



How do translators handle (near-) synonymous legal terms? A mixed-genre parallel corpus study into the variation of EU English-Polish competition law terminology¹

Łucja Biel² y Dariusz Koźbial³

Recibido: 27 de febrero 2020 / Aceptado: 8 de junio 2020

Abstract. Terminological variation, i.e. synonymy at the term level, is regarded as a recurrent problem in EU law. Working with a parallel-comparable corpus of EU English-Polish legislation, soft law and judgments in the area of competition, this study explores how source-language synonymy is handled in translation across institutional genres. The findings show that synonyms may be reflected symmetrically or asymmetrically, with variation being eliminated, partly reduced, mirrored, or increased in translation. It is quite frequent for translators to introduce additional variation and cross-variation. This is affected by: genre, source and target asymmetries, complexity of a semantic field, low termness and microdiachronic shifts. The study confirms that synonymy is one of the causes of variation in translation and calls for more conceptual clarity at the drafting stage.

Key Words: legal terminology, terminological variation, synonymy, EU translation, parallel corpora.

[es] ¿Cómo afrontan los traductores los términos legales (cuasi-)sinonímicos? Un estudio de corpus paralelo inglés-polaco de géneros diversos en la variación de la terminología legal del derecho de la competencia de la UE

Resumen. La variación terminológica, esto es, la sinonimia a nivel del término, es considerada un problema recurrente en la ley de la UE. Utilizando un corpus paralelo comparable inglés-polaco de la legislación, derecho indicativo (*soft law*) y sentencias en la área del derecho de la competencia de la UE, el presente estudio explora las formas en las que la sinonimia en la lengua origen es manejada en la traducción en diferentes géneros institucionales. Los resultados indican que los sinónimos se reflejan en la traducción de una manera simétrica o asimétrica, con la variación eliminada, parcialmente reducida, igual o aumentada. Es relativamente frecuente que los traductores introduzcan variaciones adicionales y variaciones cruzadas. A esto influyen los siguientes factores: género, asimetrías entre la lengua origen y la lengua meta, la complejidad del campo semántico, el nivel bajo de la cualidad terminológica y las oscilaciones microdiacrónicas. El estudio confirma que la sinonimia constituye una de las causas de la variación en la traducción e insiste en la importancia de la claridad conceptual en el estadio de la preparación del borrador.

Palabras clave: terminología jurídica, variación terminológica, sinonimia, traducción de la UE, corpus paralelos.

¹ This work was supported by the National Science Centre (NCN) under Grant 2014/14/E/HS2/00782.

² University of Warsaw
l.biel@uw.edu.pl

³ University of Warsaw
d.kozbial@uw.edu.pl

Sumario. 1. Introduction. 2. Key principles of EU legal terminology: autonomy, continuity and multilingual concordance. 3. Corpus studies into EU terminology: a focus on variation. 4. Material: an English-Polish parallel-comparable corpus of EU competition legislation, soft law instruments and judgments. 5. Method. 6. Parallel-corpus analysis of Polish equivalents: Findings and discussion. 6.1. How is source language variation reflected in translation? 6.2. What may affect the translators' choices when they deal with source language variation? 7. Conclusions.

Cómo citar: Biel, Ł. y Koźbiał, D. (2020) How do translators handle (near-) synonymous legal terms? A mixed-genre parallel corpus study into the variation of EU English-Polish competition law terminology, en *Estudios de Traducción* 10, 69-90.

1. Introduction

Given the essential function of terms in the representation and transfer of specialised knowledge (cf. Cabré Castellví 2003: 182), terms make a significant contribution to the meaning-making in texts. This claim is no less valid for legal terms which are points of access to legal knowledge (Biel 2014: 41). Owing to the systemic nature of law, legal terms should — ideally — be used consistently, clearly referring to and activating relevant concepts to control phenomena typical of natural languages, such as synonymy (variation) and polysemy. Yet, as practice and research show, this is not always the case. These phenomena are further exacerbated in multilingual supranational law, such as EU (European Union) law, which exists — thanks to translation — in 24 authentic language versions and is subject to the extreme filtering of content through other languages and the unavoidable side effects of multilingualism (cf. Biel & Sosoni 2019).

Drawing on an earlier study into the terminology of EU competition law (Biel et al. (2018a) (2018b)), which revealed terminological variation in EU English, the EU's main drafting language, we shall explore how term variation in the form of near and full synonyms is handled in legal translation into Polish, one of the EU's official languages. For this purpose, we shall identify semantically related terms of EU competition law and, using an English-Polish parallel corpus, we shall extract and study the equivalents of synonyms to verify whether translators mirror or rather reduce this type of variation in target texts. Additionally, we shall investigate what factors affect the translators' rendering of synonyms. In particular, we shall check whether it differs across genres by working with three institutional legal genres: primary/secondary legislation, soft law instruments and judgments. The study is exploratory in nature and intends to improve our understanding of causes of terminological variation. Its methodological goal is to test the applicability of parallel corpora to the study of legal terminology.

2. Key principles of EU legal terminology: autonomy, continuity and multilingual concordance

EU translation has been critically discussed and synthesised in a variety of publications, which often point to the theoretical and methodological challenges in researching it

(cf. Koskinen 2001, Kjær 2007, Biel 2014, Šarčević 2017). The roots of challenges are often sought, on the one hand, in the complex interplay of extra-textual variables and, on the other hand, in intertextual relations in which EU texts are embedded (cf. Biel 2019). The institutional, political, supranational, legal and multilingual variables foreground the key properties of EU texts: they are institutional translations which entail high legal and political risks (Biel 2019: 27-32). In the EU context, a translation has multiple intertextual dimensions: multilingual concordance, that is a horizontal relation of equivalence with other language versions, continuity (consistency) — a systemic vertical relation with preceding EU texts in the same language, and the (textual) fit — the relation with corresponding national texts in the Member State where it is applied (cf. Robertson 2015: 40-44, Biel 2019: 33-36).

The dimensions of multilingual concordance, continuity (consistency) and the (textual) fit are particularly pertinent to EU legal terminology. Terminology should maintain **continuity** (consistency) (1) internally within a given text and (2) externally with the EU *acquis*, in particular with preceding legal texts in the same domain and higher-ranking legislation (cf. Stefaniak 2017: 114-16, Prieto Ramos 2014: 314, Robertson 2015: 42). Secondly, EU law is autonomous supranational law enacted in 24 authentic language versions; hence, EU legal terms, in principle, refer to **autonomous** supranational concepts (cf. EU case law “Community law uses terminology which is peculiar to it”⁴; Doczekalska 2018: 174-75). According to the CJEU (Court of Justice of the European Union) case-law⁵, such concepts are expected to be uniformly interpreted and applied in all Member States. Thus, as part of multilingual **concordance**, terms ideally trigger the same concept in all language versions of a legal act. However, not all terms used in EU legal acts are fully autonomous supranational terms: EU law also uses generic legal concepts, such as *lawyer* or *penalty* (cf. Robinson 2014: 202). Thirdly, although EU law is drafted at the supranational level, it is derived from the Member States’ national legal systems (Bajčić 2017: 83), and is “an integral part of the Member States’ legal systems” (Doczekalska 2018: 175) where it is applied (Kjær 2007). Thus, one of the properties of EU terms is their “travelling” through supranational and national contexts (Biel & Doczekalska 2020). Finally, EU genres may vary in the degree of saturation with autonomous concepts and national terms; for example, Prieto Ramos observes that genres related to adjudication and monitoring have more “references to national legal realities” than legislation (2014: 319).

In respect of translation, there is no consensus as to what extent EU terminology should **fit** national usage; in other words, to what extent EU terms should be approximated to national terminology. On the one hand, the dominant narrative stresses that since we are dealing with autonomous concepts, the distinctness of EU terms should be clearly signalled to recipients through the source-oriented literal approach and equivalents which do not trigger national legal concepts (Baaij 2018). On the other hand, EU law interacts and should be coherent with national legal systems. A major challenge for EU translators is to identify when a term should be either foreignised or domesticated (see also Doczekalska 2018: 175).

⁴ Case C-283/81 of 6 October 1982 Srl CILFIT and Lanificio di Gavardo SpA v Ministry of Health, ECR 3415, ECLI:EU:C:1982:335.

⁵ E.g. Case C-467/08 of 21 October 2010 Padawan SL v Sociedad General de Autores y Editores de España (SGAE), ECR 2010 I-10055, ECLI:EU:C:2010:620.

Another aspect of EU terminology is its presumed high **standardisation**, due to the institutional nature of EU translation. EU institutions develop resources to support translators' terminology mining and standardise their terminological choices, e.g. IATE, EuroVoc or CuriaTerm multilingual termbases.

Despite these efforts, terminological variation (synonymy) is regarded as a "recurrent additional problem" in EU texts (Strandvik 2012: 39). Synonymy is a degree of semantic resemblance between two terms: there may be full (absolute) synonyms which share the conceptual content and near (partial) synonyms which have an overlapping content or different distribution (cf. Chromá 2011: 40-42; Goźdz-Roszkowski 2013: 96-97). As demonstrated by Goźdz-Roszkowski (2013: 108), the relation between related legal terms tends to be complex with frequent overlaps, exclusions and domain-specificity. It should be stressed that variation is considered to be an undesirable phenomenon in legal language, especially in multilingual law: "in the multilingual context of EU legislation, synonymy is always a problem and works to the detriment of clarity" (Strandvik 2012: 39). The *Joint Practical Guide* recommends EU drafters to avoid synonyms (EU 2015: 11). They may be confusing (Strandvik 2012: 40) and cause legal uncertainty (Bajčić & Martinović 2018: 212). As suggested by Strandvik, when four terms refer to the same concept, especially when they are not defined, "it is most likely that four different instances of 'grey ice' will be produced in all languages" (2012: 39-40).

3. Corpus studies into EU terminology: a focus on variation

The usefulness of electronic corpora in Translation Studies (TS) is unquestionable, as they have become one of the main methodologies in quantitative research, including legal translation. It is especially institutional translation which has observed an impressive emergence of research projects exploring legal translation equally at the micro level of individual lexico-grammatical patterns, the macro level of institutional varieties of legal languages (e.g. Eurolects) and the meta level of contrastive studies across a range of institutions (Biel 2020). Most studies use comparable corpora to juxtapose two or more corpora, e.g. across genres, diachronically, or translations to nontranslations. While comparable corpus research is diverse and methodologically sophisticated, parallel corpus studies have not made much progress yet. A parallel corpus is a translation corpus which is at least bilingual and contains source texts aligned with corresponding target texts. Thus, it may provide insight into the relation between source texts and target texts. However, because of technical limitations, parallel corpus research often entails time-consuming small-scale manual analyses, without fully enjoying the benefits of the quantitative approach.

Another niche is the study of terminology through quantitative approaches. There are few examples of corpus studies into EU terminology. We shall discuss those which, by reporting variation, are most relevant to our study: a comparable corpus study by Biel (2014), parallel corpus studies by Vigier and Sánchez Ramos (2017), Prieto Ramos and Guzmán (2018) and Prieto Ramos and Morales Moreno (2019), and corpus studies specifically into EU competition law by Biel et al. (2018a), Sosoni (2018) and Dobrić Basaneže (2018).

Working with a corpus of 1958-2006 *acquis* translated into Polish, Biel (2014: 271-76) observes an increased variation and instability of terminology when compared

to nontranslated legislative Polish. The increased variation concerns orthography, morphology and inflection, especially in the case of borrowings, as well as a more problematic denominative variation, e.g. variation of constituents in multi-word terms or legal variants used interchangeably with semi-legal and colloquial variants (Biel 2014: 271-76; see also Biel 2018: 316). It is attributed to interference, low termness of neologisms, inadequate knowledge of TL terminology, insufficient terminological resources and the instability of the Polish Eurolect at its formative stage (Biel 2014: 276).

Vigier & Sánchez Ramos (2017) use an English-Spanish parallel corpus of CJEU judgments to study how national legal terms — names of English and Spanish criminal courts — are handled in institutional translation. Their findings show a strong preference for the foreignising approach through the direct loan technique (95% of cases) (2017: 13). More importantly, Vigier & Sánchez Ramos indicate a lack of consistency (variation) of equivalents, in particular in how such borrowings are approximated through descriptors or glosses and treated grammatically, ascribing the latter to the lack of consistency in source texts (2017: 11, 14).

Using a parallel corpus, the study by Prieto Ramos and Guzmán (2018) investigates how the term *due process* was translated into Spanish by three institutions — the EU, the United Nations (UN) and the World Trade Organization (WTO) — over the span of 15 years. Their findings reveal a considerable variation of Spanish equivalents, with a clearly higher intertextual variation and a marked increase of intratextual variation over time in EU documents (Prieto Ramos & Guzmán 2018: 89). The study is supplemented with a qualitative analysis of the adequacy of equivalents and shows the lowest, but growing, adequacy of EU translations (Prieto Ramos & Guzmán 2018: 97). The finding of increased variation in EU Spanish translations was confirmed by Prieto Ramos and Morales Moreno (2019), who conducted a microdiachronic lexicometric analysis of four English terms, *governance*, *tariff peak*, *tariff escalation*, *hedge fund*, and their Spanish equivalents in three institutions (the EU, the UN, and the WTO). The increased variation and lower adequacy of EU Spanish texts is explained with the larger fragmentation of texts and translation services, and a lower thematic convergence (Prieto Ramos & Guzmán 2018). Both studies stress that translation introduces variation and call for an increased standardisation of terminology.

The next group of studies (Biel et al. 2018a, 2018b; Sosoni 2018; Dobrić Basaneže 2018) was an outcome of an EU-funded project (see footnote 17) and focuses on terminology and phraseology in EU competition law. Working with a parallel corpus of English, Greek, Italian and Spanish legal acts, Sosoni (2018) points to discrepancies in terminological equivalents between languages, in particular the increased explicitness of Greek equivalents compared to Italian and Spanish. Having compiled a glossary of collocations of company law terminology on an English-language corpus of EU competition legislation, soft law and judgments, Biel et al. (2018a) report increased variation (inconsistency) as one of the distinctive features of EU English competition terminology and collocations. It includes in particular non-denominative variation with a range of graphical, orthographic, morphological, inflectional and permutation variants, e.g. *anti-competitive*, *anticompetitive*, *anti competitive*; *investigatory*, *investigative*; *action for damages*, *damages action*, as well as denominative variation (synonyms) identified at the level of terms and collocations, e.g. *restriction/limitation of competition*, *to form/establish/conclude a cartel* (Biel et al. 2018a: 268-69). Variation can also affect other types of phraseological units, as was demonstrated by Dobrić Basaneže (2018) in her study of

binomials; that is, sequences joined by *and* or *or*, such as *research and development*, *goods and services*, *jointly and severally*, *to rescue and restructure*. Working with the parallel English-Croatian corpus of legal acts and soft law in the field of competition, Dobrić Basanež (2018: 235, 241) observes a certain degree of reversibility and modifications of binomials in translations into Croatian.

The above reported studies indicate the increased terminological variation in EU translations, evidencing that translation introduces both non-denominative and denominative variants. It has also been pointed out that variation exists not only in translations, but also in English language versions of legal acts presumed to be source texts. Drawing on the latter, this study will approach variation through the lens of semantically related terms identified in English source texts and their treatment in Polish translations.

4. Material: An English-Polish parallel-comparable corpus of EU competition legislation, soft law instruments and judgments

The study is based on a thematically-controlled parallel corpus in order to better control the domain-specific meaning of terms subject to analysis. The corpus focuses on competition law, the role of which is to protect competition and consumers (Woods et al. 2017: 628-29). Our interest in EU competition law is motivated by several factors. Firstly, competition law is one of the most mature and complex areas of EU law, which combines law and economics (Woods et al. 2017: 680). It shows EU law at its fullest: it has greatly contributed to European integration by opening up national markets and strengthening the EU Common Market (Patel & Schweitzer 2013: 1). This branch of law experiences a larger universality of concepts partly to its internationalisation and the transnational nature of competition itself (Dabbah 2010; cf. Biel & Sosoni 2019 for discussion).

For the purposes of this study, we compiled a mixed-genre parallel-comparable corpus of English and corresponding Polish texts, totalling about 1 million English and 1 million Polish words (see Table 1). The corpus covers three genres: primary and secondary legislation, non-binding soft law instruments and judgments, all of which shape EU competition rules to a larger or smaller extent.

Table 1. Design of the parallel English-Polish EU Competition corpus

Sub-corpora	No. of texts	No. of words English corpus	No. of words Polish corpus
EULaw2011-15	12	158,012	159,216
EUSoftLaw2011-15	4	53,905	47,782
EUD judgments2011-15	42	700,246	700,095
EULaw2002* ⁶ -10	17	74,061	64,421
<i>Total:</i>	<i>75</i>	<i>986,224</i>	<i>971,514</i>

⁶ The time range refers to the secondary legislation only. This sub-corpus also includes relevant passages from the Treaty on the Functioning of the European Union (TFEU), consolidated version OJ C 115, 9.5.2008, which originated as the Treaty of Rome coming into force on 1 January 1958 (For more information, see <https://ec.europa.eu/competition/information/treaty.html/>).

The main corpus EULaw2011-15 comprises EU legal acts from the period of 2011-2015. Since this study is part of a larger study into the Polish Eurolect⁷, the time frame corresponds to that of the PL Eurolect project. The main corpus, EUCompLaw2011-15, comprises 12 legal acts: 8 Commission regulations, 1 Commission decision, 2 Council regulations, and 1 directive of the European Parliament and the Council⁸. It covers all secondary legislation of general application in the EU included in the subject *08 Competition policy* of EUR-Lex⁹ for this period, except for implementing and amending regulations and numerous decisions, mainly on state aid, addressed to specific countries only. The texts were included in the corpus in full, that is together with non-normative parts and annexes.

The remaining sub-corpora from the period of 2011-2015 cover soft law (EUSoftLaw2011-15) and judgments (EUJudgments2011-15). The EUSoftLaw2011-15 sub-corpus consists of Commission communications and a recommendation (4 files). The EUJudgments2011-15 corpus is composed of all 42 CJEU judgments (10 Court of Justice judgments and 32 General Court judgments) in the subject matter of competition delivered in the period of 2011-2015 and limited to those granting an application and with English as the language of the case.

We also included a fourth reference corpus of EU secondary legislation of general application (regulations and directives), except for implementing and amending acts, published between 2002 and 2010, directly preceding the period we are interested in, as well as the relevant excerpt of the fundamental primary legislation (Articles 101-109 of the TFEU) which sets out the EU competition policy and establishes key authoritative terminology in the field which should be continued in lower-ranking secondary legislation.

The texts were downloaded from the EU's official databases¹⁰ and saved as plain text files with the UTF-8 encoding. The files were automatically aligned with LF-Aligner 4.1¹¹ and manually verified for noise. The aligned files of each sub-corpus were merged in TMXEditor¹² and uploaded to Sketch Engine¹³ (Kilgarriff et al. 2014), which was the main tool used in the analysis. We also used Wordsmith Tools 7.0 (Scott 2016) to verify the distribution of terms across texts.

There are crucial differences in the status and translation arrangements among the three institutional genres included in the corpus. Legislation is the most authoritative prescriptive genre which is covered by full multilingualism: it is expressed in 24 authentic language versions (cf. Šarčević 1997: 64). Its drafting is multi-stage and multilingual: translation is done mainly in-house by the institutions and undergoes strict quality assurance with bilingual review and legal revision (Biel et al. 2019: 69-70). Soft law covers non-binding instruments, such as recommendations, opinions, communications, notices, which are used by the European Commission as “new ways of developing policy” through guidelines, in particular in competition law (Woods

⁷ For more information on the PL Eurolect research project, see <https://eurolect.ils.uw.edu.pl/>.

⁸ For a list of files, see <https://eurolect.ils.uw.edu.pl/publications-and-materials/additional-materials/>.

⁹ <https://eur-lex.europa.eu/browse/directories/legislation.html?locale=en>.

¹⁰ EU legal acts were downloaded from EUR-Lex (<https://eur-lex.europa.eu/advanced-search-form.html?qid=1574588083056&action=update>), while CJEU judgments were downloaded from InfoCuria (<http://curia.europa.eu/juris/recherche.jsf?cid=3821537>).

¹¹ Available at <https://sourceforge.net/projects/aligner/>.

¹² Available at <https://maxprograms.com/products/tmxeditor.html>.

¹³ Available at <https://www.sketchengine.eu>.

et al. 2017: 71, 629). They may be translated in-house or outsourced to external contractors, thus running the risk of lower quality controls. The role of the third genre, judgments, is to resolve disputes. More importantly, judgments communicate an authoritative interpretation of the law and ensure that it is applied in a uniform manner throughout the EU Member States (cf. Prieto Ramos 2014: 315-16). CJEU judgments are customarily deliberated in French (McAuliffe 2016: 11), the Court's working language, and are translated into all official languages; however, only a judgment issued in the language of the case is authentic¹⁴ (cf. Derlén 2015: 68). CJEU judgments are translated by in-house and external lawyer-linguists with a law degree from a Member State of origin and a thorough knowledge of at least two official EU languages¹⁵ (cf. Koźbiał 2017). Given the current status of English as the EU's lingua franca (Biel et al. 2018a: 250-54), English language versions may, with a high likelihood, be regarded as source texts in the case of legal acts and soft law. As regards judgments, English versions are target texts translated from French although, at the same time, they are *de jure* originals since English is the language of the case in our corpus. Polish language versions are always target texts (translations). Furthermore, the genres in the corpus become involved at different stages of the life of a term: legislation and soft law create concepts, while judgments interpret and adjust their meaning (cf. Biel & Doczekalska 2020). We hypothesise that genre differences may affect the translators' decision-making process and solutions adopted in target texts.

5. Method

To identify semantically related terms we generated and analysed the keyword¹⁶ list of our main focus corpus EULaw2011-15 (1400 single-word items). The keywords were extracted with Sketch Engine's EUR-Lex2016 set as a reference corpus. We opted for the large EUR-Lex corpus of EU texts rather than a general English corpus to identify competition law terms¹⁷ and eliminate general EU terminology, e.g. names of institutions or instruments, which are common in EU documents. The list of potential synonyms was supplemented through the analysis of the EULaw2011-15 wordlist and the glossary of collocations of EU English competition terms compiled as part of an earlier corpus study¹⁸ (Biel et al. 2018a, 2018b). The potential synonyms were initially clustered into groups of semantically similar terms based on our knowledge of their collocational behaviour gained during the glossary compilation process. For example,

¹⁴ See Article 41 of the Rules of Procedure of the Court of Justice and Article 49 of the Rules of Procedure of the General Court.

¹⁵ See Article 42 of the Rules of Procedure of the Court of Justice.

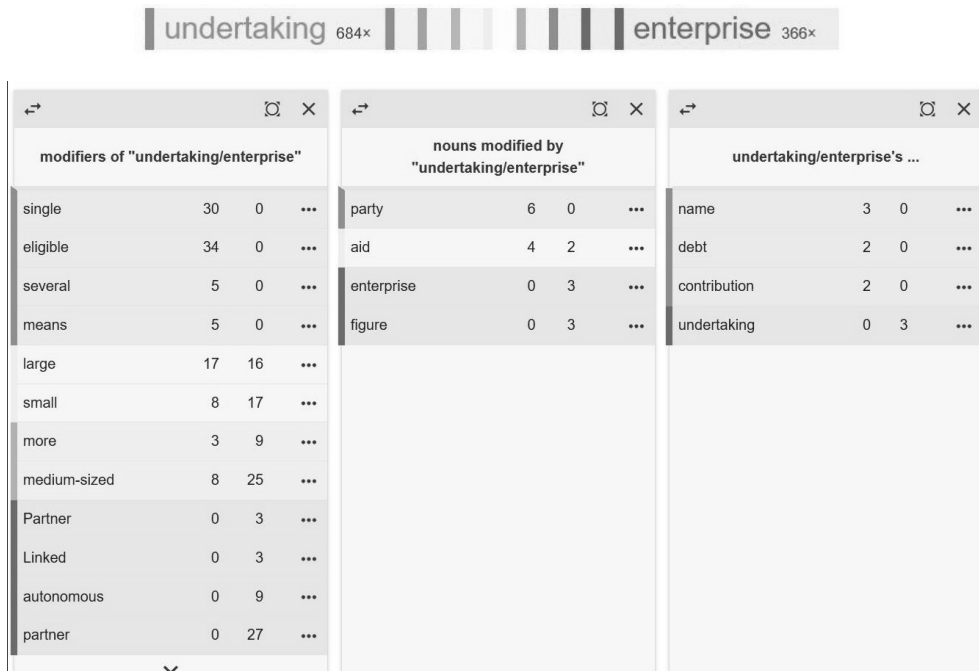
¹⁶ For more information on how Sketch Engine identifies keywords, (2009).

¹⁷ We excluded terms from other areas, e.g. *pest*, *agroforestry*, *avalanche*, *eruption*, *biofuel*, which were identified as keywords.

¹⁸ The glossary was compiled within the action grant *Training action for legal practitioners: Linguistic skills and translation in EU Competition Law* funded by the DG Competition of the European Commission and coordinated by the University of Insubria. The project consisted in the legal and English-language training of national judges in the area of EU competition law. In order to compile a glossary, we first identified over 100 term candidates in a corpus of legal acts, soft law, case law and academic textbooks, working with the keyword and term extraction functions and next analysed their collocational environment through concordances, collocations and word sketch functions (cf. Biel et al. 2018a).

terms *undertaking*, *enterprise*, *company*, *entity*, or *firm* refer to a type of business entity and have an overlapping conceptual content. The potential synonyms were further verified in EULaw2011-15 concordances through the collocational analysis proposed by Goźdz-Roszkowski (2013) for near synonyms as well as the Sketch Engine word sketch difference function¹⁹ to analyse differences in use, overlaps of contexts and the degree of semantic similarity (see Figure 1).

Figure 1. A screenshot of Sketch Engine's word sketch showing differences in the modifiers of *undertaking* and *enterprise* in EULaw2011-15



This helped us eliminate items which had a divergent distribution²⁰, e.g. *reduction / discount* (*tax reduction, discount rate*) or *advantage / benefit / privilege* (*tax advantage, benefit of the block exemption, legal professional privilege*). The list below shows clusters of semantically similar terms in the EULaw2011-15 corpus (bolded terms are those terms which were earlier identified by Sketch Engine as keywords in the focus corpus; figures in brackets indicate a raw frequency followed by a normalised frequency per million words in the focus corpus). These are 19 clusters of terms (nouns and adjectives), of which 15 have at least 2 keywords. They are arranged thematically below:

¹⁹ This function compares word sketches of two lexemes to show usage differences <https://www.sketchengine.eu/quick-start-guide/word-sketch-difference-lesson/>.

²⁰ According to the distributional hypothesis proposed by Harris, semantically similar words appear in similar linguistic contexts (cf. 1954).

1. **aid** (3249/18153), *support* (41/229), *assistance* (14/78)
2. **grant** (214/196), **subsidy** (19/106)
3. **unlawful** (26/145), **illegal** (7/39)
4. **state** (712/3978), **public** (223/1246)
5. **undertaking** (710/3967), **enterprise** (366/2045), *company* (49/274), *operator* (16/89), *entity* (39/218), *firm* (9/50), *entrepreneur* (1/6), *economic actor* (1/6)
6. **shareholder** (31/173), **member** (778/4347)
7. **cartel** (21/117), **concerted** (8/45), **agreement** (118/659), *collusion* (2/11), *coordination* (9/50), *arrangements* (8/45)
8. **acquisition** (69/385), *take-over* (2/11)
9. **infringement** (116/648), **violation** (5/28), *contravention* (4/22), *breach* (1/6)
10. **damage** (113/681), **harm** (77/430), **loss** (77/430)
11. **compensation** (150/838), **damages** (109/609), **remedy** (10/56), *relief* (3/17), *redress* (2/11)
12. **fine** (26/145), **penalty** (36/201)
13. **limitation** (48/268), **restriction** (37/207)
14. **ceiling** (139/777), **threshold** (79/441), **cap** (50/279)
15. **purchaser** (29/162), **buyer** (10/56)
16. **beneficiary** (231/1291), **recipient** (11/61)
17. **evaluation** (58/324), *assessment* (54/302)
18. **cost** (852/4760), **overheads** (11/61), *expenditure* (45/251), *expense* (10/56)
19. **charges** (20/112), *fees* (31/173), *levies* (1/6), *royalties* (1/6)

The next step of the analysis was to extract Polish equivalents from the parallel corpus. In most cases, Sketch Engine was not able to recognise equivalents per se by highlighting them in Polish texts and they had to be retrieved through parallel searches and manually from concordances. What proved particularly useful in the case of high-frequency terms was the “does contain” and “does not contain” option in Sketch Engine parallel searches. For example, after we analysed the first 30 concordances and identified the most common equivalents of a given term, we ran a “does not contain” search to obtain concordances with other variants. It was time-consuming, but still, this tool made the extraction of equivalents more efficient compared to a purely manual identification of variants.

6. Parallel-corpus analysis of Polish equivalents: Findings and discussion

6.1. *How is source language variation reflected in translation?*

The analysis of Polish equivalents of synonymous terms shows that synonyms may be reflected symmetrically or asymmetrically in translation and variation may be eliminated, partly reduced, maintained or increased (see Figure 2).

Let us start with the **elimination** of variation in target texts when a set of synonyms is rendered with a single equivalent. It concerns a few clusters:

- **buyer, purchaser** → *nabywca* [purchaser],
- **infringement, violation, contravention, breach** → *naruszenie* [breach],

- *financial ceiling, financial threshold* → *pułap finansowy* [financial ceiling],
- *current costs, current expenditure* → *wydatki bieżące* [current expenses].

For example, *buyer* and *purchaser* are rendered with a single equivalent *nabywca*, although legal Polish also has two synonyms — *nabywca* and a slightly less formal synonym *kupujący*. The term *buyer* is used in EULaw2002-10, where it is consistently translated as *nabywca*, except for a single occurrence of *kupujący*, which is not continued in later translations. A new denomination *purchaser* is introduced in Directive 2014/10/EU in defined terms: *indirect purchaser* (*nabywca pośredni*) and *direct purchaser* (*nabywca bezpośredni*). The term *nabywca* is used consistently in Polish judgments, while English judgments use *buyer* (3) and *purchaser* (13) interchangeably.

Cluster #9 terms — *infringement, violation, (in) contravention (of) and breach* — are distributed in slightly different contexts, with *infringement of competition law* being a defined term, *violation* collocating mainly with *Union law* and *breach* collocating with an *article* or *obligation*. They were all rendered as *naruszenie*, a corresponding Polish term of a broader scope.

In some cases, translators successfully identify and eliminate synonyms but at the same time introduce additional variation, leading to an only **partial reduction of variation**. There are usually one dominant equivalent and some less frequent, even one-off, variant equivalents of one of the terms.

- *acquisition* → *przejęcie* [take-over]; *the acquired establishment* → *nabyty zakład* [purchased establishment],
take-over → *przejęcie* [take-over],
- *beneficiary undertaking* → *przedsiębiorstwo będące beneficjentem* [an undertaking being a beneficiary], *przedsiębiorstwo korzystające z pomocy* [undertaking using an aid],
recipient undertaking → *przedsiębiorstwo będące beneficjentem* [an undertaking being a beneficiary].

For example, both *acquisition* and *take-over* are translated as *przejęcie*. This disambiguation is disturbed when a variant is introduced for a derivative form (conversion) — *the acquired establishment* → *nabyty zakład* [purchased establishment] instead of a consistent *przejęty zakład*. Another illustration is cluster #16 — *beneficiary* and *recipient*. These terms tend to be distributed in different contexts: *aid beneficiary* and *immunity recipient* or *recipient of decision*, except for *beneficiary undertaking* and *recipient undertaking*. Both were correctly disambiguated as *przedsiębiorstwo będące beneficjentem* [undertaking being a beneficiary], although one odd descriptive equivalent of the former, *przedsiębiorstwo korzystające z pomocy* [undertaking using an aid], was also spotted. The second scenario takes place when source language synonyms are **symmetrically mirrored** in translation and variation is maintained. It can be observed in cluster #4, *state* and *public*, when these terms function as premodifiers of *aid*:

- *state aid* → *pomoc państwa* [state aid],
public aid → *pomoc publiczna* [public aid].

State aid is one of the most prominent high-frequency terms of competition law, with a consistent literal equivalent of *pomoc państwa* in the legal acts. Commission

Regulation (EU) 1388/2014, which applies to fisheries and aquaculture, uses a related domain-restricted term *public aid* with reference to Article 95 of Regulation (EU) No 508/2014, where it is translated literally as *pomoc publiczna* [public aid] and hence the same equivalent is continued, in line with the continuity principle, in the regulation in question (*maximum intensity of public aid*). Interestingly, *pomoc państwa* is used in national law in the context of welfare support, while *pomoc publiczna* is a national term for *state aid* (cf. Biel & Sosoni 2019), so different relations between these two terms are created at the supranational and national level. Yet the strategy of mirroring is not predictable. When *state* and *public* premodify a synonym of *aid* — *support* (cluster #1) in low frequency extended phrases, *a total amount of public support* and *a maximum level of state support*, the distinction is not maintained in translation in EULaw2011-15. This set illustrates a third scenario when **variation is increased** and synonyms are multiplied in translation. In this case a considerable variation is introduced, in particular for *public support*, where the Polish modifiers evoke both *public* and *state*, while the head noun evokes both *support*, *aid* and *funds*. *State support* has non-denominative variants in a single regulation.

- **state support** → *wsparcie zapewniane przez państwo* [support ensured by the state], *wsparcie ze strony państwa* [support from the state],
public support → *publiczne wsparcie finansowe* [public financial support],
środki pomocy państwa [funds of state aid], *pomoc ze środków publicznych* [aid from public funds].

While the extended phrases per se are not grammatically complex to be rendered into Polish, what may trigger variation is the low frequency of these terms and/or the fact that they are not defined in the legal acts. Both factors contribute to their reduced termness: these terms are less salient and more difficult to be recognised by translators as terms and rendered consistently. This results in the lack of stability of target equivalents and increases variation in translation.

Increased variation may have a form of, what we refer to as, **cross-variation** when source-language synonyms interchangeably trigger a range of synonymous equivalents. Take for example:

- **evaluation plan** → *plan ewaluacji* [evaluation plan], *plan oceny* [assessment plan],
assessment → *ewaluacja* [evaluation], *ocena* [assessment].

The term *evaluation plan* is identically defined and consistently used in two English regulations from June 2014. One of them, Regulation 651/2014, opts for a calque *plan ewaluacji* (a defined term), which is used interchangeably with *plan oceny* on three occasions (intratextual variation). The other, Regulation 702/2014, favours *plan oceny*. Both regulations contain the phrase “assessment of the evaluation plan” which cross-activates synonyms to differentiate between *assessment* and *evaluation*: *ewaluacja planu oceny* [evaluation of the assessment plan] and *ocena planu ewaluacji* [assessment of the evaluation plan].

Cluster #10, which comprises three terms *harm*, *damage* and *loss*, corresponds to two overlapping Polish terms, often used interchangeably — *szkoda* and *strata*. *Loss* is mainly translated as *strata*, *harm* as *szkoda*, while *damage* (also *damages suffered/caused*) as both *szkoda* or *strata*. Some terms are rendered inconsistently with variants, which may be triggered by their interchangeable use in Polish law:

- **material damage** → *szkoda materialna* [material damage], *strata materialna* [material loss],
actual loss → *straty rzeczywiste* [real losses], *szkody rzeczywiste* [real damage].

A more complex type of cross-variation may be observed in cluster #7: *collusion, coordination, agreement, concertation, arrangement, and cartel* (see Table 2).

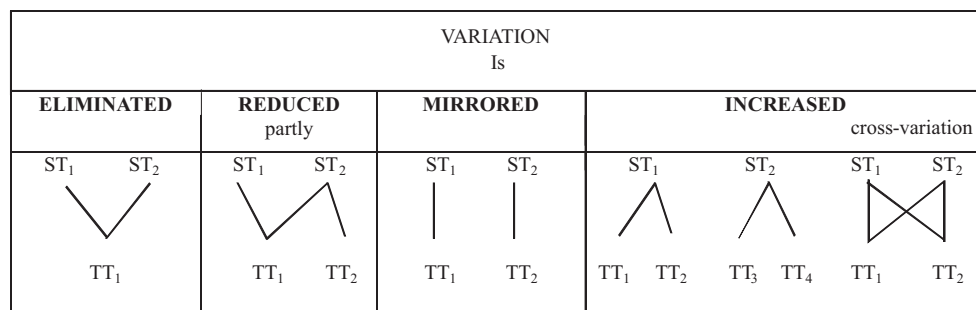
Table 2. Cross-variation of equivalents in the *collusion* cluster

collusion	coordination	concertation	arrangements	cartel
<i>zmowa</i>	<i>uzgadnianie</i>	<i>uzgodnienia</i>	<i>uzgodnienia</i>	<i>kartel</i>
<i>porozumienie</i>	<i>koordynacja</i>		<i>porozumienie</i>	<i>zmowa</i>
	<i>współpraca</i>		<i>ustalenia</i>	<i>naruszenie</i>

All those terms have an overlapping sense of anti-competitive cooperation between undertakings. Equivalents include *porozumienie* [agreement], *uzgodnienia* [arrangements], *zmowa* [collusion], *koordynacja* [coordination], etc. They are often used interchangeably as equivalents. A less rigorous approach to these terms may be observed in judgments: *collusion on prices, price collusion, coordination of prices, coordination on prices, concertation on price trends, price fixing arrangements, price arrangements, and coordination of pricing policies*.

By way of summary, Figure 2 exemplifies basic configurations of equivalents of (near-) synonymous terms in translation.

Figure 2. Basic configurations of equivalents of (near-) synonymous terms (ST = source term; TT = target term).



6.2. *What may affect the translators' choices when they deal with source language variation?*

The solutions adopted by translators are, to some extent, unpredictable and idiosyncratic — for example, a translator may not be aware of the existence of synonymy. Yet, as our data show, they are also affected by factors, such as asymmetries between source and target clusters of synonyms, complexity of a semantic field, the

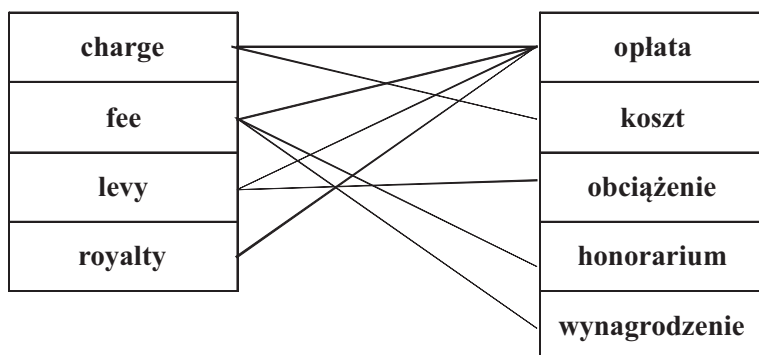
above discussed low termness, genre and microdiachronic shifts, which may result in increased variation.

The first factor is the degree of asymmetry between source and target clusters of synonyms in terms of their number, overlap of senses and level of specificity (more generic or more specific concepts within a cluster). This may be illustrated with the four English terms *charges*, *fees*, *levies*, *royalties* (cluster #13). They have an overlapping sense of payment in exchange for goods or services which often corresponds to a more generic Polish term *opłata* [~charges], as in *tax or other charges / tax and charges* → *podatki i opłaty* and *taxes, duties or levies* → *podatki, cla, opłaty*, where both *charges* and *levies* are translated as *opłaty*. These terms are quite easily disambiguated in their collocational environment due to different distribution, e.g.: *freight charges* → *opłaty frachtowe* [freight charges] or *royalties/royalty fees* → *opłaty licencyjne* [licencing charges]. Some senses trigger other denominations, for instance:

- *depreciation charges* → *koszty amortyzacji* [depreciation costs],
- *engineer fees* → *honoraria inżynierów* [engineers' fees],
- *contingency fees* → *honoraria uzależnione od wyniku sprawy* [fees dependent on the result of the case], a variant *wynagrodzenie uzależnione od wyniku sprawy* [remuneration dependent on the result of the case].

Figure 3 shows the asymmetrical overlapping senses of these terms.

Figure 3. Asymmetries in the organisation of *charge/opłata* semantic fields in English and Polish



The next example concerns *shareholder* and *member* (#6). These terms turned out to be very problematic for translators to disambiguate, partly due to source and target asymmetries, use of terms of national law and semi-legal equivalents, and shifts between a generic and specific usage. A special type of near synonymy appears in a binomial *shareholder or member*, usually as part of the term *shareholders' or members' voting rights*. This phrase is used to signal a broad scope of application to various types of owners, including companies which do not have a share capital. The binomial triggers high variation with

as many as seven variants in Polish translations of legal acts, including the correct variant *akcjonariusze lub wspólnicy* (which can be easily retrieved from IATE, see Figure 4), a range of variants with the semi-legal term *udziałowiec* (*udziałowiec lub wspólnik, udziałowiec / akcjonariusz*) and the misleading term *członek* [member] which does not refer to an owner of an entity under Polish company law (*udziałowiec lub członek; akcjonariusz, wspólnik lub członek; udziałowiec/akcjonariusz czy członek*), as well as the incorrect term *udziałowcy lub członkowie zarządu* [shareholders/partners or Management Board members]. This may be attributed to the translators' insufficient knowledge of English and Polish company law terminology. Similar problems appear when *shareholder* is used alone in a generic sense, triggering a range of options.

Figure 4. A screenshot of IATE entry 1239951 for *shareholder*

792389				
business organisation (4006) [BUSINESS AND COMPETITION]				Council
accounting (4026) [BUSINESS AND COMPETITION]				
en	shareholders	***	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	Council
	shareholders or members	***	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	Council
pl	akcjonariusze lub wspólnicy	***	<input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	COM

These errors do not appear in judgments where an English *shareholder* usually corresponds to *akcjonariusz* (64 times) and in some cases is adjusted to *wspólnik* or is translated more generically as *właściciel* [owner]. Such adjustments are possible thanks to a more tangible context in the form of facts of the case.

Another factor is the complexity of a semantic field. The more overlapping variants there are, the lower conceptual clarity and the higher cross-variation in translation, especially when the target semantic field is equally complex. This may be illustrated with one of the key terms of competition law — *undertaking* and its variants of various origin and granularity which trigger considerable cross-variation (Table 3).

Table 3. Cross-variation of equivalents in the *undertaking* cluster

undertaking	enterprise	company	firm	entity
<i>przedsiębiorstwo</i>	<i>przedsiębiorstwo</i>	<i>spółka</i>	<i>przedsiębiorstwo</i>	<i>podmiot</i>
<i>podmiot gospodarczy</i>	<i>jednostka gospodarcza</i>	<i>przedsiębiorstwo</i>	<i>firma</i>	<i>jednostka</i>
<i>przedsiębiorca</i>		<i>firma</i>		
<i>jednostka</i>		<i>podmiot</i>		

EU texts contain various terms to name an entity, with the dominant form being an *undertaking* (710/3967) and an *enterprise* (366/2045), while *company*, *firm*, *operator*, *entity*, *entrepreneur*, *economic actor* are used occasionally. This lack of conceptual clarity results from a mix of EU and national terms, and makes disambiguation difficult for translators, who introduce variants and errors. In most cases *undertaking*

(an EU term of German provenance denoting a business entity) and *enterprise* (a calque from French) are disambiguated as *przedsiębiorstwo* (687 and 330 times, respectively), which is an EU Polish equivalent for a business entity. Translators occasionally introduce national variants: *podmiot gospodarczy* [business entity], *przedsiębiorca* [entrepreneur] — the national equivalent of EU *przedsiębiorstwo*; or a hyperonym *jednostka* [unit/entity]. In cases where *undertaking* is a broader concept (“single undertaking covers all enterprises”), *enterprise* is disambiguated as *jednostka gospodarcza* [business unit]. There are cases when both terms are full synonyms, e.g. *small enterprise/undertaking*, *medium-sized enterprise/undertaking*, and they are reduced into *małe przedsiębiorstwo* and *średnie przedsiębiorstwo*, respectively. Variation increases with the term *company*, as it is translated with a national legal term *spółka* (which covers both companies and partnerships, 37 times) as well as with a more generic term *przedsiębiorstwo* (5), and, somewhat surprisingly, with a semi-legal colloquial term *firma*, which in legal Polish denotes a company name, not a company as an entity. Another confusing term is *firm* which tends to stand for a partnership and is also incorrectly calqued into *firma*: *companies*, *firms* → *spółki*, *firmy*. These inconsistencies do not make the translators’ task any easier, as translators often do not have access to background information to be able to disambiguate (near-) synonyms.

As already mentioned in passing, the translators’ choices are affected to some extent by genre. Cross-generic variation appears when one genre — usually legal acts — eliminates synonyms but other genres maintain a larger variation as is shown in Table 4 in the example of *illegal* and *unlawful aid*.

Table 4. Variation of equivalents of *illegal* and *unlawful aid*

	EULaw2011-15		EUJudgments2011-15	
<i>illegal aid</i>	<i>pomoc niezgodna z prawem</i>	7	-	
<i>unlawful aid</i>	<i>pomoc niezgodna z prawem</i>	25	<i>pomoc przyznana niezgodnie z prawem</i>	1
	<i>pomoc przyznana bezprawnie</i>	1	<i>pomoc przyznana bezprawnie</i>	6
	-		<i>pomoc bezprawna</i>	5

The English terms have a consistently different distribution in legislation: *unlawful aid* and *to declare an aid illegal*. This variation is neatly eliminated in Polish as *pomoc niezgodna z prawem* [aid inconsistent with law] and *uznaje pomoc za niezgodną z prawem* [declares an aid as inconsistent with law] in the EULaw2011-15 corpus, except for a single intratextual variant of the former *pomoc przyznana bezprawnie* [aid granted illegally] in the preamble of the regulation. Interestingly, Polish judgments do not continue the defined equivalent from the legal acts, but prefer *pomoc przyznana bezprawnie* [aid awarded illegally] and *pomoc bezprawna* [illegal aid], whereas English judgments consistently favour *unlawful aid*.

Variation may also be due to microdiachronic shifts in the use of terms. This may be illustrated with cluster #2 — *grant* and *subsidy*. These terms are regarded in IATE (1239951) as synonyms, with a single recommended Polish equivalent *dotacja* (see Figure 5).

Figure 5. A screenshot of IATE entry 1239951 for *grant* and *subsidy*

★ 1239951				COM
economic support [ECONOMICS > economic policy]				COM
commercial law [TRADE > trade policy]				
en	grant	***		COM
	<u>subsidy</u>	***		COM
	grant aid	***		COM
pl	dotacja	***		COM

The main term, *grant*, is frequent in the EULaw2011-15 corpus. It refers to a type of aid (*reimbursable grant*, *gross grant equivalent*, *cash grant*, *direct grant*) and overlaps with the meaning of *subsidy*, a much less common term used in a more specialised sense in: *wage subsidy* and *interest rate subsidy*.

- **grant** → *dotacja* [grant],
- **subsidy** → *dopłata* [additional payment], *dotacja* [grant], *subsydiowanie* [subsidisation],
interest rate subsidy → *dopłaty do oprocentowania* [additional payments to the interest rate] → *dotacja na spłatę odsetek* [grant for the repayment of interest],
wage subsidy → *subsydiowanie wynagrodzeń* [subsidisation of remuneration].

In 2012-2014 documents, the term *interest rate subsidy* was translated as *dopłaty do oprocentowania*. The year 2014 observes a replacement of this equivalent with *dotacja na spłatę odsetek*, using the same denomination for *subsidy* as for *grant*. This shift creates some clumsiness, as the phrase *aid comprised in grants or interest rate subsidies* becomes *pomoc w formie dotacji oraz dotacji na spłatę odsetek* [aid in the form of grants and grants for the repayment of interest]. Additional variation is found in judgments, in which *subsidy* corresponds to *subwencja* [subsidy] and to a hyperonym *pomoc* [aid].

Another example of a shift in equivalents concerns the term *fine* and shows how such a change may result in an inconsistent elimination of variation across genres. Council Regulation (EC) No 1/2003 introduces two types of penalties: *fines* and *periodic penalty payments*. The former may not exceed 1% of the total turnover in the preceding business year, while the latter may not exceed 5% of the average daily turnover in the preceding business year per day (Articles 23 and 24). Thus, these are not full synonyms and they often appear in a binomial *fines and periodic penalty payments*, where they are translated as *grzywny i okresowe kary pieniężne* [fines and periodic monetary penalties]. The difference is mirrored in EULaw2002-10, except for a single one-off inconsistency, where *fine* is incorrectly translated as *kara pieniężna* in the preamble to Regulation 773/2004 despite a clear reference to Article 23 of Regulation No 1/2003.

- **fine** → *grzywna* [fine], *kara pieniężna* [monetary penalty],
penalty payments → *kara pieniężna* [monetary penalty].

The Polish equivalent of *fine* changes in Directive 2014/104/EU, again despite frequent references to Regulation No 1/2003, where it appears in the context of immunity from fines under leniency programmes related to the involvement in a cartel (the directive does not use the term *penalty payments*). This usage is rendered by translators with a more general term used in national law — *kara pieniężna* [monetary penalty], except for a single case when *grzywna* is retained. This leads to cross-generic variation as later soft law instruments still translate *fine* as *grzywna* in the context of immunity. This concerns the Communication from the Commission (2015/C 256/01) which amends the Commission Notice on Immunity from fines and reduction of fines in cartel cases (C 298/17, 8.12.2006) and hence continues the equivalent *grzywna* in the term *immunity from fines* from the 2006 translation.

7. Conclusions

While most studies conducted so far have focused on variation introduced by translators in target texts, the objective of our study was to empirically verify translators' procedures when they are confronted with variation in source texts. First, the analysis confirms the existence of synonyms in the form of terminological variants in EU English source texts. While a substantial part of them tends to be distributed in different contexts, there are also synonyms with similar distribution which evoke overlapping or identical concepts.

The findings show that semantically related terms may be regarded as one of the causes of variation in legal translation. Translators are not always aware of the existence of synonyms in the same branch of law and hence their treatment of synonymous terms is to some extent unpredictable. This is contrary to the standardisation objective in institutional translation. In some cases, as predicted by Strandvik (2012: 39-40), translators mirror the synonymy; in some cases, synonyms were eliminated or partly reduced. Yet in more cases variation was increased, in particular in large clusters of near-synonyms, which are difficult to control in translation. The factors which affect the translators' choices include: asymmetries between source and target clusters of synonyms, complexity of a semantic field, low termness and microdiachronic shifts. They contribute to lack of conceptual clarity in source texts and may cause terminological inconsistencies, cross-variation and errors.

What is also alarming is that even though translators attempt to reduce source-text variation, they nevertheless introduce a considerable degree of occasional variation within equivalents of a single term, or cross-variation. Instances of such variation were identified for all of the clusters we studied. This suggests that on the one hand, translators tend to disambiguate and reduce variation (simplification); on the other hand, such simplification at one level is complicated with accidentally (or intentionally) introduced variation at other levels (e.g. within a term).

Although we did not observe significant differences between legal acts and soft law instruments, increased but acceptable terminological variation was observed in judgments. The main reason for this seems to be that judgments are functionally different from normative texts and are based on specific facts of the cases heard by the CJEU and hence provide translators with more contextual tips to adapt terminological equivalents to the context. Additionally, judgments are translated by lawyer-linguists who — we hypothesise — may feel that they have more agency

and legal knowledge to depart from the source text. Most adaptations concern shifts in the level of specificity, that is the use of hyponyms, hyperonyms as well as some alignment to national usage. Legal acts are generic and invite more literal techniques.

Our analysis confirms that source-text synonymy causes variation and that, due to asymmetries between languages, it is difficult to control synonymy and standardise variants in translation. Synonymy in source texts may be regarded as one of the root causes of terminological variation in translation. Translators are not always in a position to determine whether a terminological variant is used intentionally to refer to a new supranational concept or whether it is a synonym which may and should be eliminated. Source-text variation further complicates the already difficult task facing EU translators and adds on to their cognitive effort. Terminological resources for institutional translators could pay more attention to (near-) synonymy. Furthermore, more reflection is needed on what triggers variation in translation as synonymy is clearly not the only cause.

The main limitations of this study are the language pair, field and corpus size. Our study addresses translations into Polish, the EU's official language since 2004 and a relatively young Eurolect. Similar trends may be expected to exist in other Eurolects which are at the same stage of development (the 2004 accessions); yet older Eurolects, e.g. French, Dutch, Italian or German, which are more mature and stabilised (cf. Biel 2020), may be more resistant to terminological variation. Considering the wealth of EU documents, our corpus is relatively small and it is quite likely that variation would be larger in a corpus comprising more genres, fields and time ranges. Furthermore, it is unclear to what extent a reduction of synonymy is intentional or accidental; to answer this question, qualitative methods, such as think-aloud protocols or interviews with translators, would have to be applied. Finally, the study was feasible, since we worked with a narrowly defined group of terms on a relatively small corpus. The question remains how to operationalise more extensive studies into term variation spanning a wider range of term categories.

References

- Baaij, C. J. W., *Legal Integration and Language Diversity: Rethinking Translation in EU Lawmaking*. New York: Oxford University Press 2018.
- Bajčić, M., *New Insights into the Semantics of Legal Concepts and the Legal Dictionary* Amsterdam: John Benjamins 2017. DOI: 10.1075/tlrp.17.
- Bajčić, M. and Martinović, A., «A mutual learning exercise in terminology and multilingual law», in: Marino, S., Biel, Ł., Bajčić, M., and Sasoni, V. (eds.), *Language and Law: The Role of Language and Translation in EU Competition Law*. Cham: Springer 2018, 207-223. DOI: 10.1007/978-3-319-90905-9_12.
- Biel, Ł., *Lost in the Eurofog. The Textual Fit of Translated Law*. Frankfurt am Main: Peter Lang 2014. DOI: 10.3726/978-3-653-03986-3.
- Biel, Ł., «Observing Eurolects: The case of Polish», in: Mori, L. (ed.), *Observing Eurolects*. Amsterdam/Philadelphia: John Benjamins 2018, 295-327. DOI: 10.1075/sci.86.12bie.
- Biel, Ł., «Theoretical and methodological challenges in researching EU legal translation. Legal Translation», in: Simonnæs, I. and Kristiansen, M. (eds.), *Current Issues and Challenges in Research, Methods and Applications*. Berlin: Frank & Timme 2019, 25-39.

- Biel, Ł., «Eurolects and EU Legal Translation», in: Meng, J. and Laviosa, S. (eds.), *The Oxford Handbook of Translation and Social Practices*. Oxford: Oxford University Press 2020. DOI: 10.1093/oxfordhb/9780190067205.013.15.
- Biel, Ł., Biernacka, A., and Jopek-Bosiacka, A., «Collocations of Terms in EU Competition Law: A Corpus Analysis of EU English Collocations», in: Marino, S., Biel, Ł., Bajčić, M., and Sosoni, V. (eds.), *Language and Law: The Role of Language and Translation in EU Competition Law*. Cham: Springer 2018a, 249-274. DOI: 10.1007/978-3-319-90905-9_14.
- Biel, Ł., «The Glossary of EU English Competition Collocations and Terms.» in: Marino, S., Biel, Ł., Bajčić, M., and Sosoni, V. (eds), *Language and Law: The Role of Language and Translation in EU Competition Law*. Cham: Springer 2018b, 275-324. DOI: 10.1007/978-3-319-90905-9_15.
- Biel, Ł. and Doczekalska, A., «How do supranational terms transfer into national legal systems? A corpus-informed study of EU English terminology in consumer protection directives and UK, Irish and Maltese transposing acts», *Terminology* 26, 2 (2020), 184-213.
- Biel, Ł., Koźbiał, D. and Wasilewska, K., «The formulaicity of translations across EU institutional genres: A corpus-driven analysis of lexical bundles in translated and non-translated language», *Translation Spaces* 8 (2019), 67-92. DOI: 10.1075/ts.00013.bie.
- Biel, Ł. and Sosoni, V., «EU legal culture and translation in the era of globalisation: the hybridisation of EU terminology on the example of competition law», in: Vogel, F. (ed.), *Legal Linguistics beyond the borders: Language and Law in a World of Media, Globalisation and Social Conflicts. Relaunching the International Language and Law Association*. Berlin: Duncker and Humblot 2019, 208-229. DOI: 10.3790/978-3-428-55423-2.
- Cabré Castellví, M. T., «Theories of terminology: Their description, prescription and explanation», *Terminology* 9 (2003), 163-199. DOI: 10.1075/term.9.2.03cab.
- Chromá, M., «Synonymy and Polysemy in Legal Terminology and Their Applications to Bilingual and Bijural Translation», *Research in Language* 9 (2011), 31-50. DOI: 10.2478/v10015-011-0004-2.
- Dabbah, M. M., *International and Comparative Competition Law*. Cambridge: Cambridge University Press 2010.
- Derlén, M., «A Single Text or a Single Meaning: Multilingual Interpretation of EU Legislation and CJEU Case Law in National Courts», in: Šarčević, S. (ed.), *Language and Culture in EU Law. Multidisciplinary Perspectives*. Farnham: Ashgate 2015, 53-72. DOI: 10.4324/9781315591445-10.
- Dobrić Basaneže, K., «Binomials in EU Competition Law. Language and Law: The Role of Language and Translation in EU Competition Law», in: Marino, S., Biel, Ł., Bajčić, M., and Sosoni, V. (eds.), *Language and Law: The Role of Language and Translation in EU Competition Law*. Cham: Springer 2018, 225-248. DOI: 10.1007/978-3-319-90905-9_13.
- Doczekalska, A., «Legal Languages in Contact: EU Legislative Drafting and Its Consequences for Judicial Interpretation», in: Marino, S., Biel, Ł., Bajčić, M., and Sosoni, V. (eds), *Language and Law: The Role of Language and Translation in EU Competition Law*. Cham: Springer 2018, 163-178. DOI: 10.1007/978-3-319-90905-9_10.

- EU, European Union, *Joint Practical Guide of the European Parliament, the Council and the Commission for persons involved in the drafting of European Union legislation*. Luxembourg: Publications Office of the European Union 2015.
- Goźdz-Roszkowski, S., «Exploring near-synonymous terms in legal language. A corpus-based, phraseological perspective», *Linguistica Antverpiensia, New Series – Themes in Translation Studies* 12 (2013), 94-109.
- Harris, Z. S., «Distributional Structure», *Word* 10 (1954), 146-162. DOI: 10.1007/978-94-009-8467-7_1.
- Kilgarriff, A., «Simple maths for keywords», in: Mahlberg, M., González-Díaz, V., and Smith, C. (eds), *Proceedings of Corpus Linguistics Conference CL2009*. University of Liverpool 2009.
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., and Suchomel, V., «The Sketch Engine: ten years on», *Lexicography* 1 (2014), 7-36. DOI: 10.1007/s40607-014-0009-9.
- Kjær, A. L., «Legal Translation in the European Union: A Research Field in Need of a New Approach», in: Kredens, K. and Goźdz-Roszkowski, S. (eds.), *Language and the Law: International Outlooks*. Frankfurt am Main: Peter Lang 2007, 69-95.
- Koskinen, K., «How to research EU translation?», *Perspectives* 9 (2001), 293-300. DOI: 10.1080/0907676x.2001.9961425.
- Koźbiał, D., «Two-tiered approach to quality assurance in legal translation at the Court of Justice of the European Union», in: Svoboda, T., Biel, Ł., and Łoboda, K. (eds.), *Quality aspects in institutional translation*. Berlin: Language Science Press 2017, 153–172. DOI: 10.5281/zenodo.1048198.
- McAuliffe, K., «Hidden Translators: the Invisibility of Translators and the Influence of Lawyer-Linguists on the Case Law of the Court of Justice of the European Union», *Language and Law / Linguagem e Direito* 3 (2016), 5-29.
- Patel, K. K. and Schweitzer, H., «Introduction», in: Patel, K.K. and Schweitzer, H. (eds.), *The Historical Foundations of EU Competition Law*. Oxford: Oxford University Press 2013, 1-18.
- Prieto Ramos, F., «International and supranational law in translation: from multilingual lawmaking to adjudication», *The Translator* 20 (2014), 313-331. DOI: 10.1080/13556509.2014.904080.
- Prieto Ramos, F. and Guzmán, D., «Legal Terminology Consistency and Adequacy as Quality Indicators in Institutional Translation: A Mixed-Method Comparative Study», in: Prieto Ramos, F. (ed.), *Institutional Translation for International Governance: Enhancing Quality in Multilingual Legal Communication*. London: Bloomsbury 2018, 81-101. DOI: 10.5040/9781474292320.0015.
- Prieto Ramos, F. and Morales Moreno, A., «Terminological innovation and harmonization at international organizations: Can too many cooks spoil the broth?», in: Simonnaes, I. and Kristiansen, M. (eds.), *Legal Translation. Current Issues and Challenges in Research, Methods and Applications*. Berlin: Frank & Timme 2019, 87-110.
- Robertson, C., «EU Multilingual Law: Interfaces of Law, Language and Culture», in: Šarčević, S. (ed.), *Language and Culture in EU Law. Multidisciplinary Perspectives*. Farnham: Ashgate 2015, 33-52. DOI: 10.4324/9781315591445-9.
- Robinson, W., «Translating Legislation: The European Union Experience», *The Theory and Practice of Legislation* 2 (2014), 185-210. DOI: [10.5235/2050-8840.2.2.185](https://doi.org/10.5235/2050-8840.2.2.185).

- Šarčević, S., *New Approach to Legal Translation*. The Hague: Kluwer Law International 1997.
- Šarčević, S., «Challenges to Legal Translators in Institutional Settings», in: Prieto Ramos, F. (ed.), *Institutional Translation for International Governance: Enhancing Quality in Multilingual Legal Communication*. London: Bloomsbury Methuen Drama 2017, 9-24. DOI: 10.5040/9781474292320.0009.
- Scott, M., *WordSmith Tools version 7*. Stroud: Lexical Analysis Software 2016.
- Sosoni, V., «Language and Translation in EU Competition Law: Insights from English, Greek, Italian and Spanish Versions of Legislative Texts», in: Marino, S., Biel, Ł., Bajčić, M., and Sosoni, V. (eds.), *Language and Law. The Role of Language and Translation in EU Competition Law*. Cham: Springer 2018, 179-205. DOI: 10.1007/978-3-319-90905-9_11.
- Stefaniak, K., «Terminology work in the European Commission. Ensuring high-quality translation in a multilingual environment», in: Svoboda, T., Biel, Ł., and Łoboda, K. (eds), *Quality aspects in institutional translation*. Berlin: Language Science Press 2017, 109-121. DOI: 10.5281/zenodo.1048192.
- Strandvik, I., «Legal Harmonization Through Legal Translation: Texts that Say the Same Thing?», in: Baaij, C. J. W. (ed.), *The Role of Legal Translation in Legal Harmonization*. Alphen aan den Rijn: Wolters Kluwer 2012, 25-49.
- Vigier, F. J. and Sánchez Ramos, M. M., «Using parallel corpora to study the translation of legal system-bound terms: The case of names of English and Spanish courts», in: Mitkov, R. (ed.), *Proceedings of Computational and corpus-based phraseology, Second International Conference, Europhras 2017*. Cham: Springer 2017, 260-273. DOI: 10.1007/978-3-319-69805-2_19.
- Woods, L., Watson, P., and Costa, M., *Steiner & Woods EU Law*. Oxford: Oxford University Press 2017. DOI: 10.1093/he/9780198795612.001.0001.