

A Corpus-Based Approach to Examining Terminological Variation in EU Law

Clay, Edward

DOI:

[10.14762/jll.2022.142](https://doi.org/10.14762/jll.2022.142)

License:

Creative Commons: Attribution (CC BY)

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Clay, E 2022, 'A Corpus-Based Approach to Examining Terminological Variation in EU Law', *International Journal of Language & Law*, vol. 11, pp. 142-162. <https://doi.org/10.14762/jll.2022.142>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

A Corpus-Based Approach to Examining Terminological Variation in EU Law

*Edward Clay**

Abstract

This paper presents an original methodology for examining terminological variation in EU legal texts using techniques from corpus linguistics and terminology studies. It describes an innovative approach which enables researchers to obtain robust, empirical data on the existence and extent of intra-lingual and inter-lingual terminological variation in EU law. It firstly examines the nature of terminological variation and then sets out a methodology which can be applied to study such variation in EU legal texts to produce quantitative results. It then presents a pilot study exploring migration terminology in English and Italian EU texts to illustrate how this methodology can be applied in practice before discussing the most significant results produced and their wider implications. The pilot study uses the diachronic corpus-based approach described in the methodology to identify variation in terminology both within and between the languages under investigation. This paper shows the effectiveness and application of this novel methodology, clearly demonstrating the degree of intra-lingual and inter-lingual terminological variation present in EU law.

Keywords

terminology, EU law, corpus linguistics, migration, terminological variation

* University of Birmingham, UK, exc806@student.bham.ac.uk. The author would like to thank the anonymous reviewers for their invaluable input, as well as his children for keeping their Dad on his toes and Dr Sofia Malamatidou for her helpful suggestions. Research for the article has been kindly supported by the Wolfson Foundation.

1. Introduction

The law exists and is expressed through language, meaning that “law cannot be understood without an appreciation of the nature of language” (Ainsworth, 2014: 53). Law and language are inextricably intertwined. Therefore, gaining an in-depth understanding of how the language of the law works and evolves is an essential part of legal research. This contribution takes an empirical approach towards examining law and language, setting out a methodology for researching terminological variation in legal texts produced by the European Union. Terminological variation can be defined as the phenomenon of various terms being used to refer to the same concept. This variation can be both intra-lingual (within individual languages) and inter-lingual (across different languages). This paper presents the results of a pilot study examining variation in migration-related terms in English and Italian to illustrate how this methodology can be implemented. By harnessing an interdisciplinary approach which takes inspiration from methods previously applied in the fields of translation studies, corpus linguistics and terminology studies, it is possible to gain new insights on the dynamics at play between law and language. The objective of this paper is to demonstrate, by way of a corpus-based study, how such empirical research can be carried out and the quantitative results and analysis which can be derived from it.

This research takes a novel approach to studying both intra-lingual and inter-lingual terminological variation in the EU context. The complexity involved in the EU translation process and the vagueness around the notion of source language and source text have long been upheld as reasons which make effective, corpus-based approaches examining variation across different EU languages difficult to implement (Fischer, 2010). However, increasing numbers of researchers in this field are adopting corpus-based approaches to yield new insights into legal translation practices in multi-lingual jurisdictions (Prieto Ramos, 2019; Trklja & McAuliffe, 2019). The terminology used in EU law in all of its official languages is extremely influential, given the prestige of the EU institutions and its wide-reaching application. It also has an impact on legal terminology used outside the EU’s borders, for example in accession countries (Čavoški, 2018), and can have a significant influence on the language used by ordinary European citizens (Mariani, 2018). The importance of EU terminology both within and beyond EU legal circles cannot be overstated. This makes it all the more important to gain greater clarity on how it develops and evolves over time. An empirical approach, such as the one presented in this paper, can provide robust, quantitative data to support hypotheses and enhance qualitative observations on terminology use in legal texts. The results of the pilot study using the methodology described in this paper demonstrate the effectiveness and usefulness of this approach to researching terminological variation in EU legal language.

The methodology adopted in this research aims to capture terminological variation in EU legal texts in English and Italian, which may be synchronic (at a particular point

in time) and diachronic (at two different points in time). It is designed in such a way as to provide quantitative data on both variation within the individual languages as well as variation between the two languages. In order to illustrate how the methodology works in practice, this pilot study specifically focuses on terminology related to the issue of migration. Migration terminology was chosen as a focus of this study for two key reasons. Firstly, it constitutes a potential area of terminological variation at the EU level, since migration has only been the subject of formal legislation since the early 1990s. The Maastricht Treaty, signed in 1992, effectively created a formal intergovernmental system for Justice and Home Affairs cooperation (Peers, 2011), building on earlier informal initiatives in the area. Article K.1 of Title VI of the Treaty, officially entitled the Treaty on European Union, established that “for the purposes of achieving the objectives of the Union, in particular the free movement of persons”, Member States shall henceforth regard the areas of asylum policy, external border controls and immigration policy as “matters of common interest” (Treaty on European Union, 1992). Terminology in the earlier stages of development is more likely to show variation and it is therefore reasonable to expect some degree of terminological variation in this domain, even over a short period, given its relative newness within the EU legal context (Picton, 2014). Secondly, terminology in this area is particularly important since it can have a significant impact on the lived experiences of those who are subject to it and on the perceptions of host communities. Migration terminology is particularly important since the labels applied to the people and processes involved can have huge implications on the legal and moral obligations which receiving countries feel towards them (Sigona, 2018). Indeed, it was highlighted in a speech by the UN Special Rapporteur on the Human Rights of Migrants, François Crépeau in 2013 that

incorrect terminology contributes to negative discourses on migration, reinforces negative stereotypes against migrants, and legitimates a discourse of the criminalisation of migration which in turn contributes to further alienation, marginalisation, discrimination and violence against migrants.

These factors make EU migration law a particularly interesting area of study for research into terminological variation and further emphasises the importance of a rigorous, empirical approach to this area, since the impact of terminological variation in this area is both deep and wide-reaching.

In practical terms, the pilot study presented in this paper seeks to provide robust, quantitative data to investigate (i) whether migration-related terminology used in English and Italian EU legal texts shows variation during the time period studied and (ii) whether the terminology used in English and Italian has become more similar over time. This pilot study therefore seeks to identify changes in migration-related terminology in EU legal texts in English and Italian and to determine whether there has been any terminological convergence between these languages with regard to the specific set of terms under investigation. Firstly, this paper focuses on approaches to terminological research and terminological variation. It then describes the various stages in-

volved in the empirical methodology used and subsequently presents the results of the pilot study along with an analysis of the quantitative findings produced. Finally, it offers a discussion of the results, their implications and considers the possible causes of the terminological variation identified.

2. Terminological Variation

Traditional notions of terminology are largely based on the General Theory of Terminology (GTT) developed by the Vienna School for Terminology and pioneered by Eugen Wüster. This theory focuses largely on the standardisation of terminology within and across languages, advocating the use of one term for one concept, and tends to be applied mainly to scientific and engineering texts (e.g. Wüster, 1968). GTT holds that concepts are clear-cut and unambiguously delineated based on their characteristics and that each concept should be designated by only one term (Felber, 1984). However, more recent research has questioned the foundations of GTT by applying rigorous corpus-based studies to terminological questions (e.g. Bertaccini et al., 2010; Fernández-Silva & Kerremans, 2011). Temmerman (2000: 15) argues that “Traditional Terminology confuses principles, i.e. objectives to be aimed at, with facts which are the foundation of a science”. While practical and useful when applied to technical domains such as science and engineering, GTT fails to capture how terminology operates in other fields, such as law, where terms are often used to describe more abstract concepts. While traditional terminologists may advocate the universal application of the ‘one term, one concept’ principle, this does not describe how specialised language works in practice. This new approach to terminology has shown that rather than a neat and ordered system, more often than not categories are in fact “fuzzy and cannot be absolutely classified by logical and ontological means” (Temmerman, 2000: 16). Consequently, there is a range of possibilities available when naming concepts, which gives rise to situations where different denominations refer to the same concept on a cognitive level (Freixa Aymerich et al., 2008). Also, with regard to diachronic variation, as our knowledge about a given subject evolves over time, the terminology we use to talk about it also tends to change (Picton, 2014).

As a result, this study takes the reality of terminological variation in legal texts into account by looking beyond the traditional notion of a univocal relationship between concept and term to a less restrictive framework, which allows for terminological variation and synonymy. This also reflects the approach to terminology adopted in similar studies examining EU law (e.g. Loupaki, 2018). The corpus-based methodology set out in this paper takes the linguistic forms, i.e., the terms themselves, which appear in the texts as the starting point of the investigation from which to construct conceptual categories (Temmerman, 2000). The research is driven by the terms actually used rather

than taking abstract concepts themselves as the starting point of the study. The notion of terminological variation is used throughout this article, as opposed to onomasiological variation or denominative variation, to emphasise the specific focus on terms, at the exclusion of other units of language. This variation can be both synchronous, with alternative terms for a single concept existing concurrently, and diachronous, with different terms for a concept emerging and disappearing over a period of time.

In contrast to general-use words, terms can be thought of as “a lexical unit consisting of one or more than one word which represents a concept inside a domain” and terminology as “the vocabulary of a subject field” (de Bessé et al., 1997: 152, 154). The Best Practice for Terminologists published for users of the EU’s Interactive Terminology for Europe (IATE) database makes the concise distinction that “terminology deals with special languages while lexicology deals with everyday language” (Translation Centre for the Bodies of the European Union, 2021). It is therefore this specialised function of terminology which enables it “to gain the precision necessary for expressing restricted meaning” (Kaguera, 2002: 15). Previous research has sought to explore the relationship between terms and the concepts to which they refer. Concepts are given linguistic expression through terms in order to make the world intelligible and to make communication meaningful. Pearson (1998: 11) describes concepts as “mental constructs to which we assign labels”. In specialist discourse, it is these ‘labels’ which constitute the terms used to refer to the abstractions in our mind.

When examining terminology, it is important to distinguish between the concept itself in the physical world or our mind and its linguistic realisation. A concept can be thought of as “an image formed in the human mind that helps us to generate knowledge about the world by organizing, naming and giving meaning to its features” (Berenskoetter, 2016: 4). Although a concept is connected to a term, it cannot be equated to a concrete and narrow linguistic form. For instance, the concept of nationality, beyond its purely linguistic form, is also expressed in the physical world through flags, passports and border walls. Equally, however, the sphere occupied by the concept cannot be entirely separated from the linguistic sphere (Olsen, 2012). Indeed, concepts cannot be communicated or properly conceived without language (Temmerman, 2000). The connection between the two therefore lies in the relationship between the term and the entity to which it refers. This research acknowledges this important distinction between concept and term by adopting a methodology which categorises terms into conceptual clusters, as described in more depth in section 3 below, in a similar way to the creation of lexical profiles presented by Speelman, Grondelaers and Geeraerts (2003: 318). For example, the concept BICYCLE can be expressed in English by the term variants *bicycle*, *bike*, *pedal bike* or *cycle* depending on various factors. These terms are therefore all part of the conceptual cluster BICYCLE. The approach implemented in this research therefore recognises and visualises the relationship between the highest level of abstraction and the labels we use to describe them.

Despite this evolving and changeable nature of terminology, given the system-bound nature of law, wherever possible, legal terms should be used consistently, “clearly referring to and activating relevant concepts to control phenomena typical of natural languages, such as synonymy (variation) and polysemy” (Biel & Koźbiał, 2020: 70). Specifically with regard to EU law, consistency of terminology is often the overriding objective, even occasionally at the expense of strict accuracy or idiomatic nature of the terminology (Stefaniak, 2017: 117). Despite such attempts and the implementation of systems to enhance standardisation and minimise instability in terminology, such as the IATE database, terminological variation persists in EU texts (Strandvik, 2012). Previous corpus-based studies into terminological variation in EU texts have identified widespread instability and use of synonyms in various fields, including competition law (Biel, Biernacka & Jopek-Bosiacka, 2018), names of national criminal courts (Vigier & Sanchez-Ramos, 2017) and international law (Prieto Ramos & Morales Moreno, 2019). This research therefore seeks to build on these previous findings by presenting a new, empirical approach to studying terminological variation in EU legal texts.

3. Methodology

3.1. Corpus Design

This research study uses a corpus-based approach. A corpus can be simply defined as “a finite body of text, sampled to be maximally representative of a particular variety of language, and which can be stored and manipulated using a computer” (McEnery & Wilson, 1996: 73). This body of text can then be analysed in any number of ways in order to answer questions and provide evidence to support or refute hypotheses about language use. The benefit of using corpora in law and language research is that they provide “a means for the empirical analysis of language” (O’Keeffe & McCarthy, 2010: 7). After all, it is insufficient to base arguments about language use on what speakers intuitively believe, since their intuitions tend to be relatively untuned regarding many aspects and nuances of their language (Hunston, 2010).

The corpus used in this research, EUMigrLaw, was purpose-built using the online corpus software Sketch Engine (sketchengine.eu) (Kilgarriff et al., 2014). It is a sentence-aligned parallel corpus in English and Italian, allowing the user to compare how individual sentences are expressed in both languages. It is made up of a total of around 1.5 million tokens, where tokens are defined as the smallest individual units in a corpus, generally word forms, digits, punctuation and anything else between spaces. It contains a representative sample of legal texts on the topic of migration produced by the European Union institutions in both languages during two different time periods: 1992–1998 and 2015–2018. The reason for these time periods is explored in further

depth below. The distribution of tokens across the subsections of the corpus is set out in the table below.

Table 1: Distribution of tokens in EUMigrLaw corpus

	EUMigrLaw92-98	EUMigrLaw15-18
English	356,481	387,006
Italian	379,328	422,900

It is made up of different types of both binding and non-binding legal documents which are all freely available for consultation and searchable by topic online via the EUR-Lex service (eur-lex.europa.eu) in a range of formats and languages. Each text was checked manually before inclusion to ensure that it was appropriate for the purposes of this research. The resulting corpus is made up of five main types of legal acts, of which regulations, directives and decisions are binding, while opinions and recommendations are not. Each of these legal acts are included in the corpus to ensure a representative sample of the texts produced by the EU institutions.

As mentioned above, the texts included in the corpus cover two time periods: 1992–1998 and 2015–2018. These subcorpora are labelled EUMigrLaw92-98 and EUMigrLaw15-18 respectively. These timeframes were chosen as they both cover periods of particularly high activity with regard to legislation on the topic of migration, and terminology in EU texts has been shown to evolve significantly over even very short time spans (Peruzzo, 2013). The period 1992–1998 covers the signing of the Maastricht treaty and the Amsterdam treaty, which marks an extremely significant period in the development of migration law in the EU. Moreover, the period 2015–2018 was characterised by high levels of inwards migration into EU Member States with hundreds of thousands of people either crossing the Mediterranean or travelling overland through south-east Europe. It was therefore a period during which the EU published a large volume of binding and non-binding documents in an attempt to manage the migration situation.

3.2. Key Terms and Conceptual Categories

For the purposes of this pilot study, ten key concepts in the field of EU migration law were selected based on a manual consultation of the content of the legal texts contained in the corpus. Following this manual analysis, the concepts chosen were identified as being among the most frequently referenced across both time periods and therefore good candidates for terminological investigation in this context. It is important to note that the choice of conceptual categories is based on an examination of the language and terminology actually used in the texts in the corpus, rather than being determined in advance. The labels used for these concepts were taken from the

terms advocated by the glossary of the European Migration Network. The European Migration Network is an EU network of migration and asylum experts working to provide “objective, comparable policy-relevant information and knowledge on emerging issues relating to asylum and migration in Europe” (European Commission, 2020). The ten key concepts used in this study are listed below. For the sake of clarity, throughout this article, concepts are written in small caps and terms in italics.

ASYLUM SEEKER
DISPLACED PERSON
IRREGULAR MIGRANT
LONG-TERM RESIDENT
MIGRANT WORKER
REFUGEE
THIRD-COUNTRY NATIONAL
UNACCOMPANIED MINOR
UNION CITIZEN
VICTIM OF TRAFFICKING HUMAN BEINGS

Having established the ten key concepts, the keyword function available in Sketch Engine was used in order to extract all relevant terms in the corpus. The keyword function compares the corpus being studied against a pre-loaded reference corpus, English Web 2013. This comparison identifies those single-word and multi-word terms which appear at an unusually high frequency in EUMigrLaw and thus make that particular corpus unique. The keyword list was analysed manually to extract and categorise all relevant single-word and multi-word terms into one of the ten conceptual categories set out in the above concept list and to eliminate any terms not applicable to this particular pilot study, namely general procedural EU terminology (e.g. qualified majority) and migration terminology belonging to a conceptual category other than those chosen here (e.g. *returnee*). For example, the terms *illegal immigrants*, *irregular migrants* and *irregular immigrants* all belong to the same conceptual category and were therefore grouped under the label IRREGULAR MIGRANTS. In addition, to ensure all relevant terms have been captured, the n-gram function was used to detect any other multi-word terms not identified by the keyword search. This stage produced two lists of English terms, one for each time period, categorised according to the above concepts. The result is that each conceptual category has its own English lexical profile consisting of the different terms used to express that concept¹. Table 2 below shows an example of the conceptual category MIGRANT WORKER and the English terms appearing in EUMigrLaw92-98 which fit into this category.

¹ This stage of the process is similar to the creation of “formal onomasiological profiles” described by Speelman et al. (2003: 319), which involves the clustering of sets of alternative linguistic means to designate a single concept.

Table 2: Terms in conceptual category MIGRANT WORKER for EUMigrLaw92-98 subcorpus

Concept	Terms
MIGRANT WORKER	<i>Migrant worker</i>
	<i>Immigrant worker</i>
	<i>Foreign worker</i>

The next stage involves comparing English and Italian terms using the parallel concordance function. This function displays the Italian version of the English sentence containing the term in question, enabling the researcher to compare how each English term is expressed in Italian and log the frequency with which that specific term pair occurs. The data produced in this stage allows for an analysis of the variation present within both English and Italian in each time period. The extent of this intra-lingual variation is indicated by the number of different terms contained in each conceptual category. The greater the number of terms used in each language within a given conceptual category, the higher the level of intra-lingual terminological variation.

3.3. Terminological Distance

Since the table of results described above shows how each English term variant in each conceptual category is expressed in Italian, and the frequency with which given terms are used as equivalents, it can therefore also be used to examine inter-lingual variation (see table 3 below). This stage of the process involves determining the terminological distance between each of the term pairs used to express the various concepts in order to analyse inter-lingual terminological variation between English and Italian. Terminological distance is a systematic way of quantifying the inter-lingual variation between terms for the same concept in different languages based in their linguistic features. Taking inspiration from the method used to determine cognitive distance in the paper by Fernández-Silva & Kerremans (2011), a numerical value must be manually assigned to each pair of terms based on the differences in conceptual information between the two terms in order to determine terminological distance. This difference may be semantic, morphosyntactic or connotative in nature. In this pilot study, a terminological distance value was allocated to each English-Italian term pair by the author based on the criteria set out below. These values are allocated as follows:

A value of 0 is allocated where there is no discernible difference between the two terms on either a semantic, morphosyntactic and connotative level. An example of this is *migrant worker* and *lavoratore migrante* ('migrant worker'). A value of 0.5 is allocated if there is some level of terminological divergence of either a semantic, morphosyntactic or connotative nature. An example of a term pair with a semantic difference is *refugee* and

*profugo*². An instance of a term pair displaying morphosyntactic difference is *irregular immigrant* and *immigrato in situazione irregolare* ('immigrant in an irregular situation'), where the English term is [noun + adjective] while the Italian term is [noun + prepositional phrase]. An example of a term pair with connotative difference is *non-national* and *straniero* ('foreigner'), where the Italian term *straniero* often carries negative connotations compared to the neutral term *non-national*. A value of 1 is allocated to a pair where there is terminological divergence on semantic, morphosyntactic and connotative levels between the terms. For instance, this applies to the term pair *illegal immigrant* and *persona in una situazione irregolare* ('person in an irregular situation'). In this example, there are semantic, morphosyntactic and connotative differences between the two terms.

Once allocated, the terminological distance values were then weighted by multiplying them by the number of times a term pair appears in the corpus in order to ensure that the terminological distance values are representative. This then produces a weighted terminological distance for each pair. The sum of these weighted terminological distances was then divided by the total number of occurrences of terms in that conceptual category in order to produce the inter-lingual variation index (IVI) (Fernández-Silva & Kerremans, 2011), which is a mean value between 0 and 1 indicating the overall terminological difference between the two languages for that specific conceptual category. Table 3 below shows an example of the calculation of the interlingual variation index for the concept IRREGULAR MIGRANT in the EUMigrLaw15-18 subcorpus.

Table 3: Calculation of IVI for IRREGULAR MIGRANT in the EUMigrLaw15-18 subcorpus

Concept	ENG	IT	Absolute Frequency	Terminological distance	Weighted term. distance
IRREGULAR MIGRANT	<i>Irregular migrant</i>	<i>Migrante irregolare</i> ('Irregular migrant') ³	40	0	0
	<i>Irregular migrant</i>	<i>Migrante in posizione irregolare</i> ('Migrant in an irregular position')	1	0.5	0.5
	<i>Illegal immigrant</i>	<i>Clandestino</i> ('Clandestine (migrant)')	2	1	2
	<i>Unauthorised migrant</i>	<i>Migrante clandestino</i> ('Clandestine migrant')	2	0.5	1
	<i>Person residing without authorisation</i>	<i>Persona in soggiorno irregolare</i> ('Person on an irregular stay')	2	0.5	1
				Interlingual variation index =	4.5/47 = 0.09574

² Although commonly translated into English as *refugee*, the Italian term *profugo* applies to a person who has been forced to leave their home country but is not in a situation to request international protection, and is not therefore in strictly legal terms synonymous with the term *rifugiato* ('refugee').

³ Note that differences in adjective positioning between English and Italian are not considered a significant syntactic divergence for the purposes of terminological distance value allocation.

The sum of the various IVI values for each conceptual category can then be divided by the total number of conceptual categories (10 in the case of this pilot study) in order to produce a mean IVI for that language pair for that time period. Using this method, the interlingual variation index for the English-Italian language pair in both the earlier and later parts of the corpus can be calculated. The lower the IVI value, the more terminologically similar the two languages are for the concepts being studied. Comparing the IVI values for the earlier and later time periods can therefore offer an indication on whether the terminology used in these two languages has grown more similar or more different over time.

4. Results and Analysis

4.1. Intra-lingual Variation

The total number of terms identified for each language in each time period is shown in table 4 below, along with the mean number of terms per concept.

Table 4: Breakdown of number of terms by subcorpora

	Number of terms in EUMigrLaw92-98	Number of terms in EUMigrLaw15-18
English	29 Mean: 2.9 terms per concept	27 Mean: 2.7 terms per concept
Italian	29 Mean: 2.9 terms per concept	20 Mean: 2.0 terms per concept

Before analysing how terms for different concepts are rendered across both languages, it is interesting to observe the internal dynamic within each language with regard to terminological variation. Indeed, it is clear from the results shown in table 4 that there is widespread terminological variation within both languages over both time periods. In the earlier time period, a mean of 2.9 terms are used per concept in both English and Italian. While some of the concepts are consistently expressed with only a single term in this period, for instance REFUGEE in English (by the term *refugee*) and LONG-TERM RESIDENT in Italian (by the term *(persona) stabilitasi a titolo duraturo* ('person settled for a long period')), most of the concepts are expressed with several different terms in both languages. Indeed, there are four concepts which are expressed by three or more different terms in both languages, namely: IRREGULAR MIGRANT, MIGRANT WORKER, THIRD-COUNTRY NATIONAL and UNION CITIZEN. To illustrate this, the various terms used for the concept UNION CITIZEN in English and Italian in the EUMigrLaw92-98 subcorpus are shown in tables 5 and 6 below.

Table 5: Terms for UNION CITIZEN in English EUMigrLaw92-98 subcorpus

English	Relative frequency (per million tokens)
<i>Citizen of the Union</i>	84.16
<i>EU citizen</i>	47.69
<i>European citizen</i>	33.66
<i>National of Member States</i>	19.64
<i>Union citizen</i>	16.83

Table 6: Terms for UNION CITIZEN in Italian EUMigrLaw92-98 subcorpus

Italian	Relative frequency (per million tokens)
<i>Cittadino dell'Unione</i> ['Citizen of the Union']	108.09
<i>Cittadino degli Stati Membri</i> ['Citizen of the Member States']	39.54
<i>Cittadino dell'Unione europea</i> ['Citizen of the European Union']	31.63
<i>Cittadino comunitario</i> ['Community citizen']	26.36
<i>Cittadino europeo</i> ['European citizen']	18.45

The above example is illustrative of the level of variation and instability in terminology within both languages for many of the concepts in the earlier subcorpus. The extent of this intra-lingual variation is striking given the importance of consistency in legal texts, where one might expect a single term to be used each time the same concept is expressed for the sake of both legal interpretation and public accessibility. Indeed, although some of the variation is attributable to mere orthographical variants, such as *asylum seeker* and *asylum-seeker*, most of the differences observed are due to denominative variation.

The figures in table 4 show that the extent of this intra-lingual terminological variation reduces somewhat over the time period. This reduction in variation is only modest in English, remaining at a mean of 2.7 terms per concept, while the decrease in variation is more significant in Italian, falling to a mean of 2.0 terms per concept. Nevertheless, despite this difference between the languages, these results still indicate that more than one term is frequently used for the same concept in both languages. Indeed, although fewer terms are used on average to express concepts in the later subcorpus, most of the concepts are still expressed by at least two different term variants in both languages. The more significant decrease in variation in Italian is due largely to the fact that there are more conceptual categories with only a single term in that language compared to English. In total, for English, only 3 out of 10 concepts have a single term variant in EUMigrLaw92-98, which actually falls to 2 out of 10 in EUMigrLaw15-18. For

Italian, however, while only 1 out of 10 concepts in EUMigrLaw92-98 is expressed with a single term, this rises to 5 out of 10 for EUMigrLaw15-18.

This difference can be illustrated by examining some specific concepts. For instance, the category ASYLUM SEEKER is expressed unequivocally in Italian in EUMigrLaw15-18 by the term *richiedente asilo* ('asylum-seeker'), whereas this same category in English still contains three variants: *asylum-seeker*, *asylum seeker* and *asylum applicant*. Likewise, the concept UNACCOMPANIED MINOR is expressed in Italian in this subcorpus on every occasion by the term *minore non accompagnato* ('unaccompanied minor') while the English concept is expressed with three variants: *unaccompanied minor*, *unaccompanied child* and *UAM*. Incidentally, the use of the term *child* here has also been deemed "highly problematic" due to the "more charged" connotations of the term *child* as opposed to *minor* within the broader migration context (Bennett, 2021: 367). Indeed, not only can such terminological variation create unwanted confusion but it can also introduce or perpetuate terms with negative connotations. Since this pilot study only examines 10 key concepts, a more comprehensive study using this methodology must be undertaken to establish whether this trend of greater terminological variation in English than Italian applies more generally in EU legal texts.

The results in this section demonstrate that there is a great deal of variation and instability within both languages for a significant number of the conceptual categories identified over both time periods. Although there is a reduction in the mean number of terms per concept for each language, the mean for the later subcorpus remains significantly higher than one term per concept, indicating continued variability in expressing concepts in both languages.

4.2. Inter-lingual Terminological Variation

Having observed the extent of terminological variation within the individual languages, the extent of inter-lingual terminological variation can also be analysed. This indicates how similar the terminology used in each language is for each concept and whether this has changed over the time period under investigation. The weighting of the terminological distance values, as discussed in section 3.3, is essential since the absolute frequency of some term pairs is extremely high, over 300 in some cases (e.g. third-country national – *cittadino di un paese terzo* ('citizen of a third country')), while other term pairs only appear on one or two occasions (e.g. unaccompanied child – *minore non accompagnato* ('unaccompanied minor')). The IVI values for each conceptual category are presented in table 7 below. These resulting IVI values were then added together and divided by the overall number of categories (10) to provide the mean IVI for the English-Italian language pair in each time period. The results of these calculations are also shown in table 7 below:

Table 7: English-Italian IVI values

Concept	IVI EUMigrLaw92-98	IVI EUMigrLaw15-18
ASYLUM SEEKER	0.02500	0.04795
DISPLACED PERSON	0.46667	0.47727
IRREGULAR MIGRANT	0.79487	0.13265
LONG-TERM RESIDENT	0.50000	0.16667
MIGRANT WORKER	0.10000	0.00000
REFUGEE	0.20800	0.03709
THIRD-COUNTRY NATIONAL	0.08696	0.00871
UNACCOMPANIED MINOR	0.25000	0.02128
UNION CITIZEN	0.18824	0.06863
VICTIM OF TRAFFICKING IN HUMAN BEINGS	0.69048	0.12000
	Mean IVI EUMigrLaw92-98	Mean IVI EUMigrLaw15-18
	0.33102	0.10803

Table 7 shows a significant fall of 67.36% in the mean IVI over the time period under investigation, suggesting substantial terminological convergence between the two languages being studied here, at least with regard to the migration-related concepts under investigation in this pilot study. This conclusion can be reached from the quantitative data above because the IVI value is an indication of the mean terminological distance between two languages with regard to the set of concepts under investigation. A decrease in the IVI therefore indicates a drop in the mean terminological distance observed for the collection of migration-related terms being examined. In other words, there has been a marked increase in the semantic, morphosyntactic and/or connotative similarities between the terms being used in English and Italian.

This conclusion can be elaborated through some specific examples. Both lexical and morphosyntactic convergence can be identified in the terminology used for the concept IRREGULAR MIGRANT. The most frequently occurring Italian term in the earlier subcorpus for the concept IRREGULAR MIGRANT is *immigrato in situazione irregolare* ('immigrant in an irregular situation') (46% of terms), while in English it is *illegal immigrant* (76% of terms). In the later time period, the most frequent English term changes to *irregular migrant* (79% of terms). Indeed, it has long been argued that only actions can be illegal, not people themselves. In an attempt to combat this usage, the European Parliament's resolution of 14 January 2009 called for EU institutions and Member States to stop using the term *illegal immigrants* due to its very negative connotations, instead preferring *irregular/undocumented workers/migrants*. This research shows a very clear preference in more recent EU legal texts in English for the term *irregular migrant*. Therefore, the predominant adjective used in English for this concept shifts to become more similar to the adjective most frequently used in Italian in both time periods: *irregolare*. At the

same time, morphosyntactic change in the Italian terminology increases the convergence identified. While the structure of the English terms remains stable over time [adjective + noun] (98% of terms), an analysis of the results for this concept in the later subcorpus shows that the Italian terminological structure has shifted from the [noun + prepositional phrase] (56% of terms) structure in *immigrato in situazione irregolare* to the [adjective + noun] structure of *migrante irregolare* ('irregular migrant') (86% of terms) thus reflecting the English morphosyntax. The adjective *irregolare* now directly modifies the noun *migrante*. This apparently minor morphosyntactic change has the effect of shifting the focus from the situation in which a person happens to find herself to the irregularity of the person herself. In sum, lexical and morphosyntactic changes occurring in both English and Italian terminology for this concept over the time period has resulted in significantly reduced terminological distance between the two languages, thus corroborating the quantitative findings set out above.

Another example of terminological convergence between the two languages can be seen with the concept LONG-TERM RESIDENT, which refers to anyone with long-term resident status under the Long-Term Residents Directive (Council Directive 2003/109/EC, 2003). In both the earlier and later time periods, the only term used in English for this concept is *long-term resident*. However, the terms used in Italian for this concept differ between the two subcorpora. The preferred Italian term for this concept in the earlier subcorpus is *(persona) stabilitasi a titolo duraturo* ('person settled on a long-term basis'). In the later subcorpus, this terminology is dropped entirely, shifting to *soggiornante di lungo periodo* ('long-term resident'), which closely reflects the terminology used in English. The Italian terminology for this concept therefore moves closer to the English both semantically and morphosyntactically.

The above examples are illustrative of the terminological changes detected in both English and Italian, substantiating the quantitative findings which indicate significant terminological convergence between the languages. In any case, it is clear from the quantitative data produced here that the terminology used in the English and Italian versions of EU legal texts to express specific migration-related concepts has become significantly more similar over time. In addition to highlighting intra-lingual variation, this methodology therefore also provides robust, empirical evidence of the significant inter-lingual convergence that has occurred for the set of migration-related terms examined.

5. Discussion

The findings produced by this study are especially striking given the various systems in place to ensure consistency of terminology in the EU and minimise terminological instability and variation. The Terminology Coordination unit, known as TermCoord, as-

sists translation units in “terminology management” and provides “terminology and documentation tools and resources” (European Parliament, 2018), while the IATE database, in use since 2004, provides translators with EU-specific terminology with the aim of “enhancing the availability and standardisation” of information regarding terminology in the institution’s official languages (Translation Centre for the Bodies of the European Union, 2021). Specifically in relation to migration, the European Migration Network provides an up-to-date glossary of around 500 terms reflecting the most recent European policy on migration and asylum (European Commission, 2020). Moreover, the Joint Practical Guide produced by the EU encourages drafters to avoid “the use of synonyms and different expressions to convey the same idea” in order to comply with the key principle that legal acts should be “clear”, “simple” and “precise” (European Commission, 2016: 10–11). Despite these efforts to ensure consistency and standardisation of terminology within and between different languages, this research shows that terminological variation and instability remains widespread, suggesting other factors are at play in creating and perpetuating this dynamic of variation.

Although an in-depth investigation into the causes of such variation is beyond the scope of this paper, it is likely to result from a combination of causes highlighted by previous research. Such variation has been attributed to translators as “semi-specialists” having inadequate knowledge of the target language combined with problems with resources that have excessively large translation memories and inconsistencies between terminological databases (Biel, 2014: 276). It is also important to note the different translation processes used by the EU for different types of legal documents. While binding legislation is subject to multi-stage and multi-lingual scrutiny and quality assurance carried out mainly in-house, non-binding texts are often outsourced to external contractors, and may not therefore be subject to the same strict controls (Biel & Koźbiał, 2020: 75–76). In addition, the chronological variation also detected in this corpus study can result from a development in knowledge about a given concept, which leads to “a period of cohabitation between an old and a new term” (Freixa, 2006: 55). This is especially apparent for relatively newer domains which are still developing, leading to a sort of terminological “chaos” prior to a clearer and more fixed terminology becoming established (Bertaccini et al., 2010: 11–12).

Furthermore, specifically regarding the terminological convergence identified by this study, it is possible that the dominant usage of English within the EU institutions has impacted the terminology used in Italian (and other languages besides). In the past, the majority of the original drafts of legal texts were written in French. Indeed, in 1994, a survey carried out among EU officials showed that 59% were using French as an original drafting language, while 33% were using English (Mattila, 2006: 33). However, the ineluctable rise of English in the institutions meant that by 2013, 81% of all legislation was being originally drafted in English (Čavoški, 2017: 62). During this drafting process, the Legal Service of the Commission supports internal analysis of scientific evidence and working groups with Member States and subject experts at technical and

political levels (Čavoški, 2020). For the sake of convenience and expediency, this is usually done on the basis of one language version – usually English – and the other language versions are then adapted to fit with the base language version being worked on (Robertson, 2010). Consequently, in the context of this multi-layered legislative drafting process, from the genesis of the legislation and through each of the drafting stages, there is a *de facto* base version – usually English – which is consulted by the working groups and experts negotiating the formation of the legislation. This dominant use of English is likely to be at least in part responsible for the terminological convergence observed in this pilot study with Italian terminology being influenced by the lexical and morphosyntactic features of the English terminology.

Specifically regarding the topic of term creation, terms are created in the EU system in a two-step process. They tend to be initially established in one of the procedural languages (usually English) and are then translated into all the other languages (Fischer, 2010: 28). In addition, it is common within the EU to transpose an original term by creating a calque of that term in another language in order to preserve its formal and semantic characteristics (Mariani, 2021: 38). As a result, the choice of the terminology in these other languages will inevitably be influenced to a large extent by the initially created term. The terminological convergence detected by the methodology set out in this study can therefore be at least partially attributed to this term creation process.

6. Conclusion

This paper sets out an original methodology for exploring terminological variation in EU legal texts both within and between different languages. The methodology provides precise quantitative data on the extent of variation detected and the degree of convergence or divergence between two languages for a given terminological field. The pilot study set out subsequently aims to demonstrate the application of this methodology and the nature of the results produced. The findings of the corpus analysis undertaken in this study provide strong evidence to support arguments regarding terminological variation.

Firstly, the significant extent of intralingual variation within English and Italian is highlighted by the average number of terms used in both languages to express the key concepts identified. Although the mean number of terms per concept decreases over time, it remains high. Interestingly, for the concepts studied here, terminological variation falls much more sharply in Italian than in English, which only shows a slight decrease. Further research is required to establish how widespread this trend is. In any case, the legal implications of such variation are wide-reaching. After all, legal terms represent “units of legal knowledge” and therefore serve as “prompts and points of access to knowledge structures” (Biel, 2014: 41). If these terms are used inconsistently ei-

ther synchronically or diachronically within a language, these points of access to knowledge become blurred and confused, potentially leading to ambiguity and uncertainty among those attempting to interpret the law.

Secondly, the findings also clearly indicate significant terminological convergence between the languages, indicated by the fall in the IVI value between the two time periods. The 67.36% decrease in IVI values for the English-Italian language pair with regard to the terminology for the ten key migration concepts identified shows the extent of the terminological convergence that has taken place. Examples such as IRREGULAR MIGRANT and LONG-TERM RESIDENT illustrate the manner in which terminology in English and Italian has changed over the time period in question to become more similar, both semantically and morphosyntactically. The methodology described in this paper can therefore identify the existence and extent of terminological convergence or divergence, and further qualitative steps can be conducted to understand the precise nature of such changes.

In any case, the specific causes of the intralingual and interlingual variation identified by this pilot study require further in-depth investigation to provide evidence to support some or all of these potential causal factors. The methodology described in this paper can be replicated with larger sets of concepts on any number of topics using either purpose-built or ready-to-use EU corpora in order to examine terminological variation in that particular field. This approach can also be used with any combination of language pairs from among the official languages of the EU and the time period studied to track chronological variation can be adapted to suit the purposes of the specific research being carried out. The empirical research method set out in this paper can also be further strengthened by an additional qualitative analysis stage, examining individual concepts and terms in order to gain yet more insight into how terminology within each conceptual category has evolved. From a practical forward-looking perspective, this study provides further impetus to the recommendation by Otero Fernández (2020: 218) to ensure closer contacts between the various actors involved in the legislative procedure and continue to improve and streamline terminological databases and glossaries available to legal translators. The methodology set out in this paper could also be put to use to identify areas of EU law with demonstrably high levels of terminological variation and provide an indication to the EU's terminology services of the specific domains where standardisation work is necessary to reduce levels of synonymy and variation in the future.

References

European Union (1992). Treaty on European Union (Consolidated Version), Treaty of Maastricht , 7 February 1992, *Official Journal of the European Communities* C 325/5. Available at: eur-lex.europa.eu/resour

- [ce.html?uri=cellar:2bf14obf-a3f8-4ab2-b506-fd71826e6da6.0023.02/DOC_1&format=PDF](https://doi.org/10.1075/btl.157.16ben) (accessed 22 December 2022)
- Ainsworth, Janet (2014). Lost in Translation? Linguistic Diversity and the Elusive Quest for Plain Meaning in Law. In Cheng, Sin & Wagner (Eds.), *The Ashgate Handbook of Legal Translation* (pp. 43–56). London: Routledge.
- Bennett, Karen, (2021). Agency and social responsibility in the translation of the migration crisis. In Carbonell i Cortés & Monzó-Nebot (Eds.), *Translating Asymmetry – Rewriting Power* (pp. 361–377). Amsterdam/Philadelphia: John Benjamins. DOI: [10.1075/btl.157.16ben](https://doi.org/10.1075/btl.157.16ben).
- Berenskoetter, Felix (2016). *Concepts in World Politics*. London: Sage.
- Bertaccini, Franco, Massari, Monica & Castagnoli, Sara (2010). Synonymy and variation in the domain of digital terrestrial television: Is Italian at risk? In Thelen & Steurs (Eds.), *Terminology in Everyday Life* (pp. 11–20). Amsterdam: John Benjamins.
- Bessé, Bruno de, Nkwenti-Azeh, Blaise & Sager, Juan (1997). Glossary of terms used in terminology. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 4(1), 117–156.
- Biel, Łucja (2014). *Lost in the Eurofog: The Textual Fit of Translated Law*. Berlin: Peter Lang.
- Biel, Łucja, Biernacka, Agnieszka & Jopek-Bosiacka, Anna (2018). Collocations of Terms in EU Competition Law: A Corpus Analysis of EU English Collocations. In Marino, Biel, Bajčić & Sosoni (Eds.), *Language and Law* (pp. 249–274). Cham: Springer.
- Biel, Łucja & Koźbiał, Dariusz (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. *Estudios de Traducción*, 10, 69–90. DOI: [10.5209/estr.68054](https://doi.org/10.5209/estr.68054).
- Čavoški, Aleksandra (2017). Interaction of law and language in the EU: Challenges of translating in multilingual environment. *The Journal of Specialised Translation*, 27, 58–74. Available at: jost-trans.org/issue27/art_cavoski.pdf (accessed 22 December 2022).
- Čavoški, Aleksandra (2018). Legal Language and EU Integration. The Case of the Western Balkans. *International Journal of Language & Law (JLL)*, 7, 70–96. DOI: [10.14762/jll.2018.070](https://doi.org/10.14762/jll.2018.070).
- Čavoški, Aleksandra (2020). Science and Law in Environmental Law and Policy: The Case of the European Commission. *Transnational Environmental Law*, 9(2), 263–295. DOI: [10.1017/S2047102520000151](https://doi.org/10.1017/S2047102520000151).
- Crépeau, François (2013). UN Special Rapporteur on the Human Rights of Migrants: Concluding remarks. In:
- European Commission Legal Service (2016). *Joint practical guide of the European Parliament, the Council and the Commission for persons involved in the drafting of European Union legislation*. Available at: op.europa.eu/en/publication-detail/-/publication/3879747d-7a3c-411b-a3a0-55c14e2ba732/language-en (accessed 23 July 2022).
- European Commission (2020). *European Migration Network*. Available at: home-affairs.ec.europa.eu/networks/european-migration-network-emn_en (accessed 27 August 2022).
- European Council (2003). *Council Directive 2003/109/EC of 25 November 2003 concerning the status of third-country nationals who are long-term residents*. Official Journal L 16, 23.1.2004. Available at: eur-lex.europa.eu/legal-content/en/TXT/?uri=celex%3A32003L0109 (accessed 27 August 2022).
- European Parliament (2009). *Resolution of 14 January 2009 on the situation of fundamental rights in the European Union 2004-2008*. Available at: europarl.europa.eu/doceo/document/TA-6-2009-0019_EN.html (accessed 1 September 2022).
- European Parliament (2018). *Terminology Coordination*. Available at: termcoord.eu (accessed 1 September 2022).
- Felber, Helmut (1984). *Terminology Manual: UNESCO*. Paris: Unesco.
- Fernández-Silva, Sabela & Kerremans, Koen (2011). Terminological variation in source texts and translations: A pilot study. *Meta* 56(2), 318–335. DOI: [10.7202/1006179ar](https://doi.org/10.7202/1006179ar).

- Fischer, Márta (2010). Language (policy), translation and terminology in the European Union. In Thelen & Steurs, *Terminology in Everyday Life* (pp. 21–34). Amsterdam: John Benjamins. DOI: [10.1075/tlrp.13.03fis](https://doi.org/10.1075/tlrp.13.03fis).
- Freixa, Judit (2006). Denominative variation in terminology. *Gesture*, 8(3), 285–301. DOI: [10.1075/term.12.1.04fre](https://doi.org/10.1075/term.12.1.04fre).
- Freixa Aymerich, Judit, Fernández-Silva, Sabela & Cabré, M. Teresa (2008). La multiplicité des chemins dénominatifs. *Meta: Translators' Journal*, 53(4), 731–747. DOI: [10.7202/019644ar](https://doi.org/10.7202/019644ar).
- Hunston, Susan, (2010): How can a corpus be used to explore patterns? In O'Keefe & McCarthy, *The Routledge Handbook of Corpus Linguistics* (pp. 152–166). London: Routledge.
- Kageura, Kyo (2002). The Dynamics of Terminology. *A descriptive theory of term formation and terminological growth*. Amsterdam: John Benjamins.
- Kilgariff, Adam, Baisa, Vít, Bušta, Jan, Jakubíček, Miloš, Kovář, Vojtěch, Michelfeit, Jan, Rychlý, Pavel, Suchomel, Vít (2014). The Sketch Engine: ten years on. *Lexicography*, 1, 7–36.
- Loupaki, Elpida (2018). Synonymy in the domain of migration: Sign of changing mentalities? *Parallèles*, 30(2), 22–34.
- Mariani, Jessica (2018). *Migration in Translation: the role of terminology and trans-editing in shaping the crisis in EU Institutions*. University of Verona. Available at: iris.univr.it/handle/11562/981576 (accessed 22 December 2022).
- Mariani, Jessica (2021). Migration terminology in the EU Institutions. Overview and patterns of use of terms from 1950 to 2016. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 27(1), 35–55. DOI: [10.1075/term.00057.mar](https://doi.org/10.1075/term.00057.mar).
- Mattila, Heikki (2006): Comparative Legal Linguistics. In: 1903-1785. DOI: [10.4324/9781315573106](https://doi.org/10.4324/9781315573106).
- McEnery, Tony & Wilson, Andrew (1996). *Corpus Linguistics*. Edinburgh: University Press.
- O'Keefe, Anne & McCarthy, Michael (2010): *The Routledge Handbook of Corpus Linguistics*. 1. Ed. London: Routledge.
- Olsen, Niklas (2012). *History in the Plural: An Introduction to the Work of Reinhart Koselleck*. New York: Berghahn Books.
- Otero Fernández, Irene (2020): *Multilingualism and the Meaning of EU Law*. European University Institute. Available at: cadmus.eui.eu/bitstream/handle/1814/66308/Fernandez_2020_LAW.pdf?sequence=1&isAllowed=y (accessed 22 December 2022).
- Pearson, Jennifer (1998). *Terms in Context*. Amsterdam: John Benjamins.
- Peers, Steve (2011). *EU Justice and Home Affairs Law*. Oxford: University Press.
- Peruzzo, Katia (2013). Short-period evolution in EU legal texts: Old and new terms, old and new concepts. *Linguistica*, 53(2), 39–53. DOI: [10.4312/linguistica.53.2.39-53](https://doi.org/10.4312/linguistica.53.2.39-53).
- Picton, Aurelie (2014). The dynamics of terminology in short-term diachrony. In: Temmerman & Van Campenhoudt (Eds.), *Dynamics and Terminology: An interdisciplinary perspective on monolingual and multilingual culture-bound communication* (pp. 159–182). Amsterdam: John Benjamins. DOI: [10.1075/tlrp.16.09pic](https://doi.org/10.1075/tlrp.16.09pic).
- Prieto Ramos, Fernando (2019). The use of corpora in legal and institutional translation studies. *Translation Spaces*, 8(1), 1–11. [10.1075/ts.00010.pri](https://doi.org/10.1075/ts.00010.pri).
- Prieto Ramos, Fernando & Morales Moreno, Albert (2019). Terminological innovation and harmonization at international organizations: Can too many cooks spoil the broth? In Simonnæs & Kristiansen (Eds.), *Legal Translation. Current Issues and Challenges in Research, Methods and Applications* (pp. 87–110). Berlin: Frank & Timme. Available at archive-ouverte.unige.ch/unige:115687.
- Robertson, Colin (2010). EU law and semiotics. *International Journal for the Semiotics of Law*, 23(2), 145–164. DOI: [10.1007/s11196-010-9149-x](https://doi.org/10.1007/s11196-010-9149-x).
- Sigona, Nando (2018). The contested politics of naming in Europe's "refugee crisis". *Ethnic and Racial Studies* 41(3), 456–460. DOI: [10.1080/01419870.2018.1388423](https://doi.org/10.1080/01419870.2018.1388423).

- Speelman, Dirk, Grondelaers, Stefan & Geeraerts, Dirk (2003). Profile-based linguistic uniformity as a generic method for comparing language varieties. *Computers and the Humanities* 37(3), 317–337. DOI: [10.1023/A:1025019216574](https://doi.org/10.1023/A:1025019216574).
- Stefaniak, Karolina (2017). Terminology work in the European Commission: Ensuring high-quality translation in a multilingual environment. In Svoboda, Biel & Łoboda (Eds.), *Quality aspects in institutional translation* (pp. 109–121). Berlin: Language Science Press. DOI: [10.5281/zenodo.1048192](https://doi.org/10.5281/zenodo.1048192).
- Strandvik, Ingemar (2012). Legal Harmonization Through Legal Translation: Texts that Say the Same Thing? In Baaij (Ed.), *The Role of Legal Translation in Legal Harmonization* (pp.25–50). Alphen aan den Rijn: Kluwer Law International.
- Temmerman, Rita (2000). *Towards New Ways of Terminology Description*. Amsterdam: John Benjamins Publishing Company.
- Translation Centre for the Bodies of the European Union (2021): IATE: European Union Terminology. Available at: iate.europa.eu (accessed 22 December 2022).
- Trklja, Aleksandar & McAuliffe, Karen (2019). Formulaic metadiscursive signalling devices in judgments of the court of justice of the european union: A new corpus-based model for studying discourse relations of texts. *International Journal of Speech, Language and the Law*, 26(1), 21–55. DOI: [10.1558/ijssl.36920](https://doi.org/10.1558/ijssl.36920).
- Vigier, Francisco J. & del Mar Sánchez-Ramos, María (2017). Using Parallel Corpora to Study the Translation of Legal System-Bound Terms: The Case of Names of English and Spanish Courts. In Mitkov (Ed.), *Computational and Corpus-Based Phraseology* (pp. 260–273). Cham: Springer. DOI: [10.1007/978-3-319-69805-2_19](https://doi.org/10.1007/978-3-319-69805-2_19).
- Wüster, Eugen (1968): *The Machine Tool: An Interlingual Dictionary of Basic Concepts*. London: Technical Press.

Note: JLL and its contents are Open Access publications under the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).



Copyright remains with the authors. You are free to share and adapt for any purpose if you give appropriate credit, include a link to the license, and indicate if changes were made.

Publishing Open Access is free, supports a greater global exchange of knowledge and improves your visibility.