

Łucja Biel

Uniwersytet Warszawski

Mixed corpus design for researching the Eurolect: a genre-based comparable-parallel corpus in the PL EUROLECT project

Mieszana struktura korpusu do badania eurolektu – gatunkowy korpus porównawczo-równoległy w ramach projektu PL EUROLECT

Streszczenie

W artykule opisano mieszaną strukturę gatunkowego korpusu porównawczo-równoległego budowanego w ramach projektu *PL EUROLECT* finansowanego przez NCN (grant SONATA BIS, 2015-2018). Celem projektu jest kompleksowe zbadanie polskiego eurolektu, nowej hybrydowej odmiany języka polskiego powstającej w wyniku tłumaczenia i stosowanej w kontekście unijnym oraz dogłębne zrozumienie procesów i czynników go kształtujących, a także jego wpływu na poakcesyjną polszczyznę urzędową. Podstawą korpusu będzie struktura gatunkowa obejmująca cztery gatunki uznane za reprezentatywne dla komunikacji unijnej (akty prawne, orzeczenia, sprawozdania i urzędowe strony internetowe dla obywateli) podzielone na podgatunki – np. w ramach korpusu aktów prawnych wydzielone zostaną podkorpusy rozporządzeń, dyrektyw i decyzji. Struktura gatunkowa korpusu umożliwi zbadanie zróżnicowania wewnętrznego eurolektu i uzyskanie bardziej precyzyjnych danych ilościowych. Na strukturę gatunkową zostanie nałożony dwujęzyczny korpus równoległy zawierający wyrównane teksty w języku angielskim i polskim oraz jednojęzyczny korpus porównawczy zawierający nieprzetłumaczone teksty administracyjne w języku polskim, a także – jako punkt odniesienia – zrównoważona próba Narodowego Korpusu Języka Polskiego. Mieszana struktura korpusu ma umożliwić badanie dwóch fundamentalnych relacji, tj. ekwiwalencji – relacji eurolektu do tekstów źródłowych (korpus równoległy) oraz dopasowania tekstowego – relacji eurolektu do nieprzetłumaczonych tekstów w języku docelowym (korpus porównawczy). W strukturze korpusu uwzględniony zostanie również korpus diachroniczny polszczyzny urzędowej sporządzony dla poszczególnych gatunków z okresu przedakcesyjnego i poakcesyjnego w celu zbadania wpływu eurolektu na urzędową odmianę języka polskiego. Uzyskane dane ilościowe będą

rejestrować stan eurolektu i polszczyzny w przekroju gatunkowym w konkretnych przedziałach czasowych, i stanowić punkt odniesienia dla innych badaczy. Gatunkowe dane ilościowe otrzymane z analizy korpusowej zostaną poddane triangulacji z danymi jakościowymi (analiza dyskursu, semiotyka społeczna, badania prawnoporównawcze terminologii). Celem metodologicznym jest opracowanie interdyscyplinarnego modelu teoretycznego do badania odmian języka powstających z udziałem tłumaczy.

Keywords: parallel corpus, comparable corpus, genre-based corpus, Polish, English, Eurolect, institutional translation

Słowa kluczowe: korpus równoległy, korpus porównawczy, korpus gatunkowy, język polski, język angielski, eurolekt, tłumaczenie instytucjonalne

1. Introduction

The main objective of this paper is to discuss the mixed design of a translation-driven corpus for researching the Polish Eurolect and changes in administrative Polish after Poland's accession to the European Union. The corpus will be built as part of the project entitled *The Eurolect: an EU variant of Polish and its impact on administrative Polish* (an abbreviated name: *PL Eurolect*), funded by the Polish National Science Centre (NCN) with a SONATA BIS 4 grant awarded to the author to build a research team in the years 2015-2018. It is a follow-up of the Eurofog project funded by the Polish Ministry of Science and Higher Education in 2010-2012, the results of which were published in the book *Lost in the Eurofog. The Textual Fit of Translated Law* (Biel, 2014). The Eurofog project, which analysed the *JRC Acquis* corpus of regulations and directives (translations up to 2006) against national Polish law, is treated as a pilot since it uncovered a number of areas requiring deeper studies and, above all, the need for a better calibrated and controlled genre-based corpus of a mixed design.

2. The Eurolect – a hybrid variant of national languages and its genres

EU texts are produced in a multilingual environment with 24 official languages, which implies a constant switching and 'fusion' of languages. They are often a result of a delicate political compromise between 28 Member States and are produced under a complex array of political, procedural, institutional, legal and cultural constraints. EU texts are considered to have developed a specific language, perceived as a new legal variant of the official languages (cf. Koskinen, 2000: 53; Salmi-Tolonen, 2004: 1187). The hybridity of EU discourse is closely connected with multilingualism, leading to a constant interplay between the supranational and national elements in translation. On the one hand, EU

drafting and translation demonstrate reduced embedding in national cultures and their neutralization; on the other hand, such de-territorialization creates a new territory, with the pan-European culture based on *acquis* – the body of EU law, and constituent national cultures.

Despite its scale and social impact, the Eurolects are a surprisingly rare object of research within Translation Studies and other disciplines. The existing studies are usually small-scale and descriptive, rarely based on systematic empirical research, and, to a certain degree, findings are repetitive or superficial. In fact, very little is known about the Eurolects, not only in respect of Polish, for which this is a new phenomenon, but also for other official EU languages. Such knowledge is a prerequisite for understanding Eurolects and their impact on national languages. Another shortcoming is that most studies are legicentric as they concern EU law (including my previous project Eurofog), with other genres being heavily under-researched. This project intends to fill in this gap in order to get an in-depth insight into various generic layers of the Polish Eurolect, which may be generalisable to the Eurolects of other official languages.

The EU discourse community has developed its specific genres which have evolved as a collective multilingual effort, with the predominant influence of the procedural languages – French, English and, to a lesser extent, German. A genre is understood as “a more or less stabilized and habitual linguistic way of acting and interacting, characterized by a distinctive linguistic form or structure, associated with specific communicative purposes, and with particular social or institutional contexts” (Fairclough, 2006: 32). EU genres are therefore conventionalised uses of language associated with distinct goal orientation and practices. Four genres have been selected for the analysis as most prototypical and hence representative of EU communication: legislation, judgments, reports, and official websites for citizens. They are subject to different language policies and translation procedures.

EU legislation enjoys a very special status: it is adopted in all 24 official languages and is applicable in 28 Member States. Under the principles of multilingualism and equal authenticity, all language versions have an authoritative status, which means that they are equally valid and presumed to have the same meaning; in fact, they altogether form a single legal instrument (cf. Šarčević, 1997: 64). The ultimate goal is to ensure the uniform interpretation and application of the EU law in all the Member States (cf. Šarčević, 1997: 73). Drafting and translation are concurrent, multistage and multilingual (cf. Doczekalska, 2009: 360): proposals are drafted mainly in English and next translated into official languages in most cases in-house by the EU institutions.

The genre of **judgments** will be illustrated with the Court of Justice of the European Union, which has only one working language – French. Its Rules of Procedure provide for “the language of a case”, which may be any of the 24 EU official languages. The language of the case is used in the parties’ written and oral pleadings and in the minutes and decisions of the Court. Only documents prepared in the language of the case are authentic. In the case it is impossible to provide direct translation, pivot languages are used, e.g. German for Polish (cf. McAuliffe, 2012). Case law is translated mainly in-house by lawyer linguists who have legal training and know at least 2 official languages.

Reports of EU institutions are an example of professional communication between experts. As documents which are not legally binding, they are usually available in English, French and German. Reports may be translated into other languages depending on the target audience and importance of information. They may be translated in-house or by external contractors.

EU institutional **websites** are the main channel of communication between the EU (the institutions) and citizens of the European Union. They provide information about the EU and inform citizens of their rights and obligations; but their ‘hidden’ task is to promote the positive image of the European Union. The most representative example is EUROPA, an official website of the European Union run by the Communication Department of the European Commission on behalf of the EU institutions. Websites are prepared by the dedicated service with the requirement that they should be citizen-oriented, user-friendly and euro-jargon free (cf. European Commission, 2009). The language policy is evidence-based to decide which content is translated into all languages (i.e. if users would be seriously disadvantaged or if it is legally required) and which into only most frequent or selected languages (e.g. short-lived or specialised content)¹. Websites are translated by a specialised unit within the Directorate-General for Translation.

3. Research objectives of the PL Eurolect project

As for the Polish Eurolect, compared to the Eurolects of the old Member States (EU-15), it is a newly emerging phenomenon, a hybrid variety of administrative Polish. It started to form in the early 2000s when Poland began to translate *acquis communautaire*, the EU body of law, as a precondition for the 2004 accession. After the accession the translation process was institutionalised, when it was taken over by the EU institutions, such as the European Commission, the Council of the European Union, the European Parliament, and the Court of Justice. The nature of the Polish Eurolect is constrained by the fact

¹ http://ec.europa.eu/ipg/content/multilingualism/index_en.htm.

that it is an outcome of **translator-mediated communication**, translated from the EU procedural languages, currently mainly English, and a by-product of the unequal interaction between a majority EU culture and a minority Polish culture (cf. Biel, 2014).

Having regard to the foregoing, the main objective of the PL Eurolect project is (1) to extensively investigate the Polish Eurolect in order to understand the processes and factors behind its formation, and (2) to track the impact of the Eurolect on post-accession Polish. This objective will be researched via the following questions:

1. **External variation (the textual fit):** How does the Eurolect differ from naturally occurring, nontranslated administrative Polish?
2. **Internal generic variation:** How does the Eurolect vary internally across the four genres (legislation, judgments, reports, official websites for citizens)?
3. **Variables:** How is the Eurolect affected by a genre, source language, institutionalisation of translation process, translator profile and translation universals²?
4. **Europeanisation of administrative Polish:** How has post-accession Polish been affected by the huge inflow of EU translations (a comparison of pre-accession Polish (1999/2000) and post-accession Polish (2015))?

The methodological objective is to develop and test an interdisciplinary theoretical model for researching translator-mediated variants of language, a model which would be of general applicability extending beyond a specific language and genre.

4. The mixed corpus design

The project will be conducted in the area of **Corpus-Based Translation Studies** (CBTS), a strand of Translation Studies which applies **corpus linguistics** as methodology. The corpus-based methodology allows for inductive data-driven studies based on quantitative data. Its major advantages include: reduced subjectivity and the potential to verify hypotheses systematically on much more extensive material.

Although corpora of EU translations are abundant, the previous project has shown that the available resources should be used with care and in order to

² The hypotheses of translation universals (also more recently known as typical features, tendencies, patterns of translations) include: explicitation, simplification and disambiguation, normalisation, standardisation, conventionalisation, levelling-out (lower variation), unique items, and untypical collocational patterns (cf. Baker, 1993; Chesterman, 2004).

research more detailed questions, corpora have to be built anew based on carefully controlled criteria.

4.1. Genre-based corpus

The rationale for designing a genre-based corpus is connected with the importance of the concept of a genre as revealed by recent corpus-based studies in the domain. Studies into legal language have demonstrated a high variation of lexical bundles across legal genres (cf. Goźdz-Roszkowski, 2011); while empirical studies into translation show that differences between translations and nontranslations (in particular features of translations) are dependent on genres (Teich, 2003: 147; Delaere et al., 2012; de Sutter et al., 2012).

The core corpus of the Polish Eurolect has a genre-based structure (Figure 1). It will comprise the four prototypical genres of EU communication: legislation, judgments, reports and official websites for citizens (cf. Section 2). The core corpus is intended to be sufficiently large to allow for a variety of data. Relevant methods and tests will be applied to ensure its representativeness, balance and comparability. The detailed structure of the corpus is shown in Table 1 below.

EU legislation	<p>A subcorpus of regulations (which are directly applicable) and a subcorpus of directives (which are subject to transposition into national law). However, other types of instruments may be added at a later period. Since the pilot study has shown large differences in the distribution of lexico-grammatical patterns in non-enacting (preambles) and enacting terms (the normative part), the directives and regulations will be divided into 3 sections – preambles/citations, enacting terms and annexes, to ensure a better comparability of the Eurolect to the national language of the law, which has a different structure without extensive non-normative preambles. It was identified as one of the limitations in the Eurofog project (cf. Biel, 2014: 309).</p> <p>The corpus will also be divided into two time spans to measure the effect of institutionalisation on the translation process and the evolution of the Polish Eurolect from its formative stage to the current developed stage. The 1A corpus will contain pre-accession translations coordinated by the Polish government while the 1B corpus will contain post-accession translations from the period 2014-2015 coordinated by the EU institutions.</p> <p>1A: <i>JRC Acquis</i> corpus, Version 3.0, which contains Polish regulations (ca. 14 million words) and Polish directives (ca. 7 million words). The <i>JRC Acquis</i> corpus was built by the European Commission's Joint Research Centre and it contains texts from 1958 to 2006. The corpus design is discussed in Steinberger et al. (2006). The corpus was used in the pilot study; in this study we intend to extract pre-accession translations, that is up to 2004.</p> <p>1B: the corpus of Polish-language regulations and directives with a time span of 2014-2015, when translation quality may be deemed to have stabilised. The corpus will be compiled as part of the project. Files will be downloaded in html from the EUR-Lex portal.</p>
-----------------------	---

EU judgments	The Polish version of judgments and other decisions of the Court of Justice and the General Court to be downloaded from the Curia portal and sorted according to the so-called authentic language (language of the case, source language) – English, German and French. Such sorting will enable us to analyse the impact of source language on Polish translations (i.e. SL-specific interference)
EU reports	Technical reports prepared by the European Commission and other EU institutions; expert-to-expert communication
Official websites for EU citizens	Polish versions of websites of EU institutions, e.g. Europa portal, the website of the European Parliament and the European Commission. Special attention will be paid to pages which explain legal rights and obligations (semi-legal language).

Table 1. The core corpus: the genre-based structure of the Polish Eurolect corpus

Except for the EU law component, the remaining sections of the Eurolect corpus will exclude pre-accession translations.

4.2 Parallel, comparable and reference corpora

Owing to the involvement of translation, the core corpus of the Polish Eurolect will be embedded in a combination of monolingual comparable/reference and bilingual parallel corpora. Corpus-Based Translation Studies tend to use either the technique of parallel corpus *or* of a comparable corpus. While comparable corpora are useful for analysing the textual fit of translations, that is how they differ from nontranslations in the target language, parallel corpora give insight into equivalence, that is the relation of translations to source texts. Both relations – textual fit and equivalence – are fundamental in translations and essential to understanding their nature (cf. Chesterman, 2004). The technique of researching translations against non-translations via comparable monolingual corpora was pioneered by Mona Baker in the early 1990s (1993); however, it has lately been criticised as one-sided and methodologically flawed due to the exclusion of source texts (STs) from the study of translations and the resulting impossibility to account for ST interference in translated language (cf. Bernardini, Zanettin, 2004: 59; Pym, 2010: 82; Bernardini, Ferraresi, 2011: 228). More recent approaches to translation recommend combining two types of corpora, that is comparable and parallel ones. This approach has for example been advocated by Hansen-Schirra and Teich, who refer to such a mixed-design corpus as a *multilingually comparable corpus* (2009: 1162). In my opinion this name, which is popular especially among German scholars, does not fully reflect the nature of the corpus; therefore, I propose to use a more explicit name – **a comparable-parallel corpus**. The structure of the Eurolect comparable-parallel corpus is shown in Figure 1.

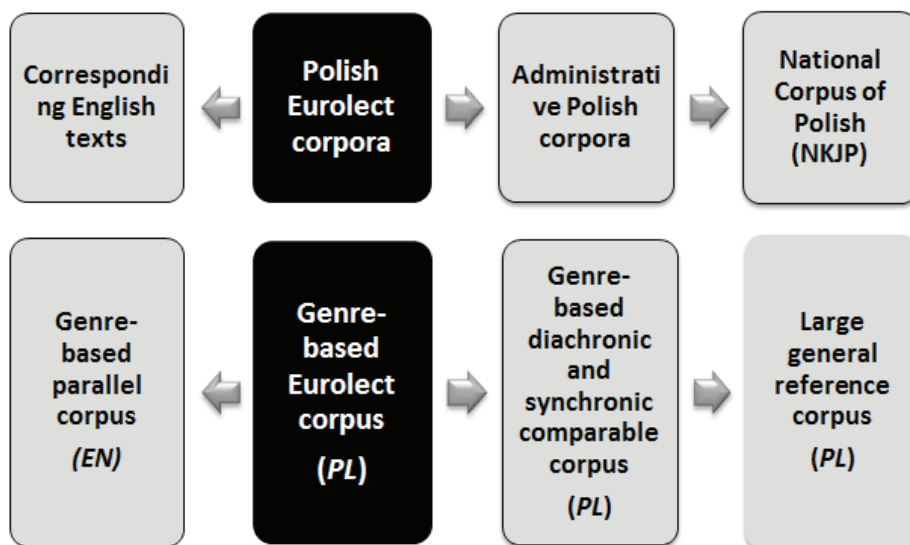


Figure 1. A genre-based comparable-parallel corpus: the PL Eurolect corpus

As for the **parallel corpus**, it will comprise the core corpus of the Polish Eurolect aligned with corresponding English texts for all the four genres. In most cases, these are “source texts”, especially for the pre-accession period; however, it may not be excluded that translations were rendered from other languages, in particular French or that a text went through a mixture of languages (the multilingual drafting process). The selection of English also accounts for the fact that it dethroned French as the key working language in the 2000s and has become a *lingua franca* in the multilingual EU (cf. Pozzo, 2012). Since the techniques of analysing parallel corpora are far less sophisticated than those of comparable corpora, the parallel corpus will be smaller to enable an in-depth analysis and will be built only for selected sections of the core Eurolect corpus. The parallel corpus will be used to study key areas of interference, causes of overrepresentation and other typical features of translations as identified through comparable corpora. Thus, it will be built at a later phase of the project in 2016-2017.

The **comparable corpus of Administrative Polish** will be built according to the same design criteria as the core Eurolect corpus to ensure that, wherever possible, the two corpora “match each other in terms of proportion, genre, domain and sampling period” (McEnery et al., 2006: 48). The corpus will be genre-based and will contain nontranslated administrative Polish. We will aim at a similar size of corresponding subcorpora in the Eurolect corpus and in the Administrative Polish corpus. The corpus will have a synchronic and diachronic component, that is it will be split into pre-accession and post-accession sections to study the impact of the Eurolect on administrative Polish (the provisional

sampling points will be 1999/2000 and 2014-2016). A reservation should be made that the full comparability of corpora in Translation Studies is not possible due to the culture-specific nature of genres (cf. Granger, 2003: 19); in this case it should be borne in mind that EU texts and national texts have a different context of use – supranational and national, respectively.

Polish national legislation	The Polish Law Corpus (PLC) compiled by the author in 2011. The corpus size is 7 million words, 755 statutes (more details see Biel, 2014). The corpus will be updated as of 2015 and a smaller control sub-corpus will be built with the same time span as the 1B corpus of EU law.
Polish judgments	Judgments of the Polish Supreme Court (Sąd Najwyższy), Civil Chamber, obtained from Professor Rafał Górski. The Civil Law Chamber Corpus has 3 million tokens and comprises 1,577 decisions, such as <i>postanowienie</i> , <i>uchwała</i> , <i>wyrok</i> , ranging from 2000-2010. The Supreme Court is the topmost tier of the Polish court system. It has powers to shape general rules of law and may resolve discrepancies in interpretation of law found in judgments of less senior courts. On these grounds the Supreme Court decisions are deemed to be most comparable to EU decisions.
Polish governmental reports	Reports published by the Polish government on comparable thematic areas.
Official websites for Polish citizens	Websites of the Polish government and other national institutions communicating with Polish citizens.

Table 2. A comparable corpus: the genre-based structure of the (national) Administrative Polish corpus

The corpus design also covers a **big general reference corpus** to avoid the “difference mindset” (cf. Baker, 2010: 153) and enable the identification of similarities between the corpora of administrative discourse. It will be a balanced sample of the *National Corpus of Polish (NKJP³)*, a big corpus of contemporary Polish which was released in 2012 and its time span ranges from 1945 to 2011. The balanced version has 240.2 million words and consists of books 29%, the press 50%, other written texts (administrative, letters) 4%, Internet 7%, spoken texts 10%. It will function as a representative sample of contemporary Polish and a benchmark for the interpretation of translation data.

5. Triangulation of data and methods

The project will not rely solely on quantitative data but, given the multifarious nature of translations, it will attempt at triangulation as recently postulated by a number of scholars (cf. Zanettin, 2012: 12). It will cover both **data and methodological triangulation**: that is, quantitative data from a broad range of comparative and parallel corpora across genres and qualitative data obtained via

³ www.nkjp.pl

small-scale analyses, applying the functional linguistics methodology (mainly discourse analysis), to study certain prominent aspects of the Eurolect, such as legitimation strategies and rhetorical patterns, the multimodal analysis of websites (social semiotics), and comparative legal analysis of supranational and national terms. It is believed that this procedure will enable the team to obtain a deep insight into the Polish Eurolect.

6. Conclusions

The project will investigate, describe and help understand the new linguistic phenomenon – the Polish Eurolect, which still has not been researched empirically on a large scale. The results of the project may be relevant for translating institutions and the EU public at large. Given the unprecedented scale of EU translations and their political, economic and societal implications, it is crucial to empirically identify ways of improving the quality of translations and increase their communicative potential by minimising departures from the conventions of Polish, and hence to reduce the colonisation of administrative Polish by the Eurolect. The knowledge about processes, factors, key areas of interference and distortions typical of the translation process will help to improve our understanding and interpretation of EU law, case law and other documents.

Acknowledgment

This study was financed by research grant no. 2014/14/E/HS2/00782 from the National Science Centre, Poland.

References

- BAKER, Mona (1993): *Corpus Linguistics and Translation Studies. Implications and Applications*. In: Mona BAKER, Gill FRANCIS, Elena TOGNINI-BONELLI (eds): *Text and Technology: In Honour of John Sinclair*. Amsterdam: John Benjamins, 233–250.
- BAKER, Paul (2010): *Sociolinguistics and Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- BERNARDINI, Silvia, FERRARESI, Adriano (2011): Practice Description and Theory Come Together – Normalization or Interference in Italian Technical Translation? *Meta: Translators' Journal* 56(1), 226–246.

- BERNARDINI, Silvia, ZANETTIN, Federico (2004): When is a universal not a universal? Some limits of current corpus-based methodologies for the investigation of translation universals. In: Anna MAURANEN, Pekka KUJAMÄKI (eds.): *Translation Universals. Do they exist?* Amsterdam: John Benjamins, 51–62.
- BIEL, Łucja (2014): *Lost in the Eurofog. The Textual Fit of Translated Law*. Frankfurt am Main: Peter Lang.
- CHESTERMAN, Andrew (2004): Hypotheses about translation universals. In: Gyde HANSEN, Kirsten MALMKJÆR, Daniel GILE (eds.): *Claims, Changes and Challenges in Translation Studies. Selected contributions from the EST Congress, Copenhagen 2001*. Amsterdam: John Benjamins, 1–13.
- DE SUTTER, Gert, DELAERE, Isabelle, PLEVOETS, Koen (2012): Lexical lectometry in corpus-based translation studies. Combining profile-based correspondence analysis and logistic regression modeling. In: Michael P. OAKES, Meng Ji (eds.): *Quantitative Methods in Corpus-Based Translation Studies: A practical guide to descriptive translation research*. Amsterdam: John Benjamins, 325–346.
- DELAERE, Isabelle, DE SUTTER, Gert, PLEVOETS, Koen (2012): Is translated language more standardized than non-translated language? Using profile-based correspondence analysis for measuring linguistic distances between language varieties. *Target* 24(2), 203–224.
- DOCZEKALSKA, Agnieszka (2009): Drafting and interpretation of EU law – paradoxes of legal multilingualism. In: Günther GREWENDORF, Monika RATHERT (eds.): *Formal Linguistics and Law*. Berlin: de Gruyter, 339–370.
- EUROPEAN COMMISSION (2009): *Web translation as a genre*. Studies on translation and multilingualism. 3/2009. http://bookshop.europa.eu/is-bin/INTERSHOP.enfinity/WFS/EU-Bookshop-Site/en_GB/-/EUR/View-Publication-Start?PublicationKey=HC8009160, (1 March 2016); DOI: 10.2782/26441.
- FAIRCLOUGH, Norman (2006): Genres in Political Discourse. In: Keith BROWN (ed.): *Encyclopedia of Language and Linguistics*. Vol. 5. Oxford: Elsevier, 32–38.
- GOŹDŹ-ROSKOWSKI, Stanisław (2011): *Patterns of Linguistic Variation in American Legal English. A Corpus-Based Study*. Frankfurt am Main: Peter Lang.
- GRANGER, Sylviane (2003): The corpus approach: a common way forward for Contrastive Linguistics and Translation Studies? In: Sylviane GRANGER, Jacques LEROT, Stephanie PETCH-TYSON (eds.): *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*. Amsterdam: Rodopi, 17–29.

- HANSEN-SCHIRRA, Silvia, TEICH, Elke (2009): Corpora in human translation. In: Anke LÜDELING, Merja KYTÖ (eds): *Corpus Linguistics. An International Handbook*. Berlin: de Gruyter, 1159–1175.
- KOSKINEN, Kaisa (2000): Institutional Illusions. Translating in the EU Commission. *The Translator* 6(1), 49–65.
- MCAULIFFE, Karen (2012): Language and Law in the European Union: The Multilingual Jurisprudence of the ECJ. In: Lawrence M. SOLAN, Peter M. TIERSMA (eds.): *The Oxford Handbook of Language and Law*. Oxford: Oxford University Press; DOI 10.1093/oxfordhdb/9780199572120.001.0001.
- MCENERY, Tony, XIAO, Richard, TONO, Yukio (2006): *Corpus-Based Language Studies. An advanced resource book*. London: Routledge.
- POZZO, Barbara (2012): English as a Legal Lingua Franca in the EU Multilingual Context. In: Cornelis J. W. BAAIJ (ed.): *The Role of Legal Translation in Legal Harmonization*. Alphen aan den Rijn: Wolters Kluwer, 183–202.
- PYM, Anthony (2010): *Exploring Translation Theories*. London: Routledge.
- SALMI-TOLONEN, Tarja (2004): Legal linguistic knowledge and creating and interpreting law in multilingual environments. *Brooklyn Journal of International Law* 29(3), 1167–1191.
- ŠARČEVIĆ, Susan (1997): *New Approach to Legal Translation*. The Hague: Kluwer Law International.
- STEINBERGER, Ralf, POULIQUEN, Bruno, WIDIGER, Anna, IGNAT, Camelia, ERJAVEC, Tomaž, TUFIŞ, Dan, VARGA, Dániel (2006): The JRC-Acquis: A multilingual aligned parallel corpus with 20+ languages. In: *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC'2006)*. Genoa, Italy, 24–26 May 2006. http://langtech.jrc.it/Documents/0605_LREC_JRC-Acquis_Steinberger-et-al.pdf, (1 March 2016).
- TEICH, Elke (2003): *Cross-Linguistic Variation in System and Text: A Methodology for the Investigation of Translations and Comparable Texts*. Berlin: de Gruyter.
- ZANETTIN, Federico (2012): *Translation-Driven Corpora. Corpus Resources for Descriptive and Applied Translation Studies*. Manchester: St. Jerome.