	partner 13
přítel	11	přítelkyně 7	
Emil		5	2.2
Anna		4	0.73

post	na	558	11.8
fanyinka		4	3.57
partnerka		6	2.75
popularita	8	44	2.36
sláva	4	úspěch 32	
přítelkyně	4	32	1.1
bratr	10	manžel 8	manželka 10
slunce		7	0.65
kolega		11	0.15

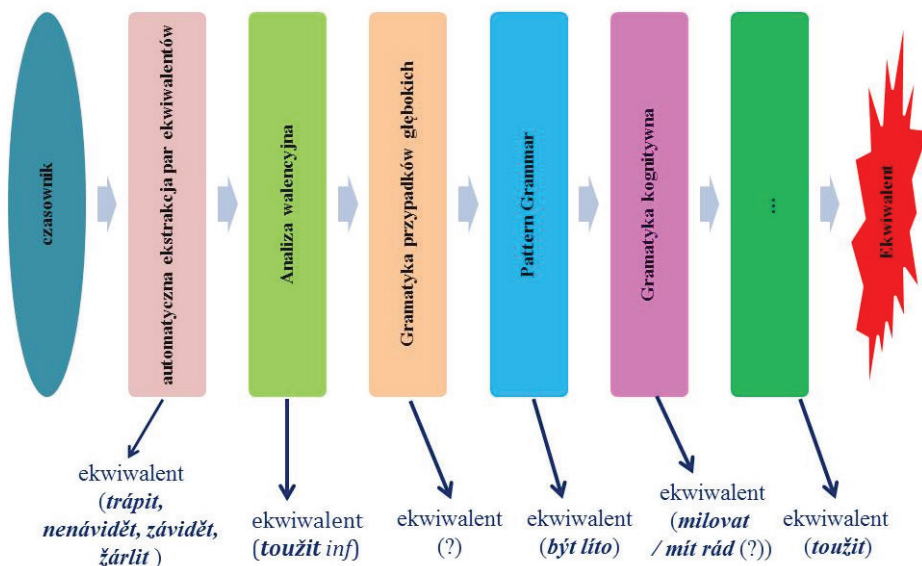
coord	303	3.4
závidět	15	4.31
odpouštět	4	3.93
nenávidět	4	2.91
zuřit	4	2.82
bít	6	2.2
milovat	5	0.62
hlídat	4	0.21

Rysunek 10. Łączliwość czasownika *žárlit* wygenerowana przez Word Sketch

5. Zakończenie

Ustalenie ekwiwalentów dla omawianych czeskich czasowników było częścią większego projektu, w ramach którego opracowywany jest algorytm ułatwiający wyszukiwanie odpowiedników czasowników wyrażających różne uczucia i emocje na podstawie danych z korpusu równoległego InterCorp (Kaczmarska, 2015a)⁴⁰. Implementując metody różnych teorii językoznawczych (zob. rysunek 11.), sprawdzamy zależność znaczenia od charakterystyki składniowej (Levin, 1993) i tym samym, na podstawie schematów składniowych, ustalamy najbliższy ekwiwalent, odnosząc się również do struktury składniowej potencjalnego odpowiednika.

40 Prace nad algorytmami ulepszającymi tłumaczenia maszynowe czy różnicującymi znaczenia jednostek wieloznacznych (np. WSD – Word Sense Disambiguation) są już na świecie prowadzone od dawna i znane także w odniesieniu do korpusów równoległych; w większości opierają się na olbrzymich danych uzyskanych z korpusów i bazują na różnych metodach matematycznych (w tym – statystycznych), np. Tian et al., 2014; Młodzki et al., 2012; Tian et al., 2010; Han et al., 2013; Kędzia et al., 2014. Opracowywane algorytmy wykorzystują też różne podejścia lingwistyczne; szerzej na ten temat – Han et al., 2013.



Rysunek 11. Schemat wypracowywanego algorytmu wspomagającego ustalanie ekwiwalentów

Jak zostało to przedstawione, cały algorytm składa się z kilku etapów analizy. Czasownik poddawany takiemu badaniu nie musi przechodzić przez wszystkie etapy; optymalny ekwiwalent może być odnaleziony na każdym poziomie analizy.

1. Pierwszym krokiem jest automatyczna ekstrakcja par ekwiwalentów. Badania pilotażowe pokazały, że już na tym etapie można znaleźć trafne ekwiwalenty wielu czasowników, co ciekawe – są to na ogół czasowniki odnoszące się do negatywnych emocji. (Kaczmarska, 2014a).

2. Drugim krokiem jest ustalenie wymagań walencyjnych. Zakładamy, iż zbieżność struktur składniowych i podobieństwo semantyczne obiektów łączących się z danymi czasownikami może być kluczem do wskazania ekwiwalentu na tym etapie. Dla niektórych czasowników znaleziono trafne ekwiwalenty tylko dla niektórych z ich znaczeń. Tak się dzieje w przypadku czasownika *toužit*⁴¹. Czasowniki, dla których nie odnaleziono ekwiwalentów na tym etapie przechodzą do etapu następnego.

41 Na podstawie wyników analizy walencyjnej i łączliwości udało się nam ustalić, iż ekwiwalentem czasownika *toužit* łączącego się z bezokolicznikiem jest polski czasownik *pragnąć* łączący się również z bezokolicznikiem (Kaczmarska, Rosen, 2013).

3. Trzecim krokiem, nieprzynoszącym, jak do tej pory, oczekiwanych rezultatów, jest identyfikowanie przypadków głębokich – roli, jakie pełnią elementy wiążące się z danym czasownikiem⁴².

4. Czwarty krok to wykorzystanie metod teorii Pattern Grammar (Ebeling, J., Ebeling S., 2013; Hunston, Francis, 2000). 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. Sprawdziło się to m.in. w przypadku frazy *být lito*. Jeżeli jednostka *být lito* łączy się z dwoma frazami nominalnymi (w celowniku i dopełniaczu), wówczas taki jej wzorzec odpowiada polskiemu ekwiwalentowi *žal*. Jeżeli natomiast *být lito* łączy się tylko z celownikową frazą nominalną (oraz ewentualnie z elementem *to*), wzorzec ten odpowiada znaczeniu polskiej jednostki (*być*) *przykro* (Kaczmarska, 2015a).

5. Kolejnym krokiem jest etap wykorzystania założeń gramatyki kognitywnej. Na tym poziomie próbujemy odkodować znaczenie słowa w kontekście konceptualizacji zjawiska nim nazwanego. W badaniach pilotażowych analizowaliśmy jednostki *mít rád* i *milovat*. O ile poradziliśmy sobie z czasownikiem *milovat*, o tyle identyfikacja znaczenia jednostki *mít rád* i jej przekład na język polski wciąż pozostają problematyczne. Metody kognitywne jednak bardzo trudno implementować do algorytmu, ponieważ przeprowadzane są w większości manualnie⁴³.

6. Ostatnie okienko to miejsce na rozwiązania przyszłościowe takie, jak np. wykorzystanie narzędzia Word Sketch dla obu języków na tym samym korpusie, czy zaawansowane badania stochastyczne⁴⁴. Wierzymy, że dzięki nim, na podstawie badania kontekstu i rozróżnienia łączliwości z różnymi obiektami,

42 Etap ten nie zostanie jednak usunięty z finalnej wersji algorytmu. Badane przez nas jednostki (będące wykładnikami różnych emocji i uczuć) są jednolite pod względem łączliwości z argumentami o pewnych wartościach – zawsze znajdziemy tu przeżywającego czy źródło i na tej podstawie nie dokonamy zróżnicowania znaczenia. Algorytm może być jednak wykorzystany w badaniu innych grup czasowników, gdzie rola semantyczna oryginału i ekwiwalentu może być istotna przy różnicowaniu znaczenia.

43 W przypadkach problemowych możemy odnieść się także do eksplikacji i naturalnego meta-języka semantycznego (Wierzbicka, 1980, 2001) lub skonstruować skalę intensywności właściwości wyrażanej przez dany czasownik (Mikołajczuk, 1997, 1999; Bratman, 1987).

44 Podczas jednego z tych badań staraliśmy się ustalić ekwiwalenty poprzez badanie liniowych i składniowych kontekstów, w których występowały czeskie czasowniki, stosując różne metody i testując teksty z korpusu równoległego. Najpierw ręcznie otagowane zostały równoległe koncordancje tych czasowników według właściwości syntaktyczno-semantycznych ich argumentów. Analiza manualna została uzupełniona przez dwujęzyczny słownik z frekwencjami. Następnie użyto klasyfikatorów stochastycznych do przewidzenia polskich ekwiwalentów na podstawie czeskiego kontekstu liniowego i składniowego. Metody były oceniane na testowych danych z korpusu z wyrównaniem słowo-do-słowa (Kaczmarska et al., 2015).

uda nam się dopasować ekwiwalenty do wszystkich znaczeń czasownika *toužit* (Kaczmarska et al., 2015), a także wszystkich innych, które nie znajdują swoich ekwiwalentów na wcześniejszych poziomach algorytmu.

Literatura

- BRATMAN, Michael E. (1987): *Intentions, Plans, and Practical Reason*. Massachusetts: Harvard University Press.
- ČERMÁK, František, ROSEN Alexandr (2012): The case of InterCorp, a multilingual parallel corpus. *International Journal of Corpus Linguistics*, 13(3), 411–427.
- EBELING, Jarle, EBELING, Signe Oksefjell (2013): *Patterns in contrast*. Amsterdam: John Benjamins.
- HAN, Aaron L.F., LU, Yi, WONG, Derek F., CHAO, Lidia S., HE, Liangye, XING, Junwen (2013): Quality Estimation for Machine Translation Using the Joint Method of Evaluation Criteria and Statistical Modeling. W: *Proceedings of the Eighth Workshop on Statistical Machine Translation*, The Association for Computational Linguistics, 365–372. <http://www.statmt.org/wmt13/pdf/WMT00.pdf>, (01-03-2016).
- HARKINS, Jean, WIERZBICKA, Anna, (red.) (2001): *Emotions in Crosslinguistic Perspective*. Berlin: Mouton de Gruyter.
- HAVRÁNEK, Bohuslav, (red.) (1989): *Slovník spisovného jazyka českého*. 2. Praha: Academia.
- HEBAL-JEZIERSKA, Milena (2008): *Wariantywność końcówek fleksyjnych rzeczowników męskich żywotnych w języku czeskim*. Warszawa: Wydział Polonistyki UW.
- HUNSTON, Susan, FRANCIS, Gill (2000): *Pattern Grammar: A corpus-driven approach to the lexical grammar of English*. Amsterdam: John Benjamins.
- KACZMARSKA, Elżbieta (2012): 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* XLVII, 247–261.
- KACZMARSKA, Elżbieta (2014a): Czeskie czasowniki oznaczające stany psychiczne – sposoby ustalania polskich ekwiwalentów na podstawie korpusu równoległego InterCorp. W: Anna STOLARCZYK-GĘBIAK, Marta WOŹNICKA (red.): *Zbliżenia. Językoznawstwo – Literaturoznawstwo – Translatologia*. Konin: Państwowa Wyższa Szkoła Zawodowa w Koninie, 45–55.

- KACZMARSKA, Elżbieta (2014b): Czy na pewno się (nie)rozumiemy? O problemach, uproszczeniach i stratach w przekładzie (na podstawie czesko-polskiej części korpusu równoległego InterCorp). W: Michala BENEŠOVÁ, Renata RUSIN DYBALSKA i Lucie ZAKOPALOVÁ (red.): *Proměny polonistiky. Tradice a výzvy polonistických studií*. Praha: Karolinum, 192–199.
- KACZMARSKA, Elżbieta (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 kognitywna). *Prace Filologiczne* LXVII, 131–150.
- KACZMARSKA, Elżbieta (2015b): *Mít rád czy milovat? O české milosti po polsku*. W: Krystyna WASZAKOWA i Marta FALKOWSKA (red.): *Pojęcia zapisane w języku*. Warszawa: Wydział Polonistyki UW, 139–156.
- KACZMARSKA, Elżbieta, ROSEN, Alexandr (2013): Między znaczeniem leksykalnym a walencją – próba opracowania metody ekstrakcji ekwiwalentów na podstawie korpusu równoległego. *Studia z Filologii Polskiej i Słowiańskiej* XLVIII, 103–121.
- KACZMARSKA, Elżbieta, ROSEN, Alexandr (2014a): Czego nie można wyrazić w języku polskim, czyli o leksykalnych w nim brakach. *Polonica* 34, 53–66.
- KACZMARSKA, Elżbieta, ROSEN, Alexandr (2014b): Praktyczny przewodnik po korpusie równoległym InterCorp. W: Milena HEBAL-JEZIERSKA (red.): *Praktyczny przewodnik po korpusach języków słowiańskich*. Warszawa: Wydział Polonistyki UW, 207–231.
- KACZMARSKA, Elżbieta, ROSEN, Alexandr (2016): Niedośłowności w dialogu czesko-polskim. W: Marcin ODELSKI, Aleksandra KNAPIK, Piotr CHRUSZCZEWSKI, Władysław CHŁOPICKI (red.): *Niedośłowność w języku*. Język i komunikacja 37. Kraków: Tertium.
- KACZMARSKA, Elżbieta, ROSEN, Alexandr, HANA, Jirka, HLADKÁ, Barbora (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.
- KĘDZIA, Paweł, PIASECKI, Maciej, KOCOŃ, Jan, INDYKA-PIASECKA, Agnieszka (2014): Distributionally Extended Network-Based Word Sense Disambiguation in Semantic Clustering of Polish Texts. W: *IERI Procedia (International Conference on Future Information Engineering)* 10, 38–44. DOI: 10.1016/j.jeri.2014.09.073
- LEVIN, Beth (1993): *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.
- LÜDTKE, Ulrike M., (red.) (2015): *Emotion in Language*. Amsterdam: John Benjamins.

- MIKOŁAJCZUK, Agnieszka (1999): *Gniew we współczesnym języku polskim. Analiza semantyczna*. Warszawa: Wydawnictwo Energeia.
- MIKOŁAJCZUK, Agnieszka (1997): Pole semantyczne 'gniewu' w polszczyźnie (Analiza leksemów: gniew, oburzenie, złość, irytacja). W: Renata GRZEGORCZYKOWA i Zofia ZARON (red.): *Semantyczna struktura słownictwa i wypowiedzi*. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego, 149–171.
- MŁODZKI, Rafał, KOPEĆ, Mateusz, PRZEPÍÓRKOWSKI, Adam (2012): Word Sense Disambiguation in the National Corpus Of Polish. *Prace Filologiczne* LXIII, 155–166.
- ROSEN, Alexandr, VAVŘÍN Martin (2014): *Korpus InterCorp – čeština, verze 7 z 19.12.2014*. <http://www.korpus.cz>, (01-03-2016)
- SIATKOWSKI, Janusz, BASAJ, Mieczysław (2002): *Słownik czesko-polski*. Warszawa: Wiedza Powszechna.
- SZYMCZAK, Mieczysław, (red.) (1995): *Słownik języka polskiego*. Warszawa: PWN.
- TIAN, Liang, WONG, Derek F., CHAO, Lidia S., OLIVEIRA, Francisco. (2014): A Relationship: Word Alignment, Phrase Table, and Translation Quality. W: *The Scientific World Journal*. Hindawi Publishing Corporation. <http://dx.doi.org/10.1155/2014/438106>, (01-03-2016).
- TIAN, Liang, WONG, Derek F., CHAO, Sam (2010): An Improvement of Translation Quality with Adding Key-Words in Parallel Corpus. W: *Machine Learning and Cybernetics (ICMLC), 2010 International Conference on*, Vol. 3, 1273–278. DOI: 10.1109/ICMLC.2010.5580888
- WIERZBICKA, Anna (1971): *Kocha – lubi – szanuje. Medytacje semantyczne*. Warszawa: Wiedza Powszechna.
- WIERZBICKA, Anna (1980): *Lingua mentalis: The semantics of natural language*. Sydney/New York, NY: Academic Press.
- WIERZBICKA, Anna (1991): *Cross-cultural pragmatics: the semantics of human interaction*. Berlin/New York, NY: Mouton de Gruyter.
- WIERZBICKA, Anna (1997): *Understanding cultures through their key words: English, Russian, Polish, German, and Japanese*. New York, NY/Oxford: Oxford University Press.
- WIERZBICKA, Anna (2001): *What Did Jesus Mean? Explaining the Sermon on the Mount and the parables in simple and universal human concepts*. New York, NY: Oxford University Press.