

EWA MICHALSKA

Comparative Statistics of Systemic and Structural Differences in the Czech and Polish Legal Codes

Abstract: This article offers a quantitative and qualitative comparison of three groups of Czech and Polish legal acts: constitutions, civil codes and commercial company legislation. Linguistic kinship between Czech and Polish does not entail stylistic kinship of legal texts. The Polish Constitution is markedly more extensive and axiological, whereas the Czech Constitution remains closer to a framework model. This relation is reversed at the level of the civil codes. The differences reflect distinct legislative strategies and have consequences for legal translation and the teaching of languages for specific purposes.

Keywords: legal linguistics; comparative statistics; Czech law; Polish law; constitution; civil code; commercial companies; legal translation; legislative style; corpus linguistics

1. Introduction

A quantitative comparison of the Czech and Polish legal systems is not a merely technical exercise. It is a cognitive instrument that makes it possible to grasp differences in the way law is conceived, in the degree of regulatory precision, and in the relationship between the constitution and ordinary legislation. Despite close linguistic kinship and a shared heritage of continental civil law, Czech and Polish legal acts differ markedly in length, level of detail, and textual organisation.

The present analysis focuses on three key groups of acts: constitutions, civil codes, and statutes governing commercial companies. The aim is not only to juxtapose figures, but also to show how quantitative differences translate into specialist translation practice, the teaching of legal language, and the understanding of distinct legal cultures. In this sense, comparative statistics become a productive method of inquiry rather than a purely formal description. The argument developed here is closely connected with my other research on legal collocations, false friends and the Latin layer of Czech-Polish legal terminology, which I am preparing in a larger forthcoming book on the subject.

1.1 Corpus-based approaches to legal language

The comparison of Czech and Polish legislation can be situated within the wider field of corpus-based legal linguistics. In the German context, Vogel, Hamann and Gauer (2018) introduced computer-assisted legal linguistics and the Ju-Reko reference corpus, bringing legislation, judicial decisions and legal scholarship into a common empirical framework. More recent methodological discussions and case studies are collected by Vogel, Walter and Tripps (2022). These approaches provide a framework for investigating recurrent linguistic patterns in legal texts while distinguishing observations about usage from arguments about legal interpretation.

Austrian research offers complementary examples of legal text processing. The BRISE-plandok corpus developed by Recski et al. (2025; first published online in 2024) contains manually annotated German-language planning regulations from Vienna. Its relevance to the present discussion lies in the explicit documentation of corpus selection and annotation, rather than

in a direct comparison of national legislative styles. Filtz, Kirrane and Polleres (2021), working in legal informatics, address the organisation and interlinking of Austrian legal data. These projects illustrate different research purposes and should not be treated as interchangeable with a representative corpus of legal language.

For Czech, Glogar (2025) presents KoPr and discusses the design of a representative legal-language corpus encompassing several legal genres. In the Polish context, Biel, Wasilewska and Koźbiał (2026) apply multidimensional analysis to institutional legal and administrative registers, comparing the Polish Eurolect with the national variety. For a Czech–Polish comparison, this line of research motivates attention to register, institutional setting and the distinction between domestic legislation and translated EU texts. It also suggests that national language alone is too broad a category to explain textual variation.

The translation and teaching implications require a similarly differentiated approach. Leńko-Szymańska and Biel (2023) examine terminological collocations in trainee and professional translations of company-law material, using comparable non-translated documents as an additional point of reference. Kontutytė (2025) investigates corpus work in the teaching of German legal language, drawing on GeLeCo. These studies offer concrete models for moving from general claims about translation difficulty to the analysis of phraseological choices and learning tasks. The broader terminology framework is developed in Biel and Kockaert (2023), while Graziadei (2025) discusses legal translation from the perspective of comparative law. Wright (2026) provides an overview of current corpus perspectives on legal discourse.

Against this background, the contribution of the present article is best framed as an exploratory comparison of the textual organisation of selected Czech and Polish statutes. Structural counts can identify differences worth investigating, but they do not by themselves establish semantic density, greater legal precision or the extent of judicial discretion. A fuller account would combine matched legal subject areas, documented corpus queries and qualitative analysis of the relevant provisions.

2. Methodology and the Cognitive Status of Quantitative Comparison

The analysis is based on current texts:

- the Constitution of the Czech Republic (*Ústava České republiky*, 1992/1993, as amended);
- the Constitution of the Republic of Poland (*Konstytucja Rzeczypospolitej Polskiej*, 1997, as amended);
- the Czech Civil Code (*Občanský zákoník*, Act No. 89/2012 Coll., in force since 2014);
- the Polish Civil Code (*Kodeks cywilny*);
- the Czech Business Corporations Act (*Zákon o obchodních korporacích*, Act No. 90/2012 Coll.);
- the Polish Commercial Companies Code (*Kodeks spółek handlowych*).

The quantitative data are approximate and serve comparative purposes. They describe structure (number of chapters, articles or paragraphs) rather than semantic content. Where possible, a corpus perspective has also been taken into account, in particular KoPr, the Representative Corpus of the Czech Legal Language.¹ A longer code is not automatically a better code (or even more thorough, for that matter). Length and density are, however, reliable indicators of a legislator's preferred strategy: whether precision is sought in the constitution, in the code, or in case law.

Why is such a comparison productive? First, it reveals differences in legislative density that remain invisible in a purely qualitative reading. Second, it makes it possible to link constitutional style with the style of the codes that follow it. Third, it opens a space for translational and didactic reflection: both the translator and the student of legal language need to know whether they are dealing with a synthetic or a casuistic text. Finally, quantitative comparison provides empirical support for claims formulated in legal linguistics, including observations developed in my other studies on collocations, false friends (*faux amis*) and the Latin heritage of the two languages.²

¹ Ondrej Glogar, Representative Corpus of the Czech Legal Language: A Tool for Jurisprudence, *International Journal for the Semiotics of Law* (2025).

² Cf. also Lucja Biel on legal translation and EU legal language, and the programme of the journal *Comparative Legilinguistics*.

Methodological qualification. The structural comparisons presented here should be distinguished from corpus-based measurements of lexical frequency and from measures of normative complexity. Article and section numbering is not a uniform unit of textual length across legal systems; amendments may introduce additional numbered provisions or repeal existing ones. The interpretation of structural differences must therefore take account of the substantive scope of each act. KoPr provides a methodological reference (Glogar 2025), but reproducible corpus queries are not reported in the present article. Claims about lexical frequency, legislative density and translation effects should accordingly be treated as exploratory hypotheses requiring further empirical testing.

3. Comparing the Constitutions: Polish Detail versus Czech Conciseness

Table 1: Structure of the Czech and Polish constitutions

Parameter	Ústava ČR	Konstytucja RP
Chapters	8	13
Articles	c. 113	243
Preamble	short, declarative	extensive, axiological
Density	more synthetic	more detailed

The Polish Constitution of 1997 is clearly more extensive. This difference is results from several historical and political factors. After 1989, Poland underwent an intensive process of building the legitimacy of a democratic state governed by the rule of law. The expanded preamble and the chapter on the freedoms and rights of persons and citizens (Arts. 30-86) had not only a normative but also a symbolic function: they recorded a break with the previous regime. The Czech Constitution, adopted in the circumstances of the dissolution of the federation, prefers conciseness and leaves more room for ordinary legislation and case law.³

³ Michele Graziadei, Legal Translation and the Quest for Authenticity, International Journal for the Semiotics of Law (2025).

A further reason for Polish constitutional amplitude lies in the decision to write a relatively complete catalogue of rights, sources of law and institutional competences into the constitutional text itself. The Czech model is closer to a framework constitution: many questions of rights are developed in the Charter of Fundamental Rights and Freedoms, which forms part of the constitutional order but is a separate act. The Polish text absorbs more of this material into a single document. For the translator this means that rendering the Polish constitution into Czech often requires compression, whereas translation in the opposite direction requires specification and the avoidance of excessive generality.

The difference also has a didactic value. Students of legal Czech and legal Polish easily assume that two closely related languages will produce similarly organised constitutional texts. Table 1 shows that this assumption is false. What looks like a simple equivalent pair - constitution / ustava / konstytucja - conceals two different textual economies.

4. Civil Codes: Continuation and Reversal of Constitutional Style

Table 2: Structure of the civil codes

Parameter	Občanský zákoník	Kodeks cywilny
Parts / books	5 parts	4 books
Paragraphs	c. 3,080	c. 1,080–1,250
Length	substantially greater	smaller

The Czech Civil Code (the so-called new OZ) is one of the most detailed civil codes in continental Europe. Its construction testifies to a striving for maximum regulatory precision and to a limitation of the role of case law in the sphere of basic institutions. The Polish Civil Code, although amended many times since 1964, retains a more synthetic character. It leaves more work to doctrine and to the courts.

It is noteworthy that, in the case of the civil codes, the relation of detail is reversed in comparison with the constitutions. The Czech constitution is concise, yet the Czech civil code is highly elaborate. The Polish constitution is elaborate, whereas the civil code remains relatively synthetic. This may be

interpreted as two different legislative strategies: the Czech Republic transfers precision to the level of ordinary statutes, Poland to the constitutional level. The codes that follow the constitution therefore do not simply copy its style; they compensate for it.

This compensation is productive for legal linguistics. If we asked only which language is more detailed, the answer would be unstable: Polish at the constitutional level, Czech at the level of the civil code. Comparative statistics force us to distinguish levels of the legal system. They also explain why a translator who has mastered constitutional Czech-Polish correspondences may still be surprised by the density of the Czech Civil Code.

Table 3: Selected terms — frequency and scope (indicative)

Term	Czech Civil Code	Polish Civil Code
závazek / zobow- iżanie	very high	high
smlouva / umowa	extremely high	very high
vlastnictví / własność	high	high
společnost / spółka	also civil-law	mainly commercial

These examples correspond to observations developed in my other research on collocations and false friends. Terms such as *společnost* / *spółka* or *závazek* / *zobowiązanie* are not merely lexical counterparts: their frequency and scope of use reflect a different regulatory density. In the Czech Civil Code, *závazek* appears across a very wide range of contexts, which requires particular caution when selecting a Polish equivalent.⁴ The same is true of *společnost*, which in Czech legal language may denote a company, an association or, outside the strict legal register, society as such. Polish *spółka* is narrower. Quantitative comparison thus leads back to qualitative contrast: numbers make visible the places where the translator must not rely on formal similarity.

5. Commercial Company Legislation

Both acts are modern and adapted to the realities of a market economy and EU law. The Czech statute, however, shows a greater tendency to regulate in detail

⁴ These observations are developed more fully in my forthcoming book on Czech-Polish legal terminology, translation theory and applied linguistics.

Table 4: Structure of company legislation

Parameter	ZOK	KSH
Paragraphs	c. 780	c. 630–650
Detail	high, more casuistic	high, more synthetic

the organs of companies, managerial liability and transformation procedures. Once again we see a continuation of the style already present in the civil code: Czech legislation prefers casuistry. The Polish Commercial Companies Code is by no means brief, but it remains somewhat more synthetic and leaves more space for judicial interpretation and commercial practice.

The similarity of the catalogues of company forms is itself instructive. After 1989 both systems reconstructed private-law vehicles for business activity under the influence of continental models and, later, of EU company law. Quantitative differences therefore cannot be explained by a different inventory of institutions. They arise from the way those institutions are written: how many paragraphs are needed to constitute a limited liability company, how much of the organ structure is fixed by statute, and how much is left to the articles of association.

For translator training, these observations can be developed into tasks that compare terminological collocations in source texts, translations and comparable non-translated legal documents, following the research design of Leńko-Szymańska and Biel (2023). Corpus-based teaching of legal language also provides a relevant point of comparison (Kontutytė 2025). Differences in the length of national statutes do not, on their own, justify shortening or expanding the legal content of a translation; such decisions require attention to the purpose of the translation and the legal concepts involved (Graziadei 2025).

6. Synthesis: What Does the Comparison Yield?

Tables 1–4 present a coherent, though not homogeneous, picture:

1. The Polish constitution is clearly more extensive and more axiological. The Czech Constitution prefers conciseness.

2. The Czech Civil Code is quantitatively dominant: a substantially higher number of paragraphs and greater regulatory density.
3. Company legislation is similar in terms of the catalogue of legal forms, yet the Czech act remains somewhat more elaborate.
4. Quantitative differences have direct translational and didactic consequences.

The comparison is productive for several reasons. First, it shows that linguistic proximity does not mean stylistic proximity. Second, it reveals different legislative strategies: the Czech Republic transfers precision to the level of ordinary statutes, Poland to the constitutional level. Third, it provides empirical support for claims formulated in my other research on collocations, false friends and the Latin heritage of the two legal languages. Finally, it allows students and translators to understand that law in both languages is not merely a set of terms, but also a set of different ways of textually organising legal reality.⁵

In other words, comparative statistics are not an ornament of legal linguistics. They are a means of knowledge. They tell us where to look more closely, which equivalents are only apparent, and how the style of a constitution prepares - or fails to prepare - the style of the codes that implement it.

7. Conclusion

A statistical comparison of Czech and Polish legal acts shows that quantitative differences are not accidental. They reflect distinct legislative traditions and different approaches to regulatory precision. The Czech system tends towards detail at the level of the codes; the Polish system towards greater code-level synthesis combined with an expanded constitutional text. Awareness of these differences is indispensable both in legal translation practice and in the teaching of languages for specific purposes.

The argument of this article will be taken further in a forthcoming larger book devoted to Czech-Polish legal terminology in the light of translation

⁵ On corpus-driven legal linguistics see Glogar (2025); on textual fit in translated law see Lucja Biel, *Lost in the Eurofog* (revised edition).

theory and applied linguistics. There the quantitative findings sketched here can be combined with a closer analysis of collocations, false friends and the Latin layer of both legal languages. Comparative statistics thus prove to be not only a method of description, but also a means of understanding distinct legal cultures.

References

- [1] BIEL, Łucja, and Hendrik J. KOCKAERT, eds. 2023. *Handbook of Terminology*. Volume 3. *Legal Terminology*. Amsterdam: John Benjamins. <https://doi.org/10.1075/hot.3>
- [2] BIEL, Łucja, Katarzyna WASILEWSKA, and Dariusz KOŻBIAŁ. 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* 31(1): 64–93. <https://doi.org/10.1075/ijcl.25126.bie>
- [3] BIEL, Łucja. 2017. *Lost in the Eurofog: The Textual Fit of Translated Law*. Second Revised Edition. Frankfurt am Main: Peter Lang. <https://doi.org/10.3726/b11383>
- [4] FILTZ, Erwin, Sabrina KIRANE, and Axel POLLERES. 2021. “The linked legal data landscape: linking legal data across different countries.” *Artificial Intelligence and Law* 29: 485–539. <https://doi.org/10.1007/s10506-021-09282-8>
- [5] GLOGAR, Ondřej. 2025. “Representative Corpus of the Czech Legal Language: A Tool for Jurisprudence.” *International Journal for the Semiotics of Law* 38: 1681–1703. <https://doi.org/10.1007/s11196-025-10254-3>
- [6] GRAZIADEI, Michele. 2025. “Legal Translation and the Quest for Authenticity.” *International Journal for the Semiotics of Law* 38: 2565–2593. <https://doi.org/10.1007/s11196-025-10283-y>
- [7] KONTUTYTĚ, Eglė. 2025. “Was können Korpora besser als andere digitale Hilfsmittel? Korpusarbeit im fachbezogenen DaF-Unterricht am Beispiel der Rechtssprache.” *Kalbotyra* 78: 13–35. <https://doi.org/10.15388/Kalbotyra.2025.78.1>
- [8] LEŃKO-SZYMAŃSKA, Agnieszka, and Łucja BIEL. 2023. “Terminological collocations in trainee and professional legal translations: A learner-corpus study of L2 company law translations.” *International Journal of Learner Corpus Research* 9(1): 29–60. <https://doi.org/10.1075/ijlcr.22008.len>
- [9] RECSKI, Gábor, Eszter IKLÓDI, Björn LELLMANN, Ádám KOVÁCS, and Allan HANBURY. 2025. “BRISE-plandok: a German legal corpus of building regulations.” *Language Resources and Evaluation* 59: 1043–1082. <https://doi.org/10.1007/s10579-024-09747-7>

- [10] VOGEL, Friedemann, Hanjo HAMANN, and Isabelle GAUER. 2018. "Computer-Assisted Legal Linguistics: Corpus Analysis as a New Tool for Legal Studies." *Law & Social Inquiry* 43(4): 1340–1363. <https://doi.org/10.1111/lsi.12305>
- [11] VOGEL, Friedemann, Tonio WALTER, and Felix TRIPPS, eds. 2022. *Korpuslinguistik im Recht. Theoretische Überlegungen und Fallstudien*. Berlin: Duncker & Humblot. *Sprache und Medialität des Rechts* 5. <https://doi.org/10.3790/978-3-428-58616-5>
- [12] WRIGHT, David. 2026. "Corpus perspectives on legal discourse." *International Journal of Corpus Linguistics* 31(1): 1–12. <https://doi.org/10.1075/ijcl.26065.wri>

About the Author

Ewa Michalska is rooted in Polish Silesia. She studied Bohemistics and completed her doctorate in Slavic languages. Her work concerns false friends and a contrastive view of Polish and Czech, above all in law, legal terminology and drafting technique, and in specialist terminology. From this line of research come her published studies and her forthcoming qualifying thesis.

Ewa Michalska (Bielsko-Biała)

ORCID 0000-0002-6373-6117