Translation of Judgments: A Corpus Study of the Textual Fit of EU to Polish Judgments

Tłumaczenie orzeczeń sądowych

Badanie korpusowe dopasowania tekstowego wyroków unijnych do polskich

AUTOREFERAT ROZPRAWY DOKTORSKIEJ

Dariusz Koźbiał

28 października 2020 r.

WYBÓR I ZNACZENIE TEMATU ROZPRAWY

Projekt ***Eurolekt – unijna odmiana języka polskiego i jej wpływ na polszczyznę urzędową;*** grant NCN nr 2014/14/E/HS2/00782 (Sonata BIS)

- kierownik projektu: prof. ucz. dr hab. Łucja Biel
- czas realizacji: 10.2015 – 03.2020
- stanowisko: doktorant-stypendysta

PRZEDMIOT ROZPRAWY

Dopasowanie tekstowe (*textual fit*):

relacja intertekstualna — stopień przystawalności tekstów przetłumaczonych do porównywalnych tekstów nieprzetłumaczonych w tym samym języku (Chesterman 2004, Biel 2014)

CELE I PYTANIA BADAWCZE

Cele rozprawy:

1. porównanie wyroków unijnych i wyroków polskich (**zróżnicowanie zewnętrzne**)
2. zbadanie różnic w obrębie eurolektu (**zróżnicowanie wewnętrzne**)
3. zbadanie wpływu wyroków unijnych na wyroki krajowe (**analiza europeizacji**)

Pytania badawcze:

1. Czy język wyroków unijnych przetłumaczonych na język polski różni się od języka wyroków krajowych?
2. Czy orzecznictwo TSUE wywiera wpływ na język wyroków krajowych?

MATERIAŁ BADAWCZY

	Korpus tłumaczeń unijnych		Korpus wyroków krajowych		Duży korpus referencyjny
	Wyroki TS 2011-2015	Wyroki Sądu 2011-2015	Okres przedakcesyjny 1999	Okres poakcesyjny 2011-2015	Narodowy Korpus Języka Polskiego (NKJP)
Liczba tekstów	897	384	401	2 564	b.d.
Liczba słów	4 693 968	4 515 866	504 605	6 833 997	ok. 240 000 000

METODOLOGIA I MODEL BADAWCZY

- Różnice:

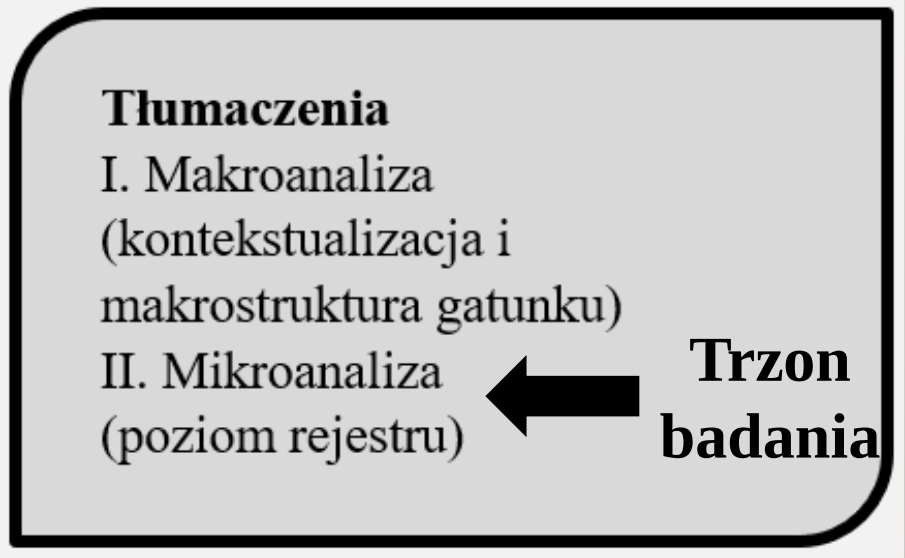
nad- i niedoreprezentowane jednostki leksykalno-gramatyczne w tłumaczeniach

- Podobieństwa:

podobna dystrybucja jednostek w tłumaczeniach i nietłumaczeniach

- Identyfikacja cech gatunkowych

Genre-register approach



METODOLOGIA I MODEL BADAWCZY

Dopasowanie tekstowe na poziomie rejestru

I. Wybrane struktury leksykalno-gramatyczne

1. przyimki proste, wtórne i złożone
2. parataksa i hipotaksa
3. deiksa
4. strona bierna i konstrukcje bezosobowe
5. leksykalne wykładniki argumentacji i schematy przyczynowo-skutkowe, warunkowe
6. okoliczniki i imiesłowy

II. Formuliczność

1. zbitki leksykalne
2. zbitki synonimiczne

III. Terminologia

1. terminy proste i złożone
2. latynizmy

STRUKTURA ROZPRAWY

Wstęp Problem badawczy i cele rozprawy	
Część I	Tło teoretyczne badania (3 rozdz.)
Część II	Materiał badawczy, wyniki, synteza (8 rozdz.)
Bibliografia	

WYNIKI – MAKROANALIZA GATUNKU

- Potwierdzenie istnienia sędziowskiej odmiany polskiego eurolektu.

- Kroki argumentacyjne stanowią główną część obu typów wyroków:
  - ok. 53,7%–80,9% – wyroki unijne (8 kroków),
  - ok. 96,6% – wyroki krajowe (4 kroki).

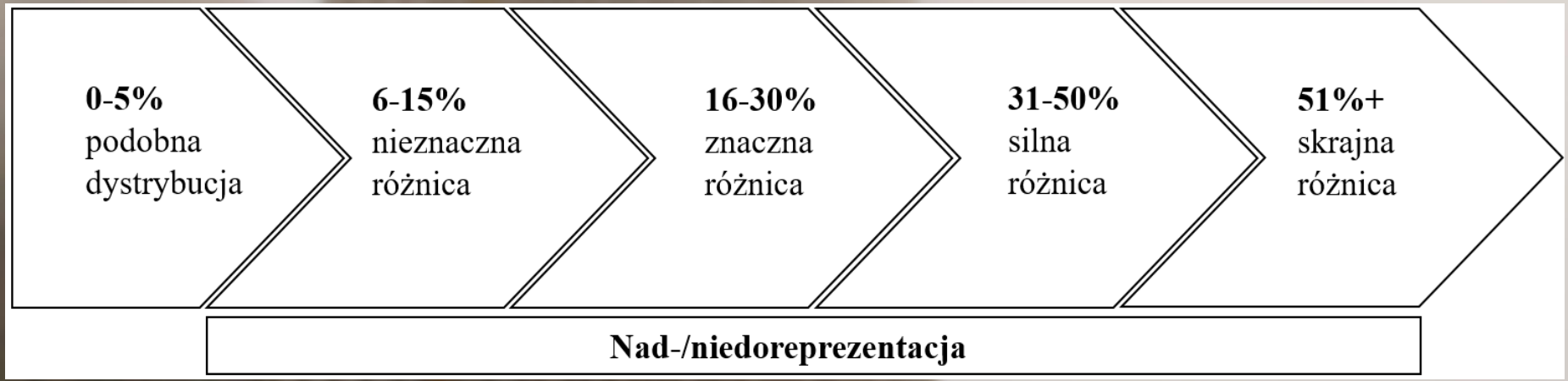
WYNIKI – MIKROANALIZA REJESTRU (I)

- Sześć kategorii na poziomie mikroanalizy:

(1) **przyimki proste, wtórne i złożone**,
(2) **hipotaksa i parataksa**,
(3) **wyrażenia deiktyczne** (zaimki osobowe, dzierżawcze, wskazujące, nieokreślone, kwantyfikatory i przymiotnikowe wyrażenia deiktyczne),
(4) **strona bierna** (czasowniki posiłkowe *być* i *zostać*) i konstrukcje bezosobowe (zaimek zwrotny *się*, czasowniki *–no/to*),
(5) **wykładniki argumentacyjne i powiązane schematy** (czasowniki werdyktywne, argumentatywne i sprawozdawcze, schematy przyczynowo-skutkowe i warunkowe),
(6) **okoliczniki i imiesłowy przymiotnikowe** (czynne i bierne) oraz **przysłówkowe** (współczesne i uprzednie).

WYNIKI – WYBRANE STRUKTURY LEKSYKALNO-GRAMATYCZNE (I)

- **Pięciostopniowa skala** służąca ocenie stopnia dopasowania tekstowego na poziomie mikroanalizy (rejestru)
- Stopień nad-/niedoreprezentacji → poziom dopasowania tekstowego



- **Niski stopień przystawalności** tłumaczeń z powodu częstej nadreprezentacji wielu jednostek leksykalno-gramatycznych

WYNIKI – WYBRANE STRUKTURY LEKSYKALNO-GRAMATYCZNE (I)

Obszary nadreprezentacji:	
Przyimki złożone i wtórne	np. *w rozumieniu*, *w przypadku*, *przeciwko* **(+12%)**
Konstrukcje hipotaktyczne	np. *że*, *iż*, *czy* **(+31,5%)**
Konstrukcje parataktyczne	np. *i*, *a także*, *jak również*, *lub*, *i/lub* **(+4,5%)**
Wyrażenia deiktyczne, w tym: • przymiotnikowe wyrażenia deiktyczne • zaimki nieokreślone i kwantyfikatory • zaimki osobowe (3. os., l. poj. i mn.)	**+47,5%** np. *rzeczony*, *niniejszy*, *wspomniany* **(+168%)** np. *wszelki*, *każdy*, *żaden* **(+139%)** np. *on*, *ona*, *oni*, *one* **(+138,5%)**
Zaimki wskazujące **Zaimki dzierżawcze**	np. *ten*, *tam* **(+26%)** np. *swój* **(+15%)**
Środki depersonalizacji, w tym: • strona bierna • czasowniki posiłkowe *być* i *zostać* • konstrukcje z zaimkiem zwrotnym *się*	**+13%** **+43,5%** **+11,5%** np. *stosuje się*, *wydaje się*, *stwierdza się* **(+2,5%)**
Czasowniki argumentatywne (3. os. l. poj., cz. ter. i przeszły) **Konstrukcje celowe** **Schematy warunkowe i powiązane konstrukcje** **Schematy przyczynowo-skutkowe**	np. *przewiduje*, *przesądza* **(+63%)** np. *aby*, *w celu*, *na celu*, *do celów*, *celem* **(+250%)** np. *jeżeli*, *jeśli*, *w przypadku*, *o ile* **(+21%)** np. *ponieważ*, *jako że* **(+5%)**
Imiesłowy **Okoliczniki**	np. *dotyczący*, *rozpatrywany*, *uwzględniając*, *zważywszy* **(+169%)**; np. *ponadto*, *co więcej*, *w szczególności*, *po pierwsze*, *w istocie*, *w rzeczywistości* **(+32%)**

WYNIKI – WYBRANE STRUKTURY LEKSYKALNO-GRAMATYCZNE (I)

Obszary niedoreprezentacji:	
Spójniki: • wprowadzające kontrast • inkluzywne	 np. *natomiast*, *ale*, *zaś* **(-45%)** np. *to znaczy*, *tzn.*, *to jest*, *tj.* **(-48,5%)**
Wykładniki argumentacji: • czasowniki werdyktywne, • czasowniki sprawozdawcze	 np. *zasądza*, *oddala*, *uchyla* **(-23%)** np. *zarzucił*, *ustalił*, *ocenił* **(-19%)**
Konstrukcje bezosobowe –*no/to*	np. *wskazano*, *stwierdzono*, *orzeczono* **(-19%)**
Przyimki proste	np. *z*, *o*, *za* **(-2,5%)**

WYNIKI – WYBRANE STRUKTURY LEKSYKALNO-GRAMATYCZNE (I)

- Zbieżności w sędziowskim eurolekcie uwypuklają dysonans między tłumaczeniami a nietłumaczeniami.
- Stopień dopasowania w obrębie eurolektu jest w znacznej mierze zbieżny w obrębie pięciu z siedmiu kategorii funkcjonalnych:
  (1) **zmiany globalne** (średnia długość słów i zdań),
  (2) **przyimki**,
  (3) **hipotaksa i parataksa**,
  (4) **wyrażenia deiktyczne**,
  (5) **środki depersonalizacji**.

WYNIKI – FORMULICZNOŚĆ (II)

- Rozbieżna formuliczność (zewnętrzna i wewnętrzna):
  - eurolekt: większa i bardziej zróżnicowana liczba typów zbitek leksykalnych oraz mniejsza liczba wystąpień.
- Niewielki zakres zbitek wspólnych.
- Rodzaje zbitek:
  - rzadkie **zbitki modalne** (np. *należy uznać, że*),
  - bardzo częste **zbitki dyskursywne** (np. *w pierwszej kolejności*) i **referencyjne** (np. *stron w postępowaniu*).
- Formuliczność w obrębie eurolektu jest rozbieżna.
- Niewielka dystrybucja zbitek synonimicznych:
  - drugorzędna rola w kształtowaniu formuliczności obu podgatunków.

WYNIKI – TERMINOLOGIA (III)

- Wyroki unijne i krajowe:
  - znaczna rozbieżność zewnętrzna i wewnętrzna,
  - różnice pod względem najczęściej wywoływanych pojęć (różna dystrybucja terminów i niski stopień pokrycia),
  - wspólna baza pojęciowa oparta na terminach prostych.
- Identyfikacja 10 terminów prostych kluczowych dla gatunku:
  sąd, strona, spółka, państwo; wyrok, prawo, decyzja, wniosek; postępowanie, koszt
- Identyfikacja 25 latynizmów kluczowych dla obu podgatunków wyroków:
  in fine, ex tunc, lex specialis, ratio legis, sine qua non, a contrario, de facto, ex post, ex nunc, erga omnes, a priori, in concreto, prima facie, sensu stricto, ad hoc, a fortiori, de minimis, etc., inter partes, quasi, lex specialis derogat legi generali, de iure, ipso iure, in abstracto, per se
- Latynizmy odgrywają znacznie większą rolę w wyrokach krajowych.

WYNIKI – CECHY GATUNKOWE

- **Nadrzędnymi cechami obu podgatunków są:**
  1. skrajna nadreprezentacja przyimków,
  2. silna nadreprezentacja konstrukcji para- i hipotaktycznych,
  3. silna nadreprezentacja okoliczników i imiesłowów,
  4. znaczna nadreprezentacja wyrażeń deiktycznych.
- **Wyroki cechuje także:**
  1. długa średnia długość zdania,
  2. znaczna dystrybucja wyrażeń deiktycznych i form bezosobowych,
  3. częste występowanie czasowników werdyktywnych (np. *orzeka*), argumentatywnych (np. *zważył*, w tym epistemicznego czasownika modalnego ***należy***) i sprawozdawczych (np. *wskazał*),
  4. znaczna formuliczność i niewielka liczba pokrywających się zbitek leksykalnych oraz niska dystrybucja zbitek synonimicznych,
  5. wspólna podbudowa pojęciowa.

WYNIKI – ANALIZA EUROPEIZACJI

- Brak bezpośredniego wpływu tłumaczeń na nietłumaczenia na poziomie:
  (1) makroanalizy gatunku,
  (2) mikroanalizy rejestru.
- Jedynie częściowa korelacja mikrodiachronicznej zmiany językowej w wyrokach krajowych z nad- i niedoreprezentacją wybranych jednostek językowych w tłumaczeniach.
- Brak jednoznacznej europeizacji wyroków krajowych przez tłumaczone wyroki unijne.

ZNACZENIE TEORETYCZNE I PRAKTYCZNE ZASTOSOWANIA

- Punkt odniesienia dla dalszych badań nad pokrewnymi gatunkami tekstów.
- Korpus JURIDIC jako narzędzie uwrażliwiające na zagadnienie dopasowania tekstowego tłumaczeń względem porównywalnych nietłumaczeń.
- Zastosowanie w kształceniu tłumaczy (*corpus-based data-driven (discovery) learning*) i w praktyce tłumaczenia.

OGRANICZENIA BADANIA I IMPLIKACJE DLA DALSZYCH BADAŃ

- Wykorzystanie wyłącznie jednojęzycznego korpusu wyroków.
- Konieczność lepszej triangulacji metodologicznej.
- Potrzeba przywiązania większej wagi do badania recepcji tłumaczeń przez ich odbiorców.

BIBLIOGRAFIA

- Biel, Ł. 2014. *Lost in the Eurofog: The Textual Fit of Translated Law*. Frankfurt am Main: Peter Lang.
- Chesterman, A. 2004. “Hypotheses about translation universals.” w: Claims, Changes and Challenges in Translation Studies. Selected contributions from the EST Congress, Copenhagen 2001. Red, Gyde Hansen, Kirsten Malmkjær i Daniel Gile, 1-13. Amsterdam: Benjamins.
- Scott, Mike. 2017. *WordSmith Tools version 7*. Stroud: Lexical Analysis Software.