

# *Analiza wielowymiarowa (MDA) polskiego Eurolektu i krajowych odmian języka administracyjnego: wymiary zróżnicowania*

An MDA analysis of the Polish Eurolect and the national variety:  
Dimensions of variation across institutional legal and administrative registers

---

ŁUCJA BIEL, KATARZYNA WASILEWSKA, DARIUSZ KOŹBIAŁ  
INSTYTUT LINGWISTYKI STOSOWANEJ  
UNIwersytet Warszawski



# Zespół EUMultiLingua

---

Łucja Biel

Dariusz Koźbiał

Katarzyna Wasilewska



# Eurolekt

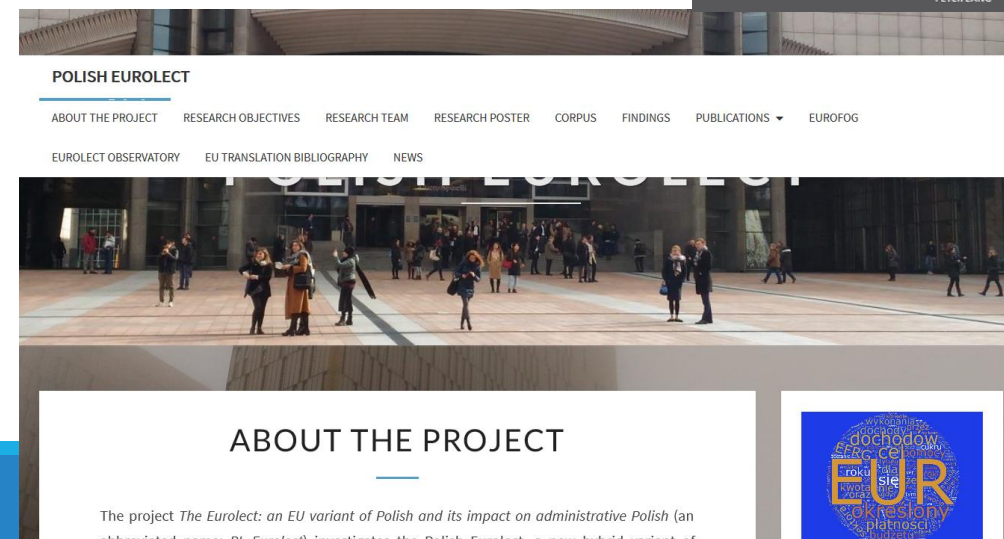
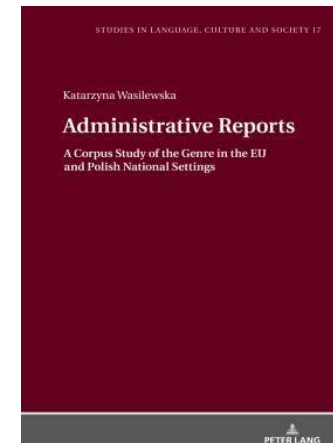
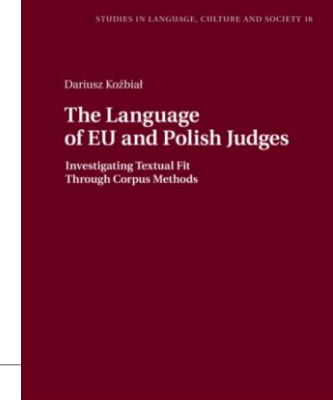
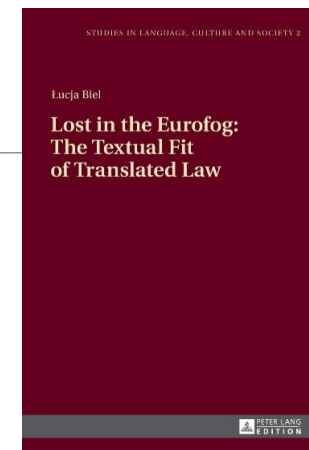


- UE – ekstremalna wielojęzyczność, 24 języki urzędowe
- Eurolekt: „*constrained language*”
  - **hybrydowa** odmiana języka administracyjnego powstająca w instytucjach UE, realizująca potrzeby językowe UE jako organizacji ponadnarodowej
  - współtworzona przez tłumaczy i prawników-lingwistów
  - Ma 24 równoległe „**lustrzane**” realizacje w językach urzędowych UE; duży wpływ EN i FR
  - Skomplikowany **wieloetapowy** i **wielojęzyczny** proces tworzenia tekstów – „**filtrowanie**” przez różne języki: “the extreme visibility of the ‘translatedness’” (Koskinen 2000).



# Badania nad polskim eurolektem

- *Jak eurolekt różni się od krajowego języka administracyjnego?*
- *Jak eurolekt jest zróżnicowany wewnątrznie?*
- *Jakie czynniki kształtują eurolekt?*
- *Jak eurolekt zmienia się w czasie?*
- *Jak eurolekt wpływa na polszczyznę administracyjną?*



# Analiza MDA

Wymiary różnicowania  
(*Linguistic Variation*)

---

# Czym jest analiza wielowymiarowa **MDA** (multidimensional analysis)?

- **Metoda statystyczna** opracowana przez Douglasa Bibera w 1988 r., początkowo do badania różnic pomiędzy językiem mówionym a pisanym, następnie rozwinięta i używana w badaniach nad różnicami pomiędzy odmianami tekstów (text varieties)

- Trzy etapy:

1) Zliczanie cech leksykalnych

2) Analiza czynnikowa

3) Identyfikacja wymiarów

- ujęcie **mikro** -> szczegółowa analiza cech językowych + ujęcie **makro** -> układ (konfiguracja) rejestrów

- **MDA Bibera** (6 wymiarów (dimensions), gotowy szablon) vs. **MDA od podstaw** (podejście indukcyjne w badaniach eksploracyjnych)

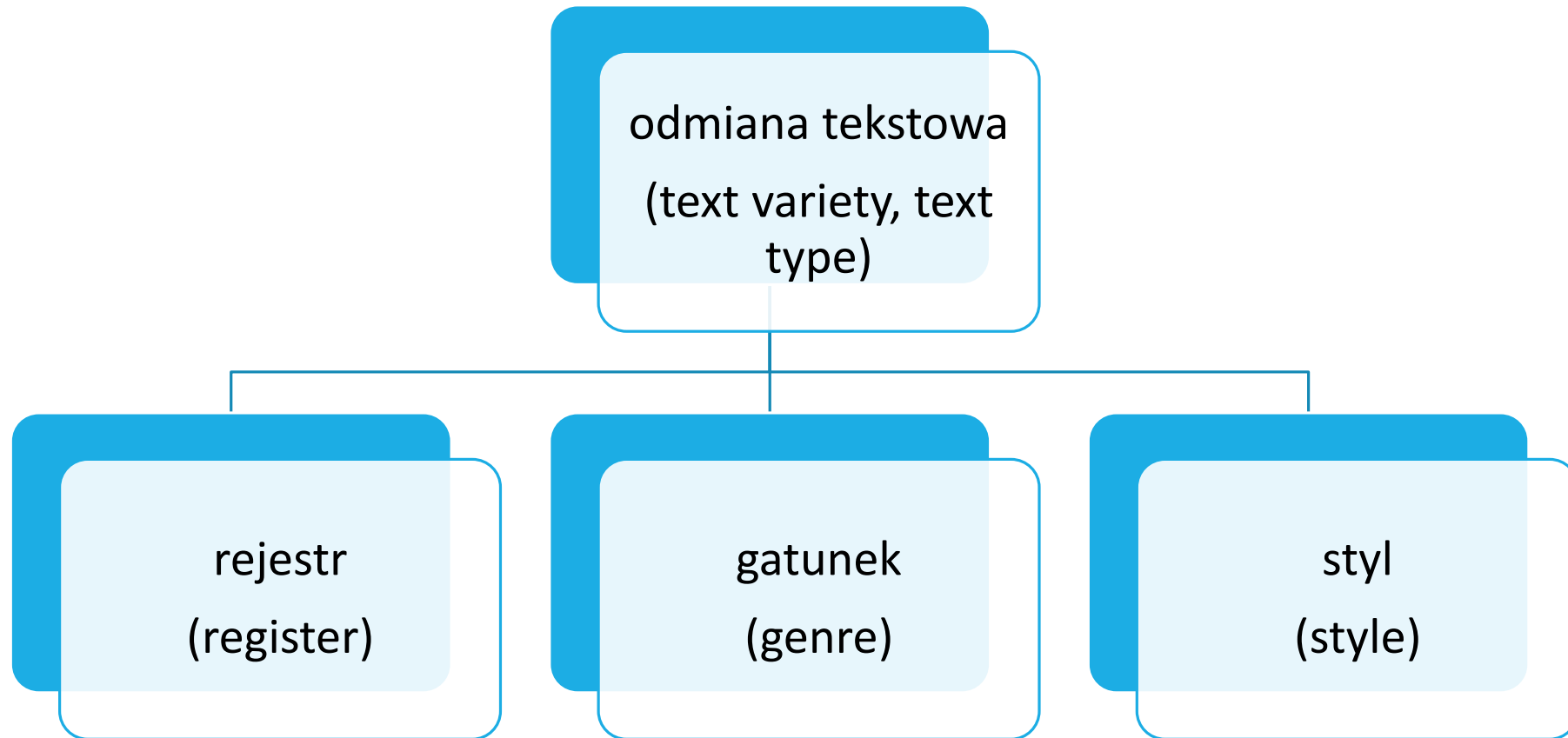
- większość badań dot. j. angielskiego lecz istnieją również w innych językach, np. koreańskim, hiszpańskim, czeskim – brak polskiego

- bardzo rzadko wykorzystywane w dyskursie instytucjonalnym i w przekładoznawstwie



# Odmiany tekstowe

---



# MDA od podstaw

---

## Procedura:

1. Przygotowanie korpusu
2. Wybór zmiennych językowych
3. Pozyskanie obserwacji z każdego tekstu
4. Wyodrębnienie czynników
5. Interpretacja czynników i przekształcenie ich w wymiary



## Narzędzia:

- **Sketch Engine** tagset NKJP (stop listy, wykorzystanie tagów, wyszukiwania przy użyciu CQL (corpus query language))
- **Lancaster Stats Tools** (Brezina 2018) <https://corpus-stats.lancs.ac.uk/>

# 1. Przygotowanie korpusu

---

## **Punkt wyjścia:**

Gatunkowy korpus porównawczo-równoległy w ramach projektu PL EUROLECT:

- 4 główne gatunki komunikacji instytucjonalnej: [akty prawne](#), [wyroki](#), [sprawozdania administracyjne](#), [strony internetowe](#)
- Teksty unijne i krajowe z okresu 2011-2015

## **Analiza wielowymiarowa MDA:**

Próbkowanie korpusu w celu otrzymania tekstów o zrównoważonej długości (1000-5000 tokenów, za Cvrček et al., 2020; 2021) aby uniknąć zniekształcenia rozkładu prawdopodobieństwa w poszczególnych podkorpusach:

- odrzucenie obserwacji odstających (outlierów; tekstów o długości ponad 3 odstępy ćwiartkowe od skrzynki na wykresie skrzynkowym) oraz dzielenie w aktach prawnych, wyrokach, sprawozdaniach;
- połączenie lub dzielenie tekstów stron internetowych

Losowy wybór próbek do progu 300K tokenów/korpus

# Korpus Eurolekt

Corpus	tokens SE	mean	median	min	max	texts
EU PL Legislation 1: Regulations	2 196 590	2 375	143	40	197 531	925
EU PL Legislation 2: Directives	885 080	9 620	6 155	130	74 611	92
Total	3 081 670	3 030	143	40	197 531	1 017
PL Legislation: Statutes	1 902 732	14 094	8 261	54	110 693	135
EU PL Judgments 1: Court of Justice	5 453 189	6 079	5 257	1 637	25 840	897
EU PL Judgments 2: General Court	5 230 696	13 622	10 258	1 508	72 568	384
Total	10 683 885	8 340	6 172	1 508	72 568	1 281
PL Judgments: Supreme Court	7 799 179	3 042	2 700	480	17 387	2 564
EU PL Reports: EU institutions	2 716 646	5 455	4 606	409	60 139	498
PL Reports: Polish institutions	2 946 220	26 784	17 045	762	155 502	110
EU PL websites: EU institutions	759 922	777	493	22	4 867	978
PL websites: Polish institutions	732 262	617	318	16	17 050	1 186

# Korpus Eurolekt MDA

<b>FINAL CORPUS 300K</b>	<b>words SE</b>	<b>tokens SE</b>	<b>mean tokens</b>	<b>median tokens</b>	<b>min tokens</b>	<b>max tokens</b>	<b>texts</b>
EU PL Legislation 1: Regulations	241 990	300 213	3 228	3 397	1 031	4 910	93
EU PL Legislation 2: Directives	247 521	300 566	3 536	3 681	1 043	4 992	85
PL Legislation: Statutes	221 798	300 124	4 168	4 275	1 151	4 996	72
EU PL Judgments 1: Court of Justice	236 795	300 531	3 536	3 481	2 508	4 928	85
EU PL Judgments 2: General Court	236 523	300 125	4 002	4 021	2 917	4 987	75
PL Judgments: Supreme Court	238 818	300 236	2 566	2 501	1 065	4 892	117
EU PL Reports: EU institutions	246 854	300 142	3 450	3 566	1 064	4 865	87
PL Reports: Polish institutions	220 571	301 507	4 568	4 718	3 062	4 999	66
EU PL websites: EU institutions	250 306	301 177	2 215	2 181	1 024	3 817	136
PL websites: Polish institutions	247 784	300 708	2 387	2 382	1 065	4 650	126

## 2. Wybór zmiennych językowych

- Brak wcześniejszych MDA dla j. polskiego  
→ Standardowe zmienne MDA dopasowane do j. polskiego + kategorie na podstawie funkcji dyskursywnych opartych na naszych wcześniejszych badaniach (np. Biel, 2014; Koźbiał, 2020; Wasilewska, 2021)

- Najlepiej min. 5 razy więcej tekstów niż zmiennych – cech językowych (Biber, 1988: 65)

→ 92 zmienne – 942 teksty

Częstotliwość i rodzaje części mowy

**Rzeczowniki** odczasownikowe

**Czasowniki:** czasowniki według czasów i osób; czasowniki modalne; imiestłowy

**Bezosobowość:** strona bierna, konstrukcje bezosobowe

**Zaimki:** osobowe, nieokreślone, wskazujące, uogólniające

**Spójniki** podrzędne i współrzędne

**Przymyki złożone** według funkcji

**Liczebniki**

**Typy zdań:** rozkazujące, pytające, wykrzyknikowe

**Nawigacja tekstowa:** wyrażenia deiktyczne, nawigacyjne i organizujące tekst

**Zdania warunkowe i konstrukcje hipotetyczne**

Wykładniki **celu**

Wykładniki **uzasadnienia**

Wykładniki **przyczynowości i wnioskowania**

Wykładniki **przyzwolenia i kontrastu**

Wykładniki **negacji**

Wykładniki **dodawania i alternatywy**

Wykładniki **wyjaśniania i egzemplifikacji**

**Konstrukcje zakotwiczące**

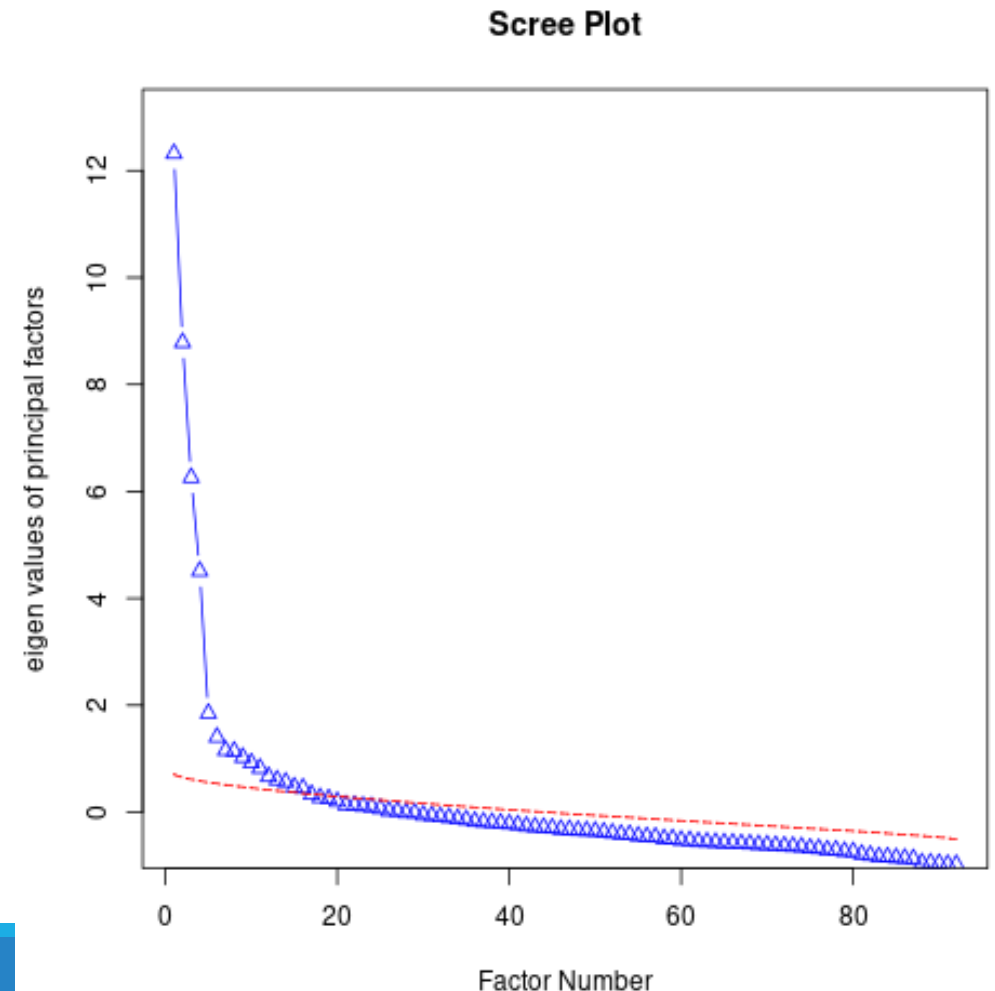
**Przysłówki modalne postawy**

# 3. Pozyskanie obserwacji z każdego tekstu

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	FilenameABI	Filename	Register	N_f	N_NF	N_t	V_f	V_NF	V_t	Adj_f	Adj_NF	Adj_t	Adv_f	Adv_NF	Adv_t	Con_f
2	EU_L_R_1	2011R1012011PL	EU_legis_R	635	38,77	226	96	5,86	46	410	25,03	130	26	1,59	13	178
3	EU_L_R_2	2011R10772011PL_1	EU_legis_R	1208	43,58	279	178	6,42	77	557	20,09	150	54	1,95	15	173
4	EU_L_R_3	2011R11692011PL_2	EU_legis_R	1297	39,29	308	223	6,76	67	752	22,78	221	64	1,94	29	215
5	EU_L_R_4	2011R1812011PL_1	EU_legis_R	901	42,52	240	135	6,37	44	458	21,61	171	42	1,98	18	177
6	EU_L_R_5	2011R2112011PL	EU_legis_R	1111	40,83	246	229	8,42	107	608	22,34	159	56	2,06	18	144
7	EU_L_R_6	2011R3272011PL	EU_legis_R	535	42,16	169	93	7,33	36	270	21,28	118	16	1,26	11	60
8	EU_L_R_7	2011R4922011PL	EU_legis_R	1364	43,01	336	212	6,69	85	719	22,67	198	45	1,42	24	204
9	EU_L_R_8	2011R632011PL	EU_legis_R	598	42,68	154	96	6,85	48	294	20,99	104	20	1,43	9	79
10	EU_L_R_9	2011R6772011PL_1	EU_legis_R	1099	45,75	278	149	6,20	57	573	23,86	169	35	1,46	16	130
11	EU_L_R_10	2011R6772011PL_2	EU_legis_R	1122	48,74	267	140	6,08	64	530	23,02	166	18	0,78	13	111
12	EU_L_R_11	2011R6922011PL	EU_legis_R	756	42,91	205	114	6,47	52	382	21,68	142	34	1,93	16	88
13	EU_L_R_12	2012R10242012PL	EU_legis_R	1621	44,34	301	252	6,89	89	802	21,94	200	82	2,24	31	193
14	EU_L_R_13	2012R10262012PL	EU_legis_R	560	38,17	174	94	6,41	43	390	26,58	132	25	1,70	11	93
15	EU_L_R_14	2012R11512012PL_1	EU_legis_R	1254	38,63	283	222	6,84	84	757	23,32	223	53	1,63	22	272
16	EU_L_R_15	2012R11512012PL_2	EU_legis_R	1136	38,59	244	216	7,34	78	729	24,76	186	82	2,79	24	182
17	EU_L_R_16	2012R11512012PL_3	EU_legis_R	1279	40,68	260	253	8,05	105	681	21,66	180	58	1,84	22	201
18	EU_L_R_17	2012R12152012PL_2	EU_legis_R	1454	38,11	254	335	8,78	110	858	22,49	187	75	1,97	24	217
19	EU_L_R_18	2012R12572012PL	EU_legis_R	992	40,23	229	141	5,72	63	646	26,20	145	41	1,66	17	104
20	EU_L_R_19	2012R2072012PL	EU_legis_R	595	40,61	171	99	6,76	43	308	21,02	114	32	2,18	18	91
21	EU_L_R_20	2012R2362012PL_3	EU_legis_R	1362	40,75	276	262	7,84	97	708	21,18	184	55	1,65	20	248
22	EU_L_R_21	2012R2672012PL_1	EU_legis_R	1362	41,05	324	158	4,76	54	674	20,31	167	64	1,93	13	391
23	EU_L_R_22	2012R2672012PL_3	EU_legis_R	1170	37,95	297	180	5,84	71	697	22,61	191	60	1,95	20	360
24	EU_L_R_23	2012R3892012PL_1	EU_legis_R	1184	39,70	243	211	7,08	73	807	27,06	159	56	1,88	17	169
25	EU_L_R_24	2012R5472012PL	EU_legis_R	361	41,64	130	50	5,77	25	194	22,38	85	6	0,69	6	26

# 4. Wyodrębnienie czynników

- Na podstawie wykresu osypiska – punkt załamania (kolano)
  - ~5-6 (wartość własna poniżej 2)
  - Czynniki 5 i 6 pominięto, gdyż ich siła była niewystarczająca do rzetelnej interpretacji
- Ostatecznie 4 wymiary



# 5. Interpretacja czynników i przekształcenie ich w wymiary

---

- I wymiar różnicowania: **argumentacja vs. informacja**
- II wymiar różnicowania: **angażujący instruktaż vs. zdystansowany autorytet**
- III wymiar różnicowania: **preskryptywność vs. narracja**
- IV wymiar różnicowania: **bogactwo leksykalne**

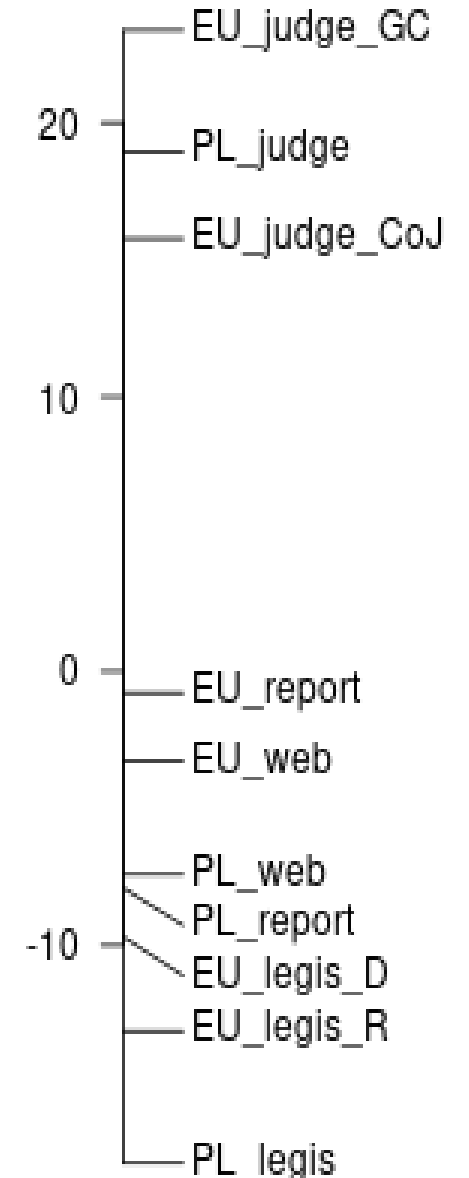
# I wymiar różnicowania: Argumentacja vs. informacja

- Zmienne o wysokich dodatnich współczynnikach (wskaźniki przyczynowości służące wnioskowaniu i uzasadnianiu, wzorce nadające strukturę argumentacji i rozumowaniu, wykładniki postaw/stanowisk) → silnie oceny/wartościujący charakter dyskursu poparty obiektywnym, bezosobowym autorytetem
- Zmienne o ujemnych współczynnikach (rzeczowniki, przymiotniki, nawigacja tekstowa) → tekst bogaty w informacje, zawierający różnicowaną terminologię oraz odniesienia międzytekstowe i wewnątrztekstowe

Code	Variable	Factor loadings
InferMark	Causal markers expressing inference	0.960
InferenceALL	Causal markers expressing inference and cause-effect prepositions (total)	0.910
Justification	Causal markers expressing justification	0.881
Con_t	Conjunctions – types	0.754
AdversativeALL	Contradictory prepositions and adversative coordinators (total)	0.712
Neg	Negation markers	0.668
Należy	Impersonal modal verb <i>należy</i> “should”	0.665
PronDef	Demonstrative pronouns	0.657
Para_adv_coor_adv	Adversative coordinators and adverbials	0.597
TextOrg	Text organization (metadiscursive) markers	0.591
Past3rd	Past tense, 3 <sup>rd</sup> person	0.583
PronDefNP	Demonstrative pronouns with nouns	0.574
CPrep_EpMod	Epistemic modality prepositions	0.572
CauseEffectPrep	Cause and effect prepositions	0.547
SubExclAby	Subordinators	0.522
StanAdvHigh	High-level stance adverbials	0.514
Nmoc_Nmozna	Possibility/permission modal verbs <i>móc można</i> “may” with negation	0.484
CPrep_Contra	Contradictory prepositions	0.481
StanAdvALL	Stance adverbials (total)	0.480
PresPartic	Present participle <i>-ąc</i>	0.444
PersPron3	Personal pronouns, 3 <sup>rd</sup> person	0.440
ConcSub	Concessive subordinators	0.357
ConcessionALL	Concessive subordinators and propositions (total)	0.354
TextNav	Textual navigation patterns referring to legal provisions	-0.354
Para_alt_coor_con	Alternative coordinating conjunctions	-0.395
N_f	Nouns – frequency	-0.399

# I wymiar zróżnicowania: Argumentacja vs. informacja

- Wyroki Sądu (rozstrzyganie sporów) są najbardziej argumentacyjne; wyroki Trybunału Sprawiedliwości (zawierające wytyczne dotyczące wykładni prawa UE) są nieco mniej argumentacyjne, ale zachowują zdecydowany ton → styl deklaratywny.
- Dyrektywy UE są bogatsze w argumentację niż polskie ustawy; polskie ustawodawstwo ma charakter głównie informacyjny (brak uzasadnień).
- Strony internetowe plasują się mniej więcej pośrodku.
- Teksty unijne są ogólnie bogatsze w argumentację niż polskie.



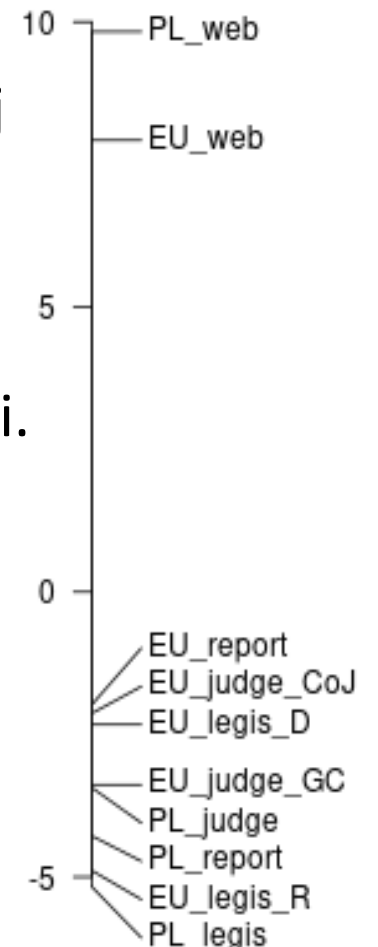
# II wymiar różnicowania: angażujący instruktaż vs. zdystansowany autorytet

- Zmienne o wysokich dodatnich ładunkach (czasowniki w 2. osobie, zaimki osobowe, tryb rozkazujący, konstrukcje warunkowe) → **zindywidualizowane przewodnictwo**, pozycjonujące instytucję jako doradcę:
  - bezpośrednio zwracanie się do odbiorcy i indywidualizowanie komunikacji
  - skracanie dystansu instytucjonalnego (obecność nadawcy)
  - tworzenie dynamicznego, zorientowanego na działanie stylu
  - angażowanie odbiorcy poprzez dialogiczność
- Zmienne o ujemnych ładunkach (strona bierna, deiksa tekstowa) → **zdepersonalizowany, autorytatywny styl**, służący budowaniu autorytetu prawnego, kondensacji informacji oraz rozstrzygnięciu konfliktów między przepisami. Ekspozuje on **autorytet, racjonalność i legitymizację**.

Code	Variable	Factor loadings
Pres2nd	Present tense, 2 <sup>nd</sup> person	0.886
PersPron2	Personal pronouns, 2 <sup>nd</sup> person	0.859
Cond_conj	Conditional conjunctions, e.g. <i>jeśli</i> "if"	0.853
Fut2nd	Future tense, 2 <sup>nd</sup> person	0.824
V_f	Verbs – frequency	0.766
ConditionalsALL	Conditional patterns (total)	0.755
Impt	Imperative	0.699
Musi	Modal verb <i>musieć</i> "must"	0.678
ModalsALL	Modal verbs (total)	0.599
Fut3rd	Future tense, 3 <sup>rd</sup> person	0.537
Past2nd	Past tense, 2 <sup>nd</sup> person	0.528
PersPron1	Personal pronouns, 1 <sup>st</sup> person	0.522
InterrogCl	Interrogative clauses	0.512
Past1st	Past tense, 1 <sup>st</sup> person	0.434
Pres1st	Present tense, 1 <sup>st</sup> person	0.390
AdjPartPas	Passive adjectival participles	-0.354
TextALL	Text deixis, navigation, and organisation (total)	-0.377

## II wymiar różnicowania: angażujący instruktaż vs. zdystansowany autorytet

- Strony internetowe są najbardziej zaangażowane, zwłaszcza polskie strony rządowe; strony unijne są nieznacznie bardziej formalne i mniej spersonalizowane.
  - Akty prawne są skrajnie zdystansowane, z wyjątkiem bardziej elastycznych dyrektyw.
  - Sprawozdania UE są bardziej angażujące niż ich polskie odpowiedniki.
  - Orzeczenia Trybunału (w post. prejudycjalnych) wykazują pewną dialogiczną, ponieważ Trybunał wchodzi w interakcję z organami krajowymi.
- 
- **Większość komunikacji instytucjonalnej pozostaje jednak zdystansowana.**



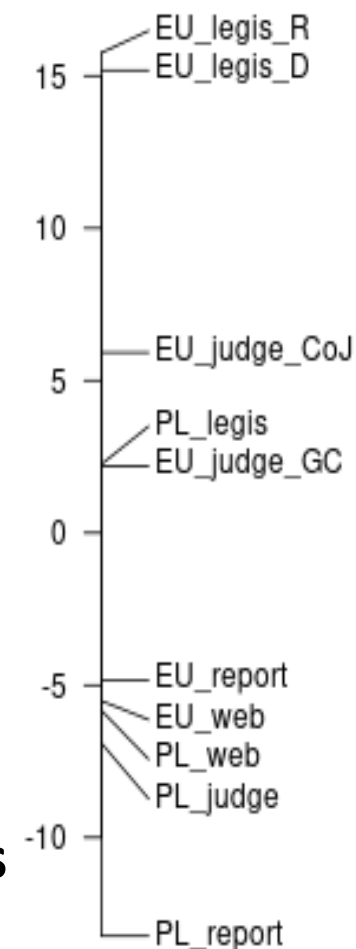
# III wymiar różnicowania: preskryptywność vs. narracja

- Ukazuje kontrast między stylem pisania zorientowanym na teraźniejszość i przyszłość, charakteryzującym się precyzją odniesień i normatywnością a bardziej opisowym i narracyjnym stylem zorientowanym na przeszłość
- Zmienne o wysokich dodatnich ładunkach (przyimki deiksy tekstowej, schematy nawigacyjne, imiesłowy bierne) → precyzyjne odniesienia prawne, umacnianie autorytetu prawnego, legitymizacja normatywnego charakteru tekstów, funkcje performatywne
- Zmienne o ujemnych ładunkach (wysoka częstotliwość rzeczowników) → narracja opisowa zorientowana na przeszłość, teksty bogate w informacje; dyskurs ma charakter opisowy, a nie nakazowy.

Code	Variable	Factor loadings
TextALL	Text deixis, navigation, and organization (total)	0.763
CPrep_Deixis	Text deixis prepositions	0.658
Para_alt_coor_con	Alternative coordinating conjunctions	0.621
Pres3rd	Present tense, 3 <sup>rd</sup> person	0.618
TextNav	Textual navigation patterns referring to legal provisions	0.587
Cond_qual_cl	Qualifying conditional clauses, e.g. <i>w przypadku</i> “in the case of” with complementizers	0.559
DeixisNP	Adjectival deictic markers with nouns	0.552
Cond_fram_cl	Framing conditional clauses, e.g. <i>pod warunkiem</i> “on condition” with complementizers	0.520
AdjPartPas	Passive adjectival participles	0.494
PassiveALL	Passives (total)	0.492
ConditionalsALL	Conditional patterns (total)	0.473
PronDefNP	Demonstrative pronouns with noun	0.471
SubExclAby	Subordinators	0.470
Adj_f	Adjectives – frequency	0.394
ImpSie	Impersonal present tense pattern “się”	0.387
PronGen	Generalizing pronouns	0.385
Con_f	Conjunctions – frequency	0.381
PronDef	Demonstrative pronouns	0.358
ImprsALL	Impersonal patterns (total)	0.353
Past3rd	Past tense – 3 <sup>rd</sup> person	-0.365
N_f	Nouns - frequency	-0.528

### III wymiar różnicowania: preskryptywność vs. narracja

- Akty prawne UE mają najbardziej preskryptywny charakter spośród wszystkich rodzajów tekstów → silna **preskryptywność i legitymizacja** jako cechy charakterystyczne stylu redagowania w ramach UE; polskie akty prawne mają mniej preskryptywny charakter w porównaniu z aktami prawnymi UE
  - Polskie wyroki mają silnie narracyjny charakter ze względu na sposób przedstawiania historycznych aspektów spraw; wyroki UE mają bardziej preskryptywny charakter, ponieważ mają charakter deklaratywny.
  - **Hybrydowość wywołana poprzez proces tłumaczenia**; niektóre cechy o wysokich ładunkach dodatnich (np. zaimki wskazujące, forma bierna) są traktowane jako interferencja z języka angielskiego.
- **Unijny styl redagowania skłania się bardziej ku preskryptywności, podczas gdy krajowy styl redagowania bardziej ku opisowości i przytaczaniu faktów**



# IV wymiar zróżnicowania: bogactwo leksykalne

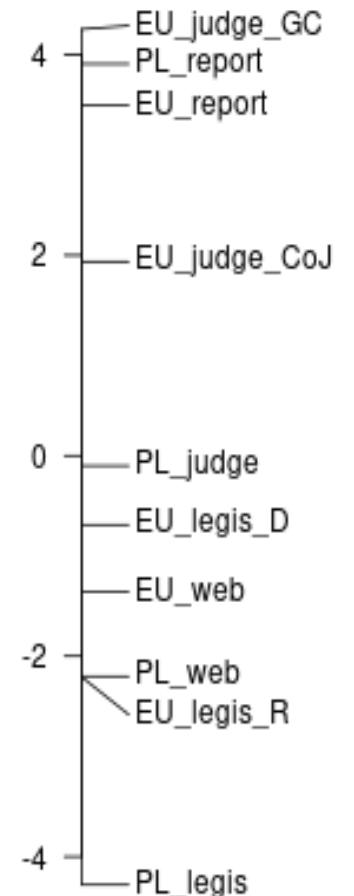
---

- Miara bogactwa leksykalnego
- Wysokie bogactwo leksykalne = wiele typów rzeczowników/przymiotników.
- Niskie bogactwo leksykalne = formuliczny język oraz utarte schematy językowe o ograniczonym wykorzystaniu synonimów

Code	Variable	Factor loadings
Adj_t	Adjectives – types	0.981
N_t	Nouns – types	0.904
Adv_t	Adverbs – types	0.748
V_t	Verbs – types	0.741
Prep_t	Prepositions – types	0.699
Con_t	Conjunctions – types	0.351

# IV wymiar różnicowania: bogactwo leksykalne

- Sprawozdania (UE + PL) cechują się największym różnicowaniem leksykalnym; wyroki również osiągają wysokie wartości ze względu na różnicowaną tematykę i argumentację.
- Akty prawne wykazują najniższy poziom bogactwa leksykalnego — formuliczny charakter.
- Strony internetowe wykazują zaskakująco niski poziom, mimo większej swobody sprzyjającej różnicowaniu.
- Teksty unijne są ogólnie bardziej różnicowane leksykalnie niż ich polskie odpowiedniki → tłumaczenie wprowadza niepożądane różnicowanie leksykalne w instytucjonalnych rejestrach prawnych.



# Wnioski

---

MDA skutecznie **odwzorowuje różnicowanie funkcjonalnie zbliżonych dyskursów instytucjonalnych**: wskazuje na odmienne sposoby komunikacji instytucji ponadnarodowych i krajowych → Teksty unijne cechują się **hybrydowością wywołaną przez proces tłumaczenia** i różnią się od polskich odpowiedników

- Akty prawne i orzeczenia UE charakteryzują się większym stopniem normatywności, częstszym odwoływaniem się do przepisów prawnych oraz bardziej argumentacyjną strukturą w porównaniu z polskimi odpowiednikami. Strony internetowe UE zawierają mniej elementów angażujących odbiorcę (*engagement*) i strategii wyjaśniających, natomiast sprawozdania UE preferują mniej zdystansowany styl.
- Teksty UE mogą wydawać się polskim odbiorcom bardziej zdystansowane, schematyczne i mniej naturalne
- Tam, gdzie jest to możliwe, tłumaczenia unijne powinny w większym stopniu być dostosowane do krajowych konwencji prawnolingwistycznych, aby poprawić ich naturalność, zrozumiałość i zwiększyć zaangażowanie odbiorców

# Ograniczenia

---

- MDA wymaga podjęcia szeregu subiektywnych decyzji metodologicznych, np., dobór zmiennych, liczba wyodrębnianych czynników czy interpretacja wymiarów (cf. Goulart & Wood 2021) → researcher bias
- Analiza oparta na jednym języku ogranicza możliwość porównań międzyjęzykowych.
- Dla języka polskiego nie ma dedykowanego tagera MDA, a ogólne tagery POS są mniej precyzyjne w przypadku złożonych wielowarstwowych kategorii.
  - Lepsze, opracowane w przyszłości narzędzia mogłyby zwiększyć precyzję i porównywalność badań.

# Thank you!

---



Comments?  
Suggestions?  
Questions?



References: Biel, Wasilewska, Koźbiał (forthcoming, 2026) „Dimensions of variation across institutional legal and administrative registers: An MDA analysis of the Polish Eurolect and the national variety.” *International Journal of Corpus Linguistics*.