

How EUR-lex parallel corpora improve the translation of general legal formulae in LELF

Patrizia Giampieri PhD

University of Perugia, Italy

patrizia.giampieri@unipg.it

ORCID: <https://orcid.org/0000-0003-4443-1593>

Abstract: Legal formulae are fixed expressions that can pose several translation challenges in view of system-related peculiarities, especially in international and multicultural settings such as the European Union. General bilingual dictionaries are often used by students and translators to cope with legilinguistic hurdles, although such tools do not always account for system-bound specificities. This may give rise to translation infelicities and/or partial translations. This paper addresses legal translation research and examines the usefulness of the EUR-Lex parallel corpora via the Sketch Engine platform to translate general legal formulae. More precisely, it focuses on neutral legal multi-words and compares corpus-driven data against entries obtained from three online general dictionaries (i.e., Cambridge, Collins, and WordReference). It assesses the translation of seven general legal formulae from Legal English as a Lingua Franca (LELF) into French and Italian. By comparing corpus-sourced equivalents with standard dictionary outputs, the study reveals that institutional parallel corpora generally outperform generic lexicographical tools. The findings also indicate that traditional dictionaries suggest decontextualised terms, whereas specialised parallel corpora offer

contextually relevant terminology. Therefore, this investigation provides useful insights for the integration of corpus-driven data into professional legal translation practice and translator-training programmes.

Keywords: Legal English as a Lingua Franca (LELF); legal translation; corpus-based translation; EUR-Lex; European legal English; legal formulae

Corpora paralleli EUR-Lex nella traduzione di formule giuridiche neutre

Sommario: Le formule giuridiche sono espressioni linguistiche difficili da tradurre a causa delle peculiarità dei singoli sistemi giuridici, soprattutto in contesti internazionali e multiculturali come l'Unione europea. I dizionari bilingui generici, spesso utilizzati da studenti e traduttori, non tengono spesso conto delle specificità legilinguistiche dei termini giuridici. Il presente contributo si colloca nell'ambito della ricerca sulla traduzione giuridica ed esamina l'utilità dei corpora paralleli EUR-Lex disponibili nella piattaforma Sketch Engine. Il presente lavoro esamina quindi alcune formule giuridiche neutre e ne propone un'analisi corpus-based, raffrontando dati estrapolati dai corpora paralleli con le voci di tre dizionari generici online (Cambridge, Collins e WordReference). In particolare, lo studio valuta la traduzione di sette formule giuridiche dall'inglese come lingua franca (Legal English as a Lingua Franca, LELF) al francese e all'italiano. Confrontando i traducenti ottenuti mediante l'analisi dei corpora con le tradizionali proposte lessicografiche dei dizionari, la ricerca dimostra che i corpora paralleli istituzionali offrono migliori soluzioni linguistiche. I risultati evidenziano che i dizionari tradizionali suggeriscono termini decontestualizzati, mentre i corpora paralleli specializzati forniscono terminologia contestualmente rilevante. Pertanto, l'indagine mostra come l'uso strumenti di ricerca corpus-based possano essere impiegati nella pratica professionale e nella formazione dei traduttori.

Parole chiave: inglese giuridico come lingua franca; traduzione giuridica; traduzione mediante corpora; EUR-Lex; inglese giuridico europeo; espressioni formulaiche giuridiche

1. Introduction

1.1 Legal language, legal systems and legal formulae

Legal language is characterised by a peculiar drafting style (Tiersma, 1999; Bhatia et al., 2004; Williams, 2004; Bhatia, 2010; Tiersma & Solan, 2012). It is hallmarked by stylistic features that make it particularly difficult and intricate (Tiersma, 1999; Williams, 2004). Amongst others, in fact, legal texts feature nominalisation (Tiersma, 1999; Williams, 2004; Coulthard & Johnson, 2010), which is a linguistic phenomenon according to which verb and noun combinations are preferred over verbs (as in the phrase “an amendment to the Constitution may be made” instead of “the Constitution may be amended”) (Williams, 2004, p. 115). Legal texts also contain Latinisms (as the expressions *ab extra* and *ab initio*) (Bhatia, 2010); syntactic discontinuities (which occur when long sentences exhibit subordinates and embeddings) (Williams, 2004); deixis (that consists of cataphoric or anaphoric particles or adverbs such as “theretofore” or “whereof” that refer back or forward to elements or concepts mentioned in a text) (Abate, 1998, pp. 14-16); passive constructions (i.e., agentless phrases used to create ambiguity and/or conceal concepts), and vagueness (which is often necessary to ensure flexibility, but may seem inconsistent with the purpose of certainty conveyed by legal texts) (Williams, 2005). In this respect, in fact, Tiersma (1999) contends that “the tension between precision and flexibility is an inherent problem for legal language” (Tiersma, 1999, p. 85).

Legal discourse is also characterised by formulaic expressions which are fixed prefabricated language patterns, or “multi-word units” (O’Keeffe et al., 2007, p. 63). The historical reasons why formulae exist in legal discourse is due to the fact that justice-related words were invested with talismanic powers in the Middle Ages, and people believed in magic and divine intervention (Tiersma, 1999). Therefore, legal transactions or proceedings involved the performance of rituals, during which formulaic phrases were employed (Tiersma, 1999).

Nowadays, legal multi-words may be preferred over plainer expressions because they have been universally accepted for centuries (Gotti, 2012, p. 52). Hence, their mention in a text supports and ensures the validity of the actions prescribed (ibid.).

One of the relevant issues concerning legal formulae is their nature and scope, as they may vary on the basis of the purpose and text where they are employed (Giampieri, 2025). A power of attorney, for instance, may contain the following opening statement: “Know all men by these presents” (Garner, 1995, p. 366), whereas a contract could mention “Now, therefore, the parties hereto have agreed as follows” (Giampieri, 2021, p. 41, 76, 79). A British English claim form probably reads “and the Claimant claims” (Civil Justice Council, 2013, p. 11), whereas an Australian one would mention “and the plaintiff claims” (Giampieri, 2025, p. 15).

Some legal formulae and lexical phrases could appear in several legal documents, as they are not genre-dependent. Examples are expressions such as “having regard to (the opinion of)”; “by virtue of”, or “under no circumstances” (Giampieri, 2025, p. 82, 113, 166). For the purposes of this paper, such formulaic multi-words are referred to as “general legal formulae” (ibid.). Some scholars also define them as neutral, or non-technical (Šarčević, 1997, p. 255).

1.2 Translating legal formulae

For the reasons stated above, legal fixed expressions are challenging, not only to the layperson (Williams, 2004), but also to the translator (Garzone, 2007). It is argued, in fact, that students in translation studies are not usually trained in the language of the law and they gain experience in the field through practice and lifelong education (Prieto Ramos, 2020). Even when trained, translators must become acquainted with both source and target language conventions, as well as prefabricated phrases (Garzone, 2007). Also, they need to make comparisons between the source and target legal systems (Šarčević, 1997), taking into account that national legal traditions evolve over time and are influenced by other legal systems (Stein, 2009),

Therefore, translators are confronted not only with the difficulties of two languages, but also with two different and dynamic legal systems (Engberg, 2020).

At European level, system-specificity challenges are apparently not present, as there is no legal tradition in which European law is rooted (Jacometti & Pozzo, 2018). As a matter of fact, the civil law and common law jurisdictions converge and merge into European law (Orts Llopis, 2009; Anesa, 2019). For this reason, the legal English adopted at EU level is a *lingua franca* (Jacometti & Pozzo, 2018), also referred to as LELF (Legal English as a *Lingua Franca*) (Campos-Pardillos, 2010; Anesa, 2019). Its *raison d'être* is driven by the need to foster comprehensibility and integrate practices and procedures across the EU (Anesa, 2019). LELF is obviously employed in multinational institutional contexts, such as international courts and bodies (Anesa, 2019). It would obviously be inappropriate and misleading to use LELF in settings where the legal terminology is bound to a specific national legal system and centuries-old legal and linguistic traditions (Giampieri, 2025).

1.3 Legal language and corpora

Corpora are collections of texts available in electronic format (Bowker & Pearson, 2002). They are argued to be accurate language resources to tackle sector-specific language (Aston, 2001), especially if they are DIY (do-it-yourself) (Zanettin, 2002), or “home-made” (Priya et al, 2018), i.e., compiled by the user. Corpora, in fact, show patterns of language usage (Farr & O’Keeffe, 2019), and are particularly useful in specialised translation. For these reasons, it is argued that they are resourceful to both analyse and translate legal language (Fantinuoli & Zanettin, 2015; Biel, 2017; Vigier Moreno & Sánchez Ramos, 2017). For example, they can be consulted to translate and reflect on past translation choices (Biel, 2017). Additionally, parallel corpora (such as the ones on the judgments of the European Union Court of Justice) can be applied in the translation training classroom (Vigier Moreno & Sánchez Ramos, 2017).

1.3.1 Parallel corpora

There are several types of corpora users can resort to. Parallel corpora consist of original texts and their translations into one or more languages (Zanettin, 2014). They feature aligned texts which are comparable between them in terms of content and terminology. For these reasons, they are considered useful because they help notice past translation choices (Zanettin, 2014) and find acceptable translation options. Many parallel corpora have been built by and within the European Institutions (Fantinuoli & Zanettin, 2015). In these cases, it is difficult to distinguish between a source text and the corresponding target texts, as the parallel corpora of EU Institutions are composed of original texts, at least as far as the legal aspect is concerned (Fantinuoli & Zanettin, 2015).

An example of parallel corpora featuring European legal documents are the EUR-Lex datasets uploaded to the Sketch Engine platform (Kilgarriff et al., 2004). These corpora are available in 24 EU languages, and are built from the html files of the EUR-Lex online database. The EUR-Lex platform contains EU legal documents, such as legal acts (directives, regulations, decisions, and recommendations), the case-law of the Court of Justice of the European Union, and other legal texts issued by EU bodies (such as reports and communications).

2. Aim of the paper and research questions

Given the high recurrence of formulaic expressions in the language of the law and the difficulties in tackling them due to both semantic and lexicological traditions, as well as system-specificity reasons, this paper is aimed at exploring whether parallel corpora within EU Institutions such as the EUR-Lex corpora available in the Sketch Engine platform can be helpful in contexts where LELF is used. In addition, online bilingual dictionaries are consulted with a view to verifying whether there are other possible and acceptable translation options. The reason why (general) bilingual dictionaries are examined

in this paper is due to the fact that many professional translators, students in Translation Studies and legal professionals consult such generic language resources instead of *ad hoc* tools (e.g., legal and/or technical dictionaries) (Giampieri, 2024, p. 54). Additionally, as this paper addresses neutral formulae, general bilingual dictionaries may provide adequate translations.

In light of the above, the following research questions are posed: 1) Can the EUR-Lex corpora uploaded to the Sketch Engine platform be considered reliable language resources when tackling and translating LELF legal formulae? 2) If so, how can they be consulted in order to produce formulae from English (as a *lingua franca*) into French and Italian?

3. Methodology

3.1. The legal formulae addressed

The French and Italian EUR-Lex corpora uploaded to the Sketch Engine platform are analysed to translate the following legal formulae into French and Italian, respectively: “having regard to the opinion of”; “hereinafter referred to as” (or “hereinafter called”); “by virtue of”; “as enshrined in”; “by way of derogation from”; “under no circumstances”, and “in witness whereof”. These general formulae appear in a wide variety of legal texts, irrespective of the subject-matter addressed (Bhatia, 2010). They are, in fact, mentioned in both public and private legal documents, such as laws, contracts, powers of attorneys, courtroom speeches, international treaties, and so on (Giampieri, 2025). Some of them (e.g., “having regard to the opinion of”) characterise statutory instruments or public deeds, whereas others (such as “as enshrined in”) are frequently used in international documents and are examples of LELF. By scrutinising the legal documents composing the Digital Library of the United Nations, for example, the expression “as enshrined in” comes to the fore in several instances.

Additionally, most of the formulae above mentioned appear in the Casemine portal, which addresses the British, Irish, North-American and Indian case-law. The same can be said of the BoLC (Bononia Legal Corpus) (Rossini Favretti et al., 2007), which tackles English acts and statutory documents. As far as private legal documents are concerned (e.g., contracts and agreements), the LawInsider contract database and the Terms of Service (ToS) corpora implemented by the author (Giampieri, 2024) yields many results with the multi-words tackled in this paper. It is, hence, interesting to explore how they can be translated into French and Italian in international settings.

3.2. How French and Italian equivalents are found

To source the contextual equivalents of the selected formulae, search queries are formulated through the Parallel Concordancer of the Sketch Engine platform via the 2024 EUR-Lex database. This dataset includes sub-corpora for various EU language combinations and addresses EU Legislation (such as directives, regulations, decisions), and the EU Case-Law (i.e., judgments of the Court of Justice of the European Union). The source language is English, whereas the target languages are French and Italian. Exact multi-word searches are carried out to ensure the comparability of the data across the three languages. For each formula, the first 100 aligned results are extracted to analyse authentic word usages in context.

In addition to the search in the EUR-Lex parallel corpora, online lexicographical resources are consulted. More precisely, the following bilingual dictionaries are considered: Cambridge, Collins and WordReference. WordReference is explored along with traditional lexicographical tools because it is frequently used by translator trainees and legal practitioners in their daily workflow (Giampieri, 2024).

4. Analysis

In this section, the EUR-Lex parallel corpora are queried in order to find French and Italian equivalents. Afterwards, online dictionaries are consulted to corroborate translation options and/or find alternative renditions.

The first legal formulae analysed is “having regard to the opinion of”. This multi-word sets the ground for new decisions or provisions. It refers, in fact, to the opinions or decisions of bodies or institutions on the basis of which provisions are established or new actions are undertaken. Table 1 reports results on “having regard to the opinion of”.

English Formula	Source	French equivalent	Italian equivalent
Having regard to the opinion (of)	EUR-Lex	<i>Vu l'avis (de)</i>	<i>Visto il parere di</i>
	Cambridge	(only results for “opinion”) <i>Opinion</i> <i>Avis</i>	(only results for “opinion”) <i>Opinione</i> <i>Parere</i>
	Collins	(only results for “opinion”) <i>Opinion</i> <i>Avis</i>	(only results for “opinion”) <i>Opinione</i> <i>Parere</i>
	WordReference	(only results for “opinion”) <i>Opinion</i> <i>Avis</i>	(only results for “opinion”) <i>Giudizio</i>

Table 1. French and Italian equivalents of “having regard to the opinion of”

As observable in Table 1, dictionaries show fragmented results, as they only propose translations of the single word “opinion”, i.e., “avis” and “opinion” (French), and “parere”, “opinione” and “giudizio” (Italian). By contrast, the EUR-Lex platform displays full equivalents of the source phrase, such as “vu l’avis de” (French) and

“visto il parere di” (Italian).

Table 2 illustrates corpus-driven source and target phrases in context. As can be noticed, the use of the English formula and of its French and Italian equivalents is similar.

English	French	Italian
Having regard to the opinion of the Economic and Social Committee.	<i>Vu l'avis du Comité économique et social.</i>	<i>Visto il parere del Comitato economico e sociale.</i>
Having regard to the opinion of the European Parliament.	<i>Vu l'avis du Parlement européen.</i>	<i>Visto il parere del Parlamento Europeo.</i>

Table 2. Usages in context of “having regard to the opinion of” and its French and Italian equivalents.

The second formula is “hereinafter referred to as”, or “hereinafter called”. These expressions abbreviate the names of the parties to a contract or a convention. They also shorten the names of entities referred to in a document. Table 3 below exhibits the equivalents retrieved from the EUR-Lex corpora and from dictionaries.

English Formula	Source	French equivalent	Italian equivalent
Hereinafter referred to as	EUR-Lex	<i>Ci-après dénommé</i>	<i>In prosieguo denominata (In) appresso denominato Qui di seguito denominato / definito In seguito denominato Nel seguito denominato / designato / chiamato D'ora in poi denominato</i>
Hereinafter called	Cambridge	(only results for “hereafter”) <i>Ci-après</i>	(only results for “hereafter”) <i>D'ora in poi</i>

	Collins	(only results for “hereafter”) <i>Ci-après</i>	(only results for “hereafter”) <i>D’ora in poi</i> <i>Da qui in avanti</i> <i>In futuro</i>
	WordReference	(only results for “hereinafter”) <i>Ci-après</i>	(only results for “hereinafter”) <i>Di seguito</i> <i>Nel seguito</i>

Table 3. French and Italian equivalents of “hereinafter referred to as” and/or “hereinafter called”

As with the first formula analysed, dictionaries only propose partial translations and suggest equivalents of single words such as “hereafter” or “hereinafter”. Also, it is worthwhile mentioning that the expression “*in futuro (denominato)*”, as partly indicated in the Collins dictionary, generates no occurrences in EUR-Lex.

Conversely, the EUR-Lex corpora feature full equivalents in both languages: the French one is “*ci-après dénommé*”, whereas the Italian ones are “*nel seguito denominato*”, “*nel seguito designato come*”, “*nel seguito chiamato*”, “*qui di seguito denominato*”, “*da qui in avanti denominato*”, “*qui di seguito definito*”, and “*d’ora in poi denominato*”.

A sample phrase sourced from the EUR-Lex parallel corpora is the following one: “these countries and territories (hereinafter called the ‘countries and territories’)”, which corresponds to “*ces pays et territoires, ci-après dénommés ‘pays et territoires’*” in French and “*questi paesi e territori, qui di seguito chiamati paesi e territori*” in Italian. As can be grasped, the word usages in context are the same across the three languages. Table 4 illustrates these and other phrases sourced from the EUR-Lex parallel corpora.

English	French	Italian
These countries and territories (hereinafter called the “countries	<i>Ces pays et territoires, ci-après dénommés “pays et territoires”.</i>	<i>Questi paesi e territori, qui di seguito chiamati paesi e territori.</i>

and territories”).		
A Special Council Of Ministers (hereinafter referred to as ‘the Council’).	<i>Un Conseil Special de Ministres, ci-après dénommé “le Conseil”.</i>	<i>Un Consiglio Speciale dei Ministri, appresso denominato il “Consiglio”.</i>
The European Investment Bank established by Article 129 of this Treaty (hereinafter called the “Bank”).	<i>La Banque Européenne D’investissement instituée par l’article 129 du Traité, ci-après dénommée la “Banque”.</i>	<i>La Banca Europea per gli Investimenti istituita dall’articolo 129 del Trattato, in seguito denominata la “Banca”.</i>
The Food and Agriculture Organisation of the United Nations (hereinafter referred to as “the Organisation”).	<i>L’Organisation des Nations unies pour l’alimentation et l’agriculture (ci-après dénommée “Organisation”).</i>	<i>L’Organizzazione delle Nazioni Unite per l’alimentazione e l’agricoltura (in prosieguo denominata “l’Organizzazione”).</i>

Table 4. Usages in context of “hereinafter referred to as” and “hereinafter called” and some of their French and Italian equivalents.

The next formula is “by virtue of”, which refers to articles, paragraphs or provisions of conventions, treaties, contracts, agreements, or laws. Table 5 below contains data sourced from the three parallel corpora and from dictionaries.

English Formula	Source	French equivalent	Italian equivalent
By virtue of	EUR-Lex	<i>En vertu de</i>	<i>In virtù di In forza di A norma di</i>
	Cambridge	-	-
	Collins	<i>En vertu de</i>	(only results for “virtue”) <i>Virtù</i>
	WordReference	<i>Grâce à En vertu de</i>	<i>In virtù di Grazie a</i>

Table 5. French and Italian equivalents of “by virtue of”

It is evident that some dictionaries provide full equivalents, whereas others indicate partial or no results (e.g., the Collins dictionary only suggests a translation of “virtue” into Italian, whereas the Cambridge dictionary produces no translation). For instance, the French “*en vertu de*” appears in the French section of the EUR-Lex. This equivalent is also featured in two dictionaries (i.e., Collins and WordReference, respectively). In Italian, the corpus provides diverse translation options, such as “*in virtù di*”; “*in forza di*” and “*a norma di*”, although only the first one appears in the WordReference dictionary.

The expressions “*grâce à*” and “*grazie a*” (as proposed by WordReference) do not collocate with the equivalents of “article” or “provision” in the French and Italian versions of the EUR-Lex corpora. For example, the French EUR-Lex sub-corpus generates instances of “*grâce à leur composition*” (back-translation: “thanks to their composition”), and “*grâce à l’insistance de la Commission*” (back-translation: “thanks to the Commission’s insistence/perseverance”). The dataset does not contain phrases such as “*grâce à l’article*”, and it only shows one occurrence of “*grâce à la disposition de l’article*” (back-translation: “thanks to the provision of the article”). The same can be said of the Italian language, where “*grazie all’articolo*” produces no results. Conversely, there are a few hits with “*grazie alle disposizioni di*” (back-translation: “thanks to the provisions of”). In practice, the expressions “*grâce à*” and “*grazie a*” have broader usages than “by virtue of”. For these reasons, they can be considered acceptable equivalents only partly.

Some sample phrases sourced from the EUR-Lex corpora containing “by virtue of” at its equivalents in French and Italian are the following ones: “by virtue of the provisions of article 4 (4) of that regulation”, which corresponds to “*en vertu des dispositions de l’article 4, paragraphe 4, du règlement*” and “*in forza dell’art. 4, n .4, del regolamento stesso*” in French and Italian, respectively. Other examples are listed in Table 6 below.

English	French	Italian
By virtue of the provisions of article 4 (4) of that regulation.	<i>En vertu des dispositions de l'article 4, paragraphe 4, du règlement.</i>	<i>In forza dell'art. 4, n. 4, del regolamento stesso.</i>
By virtue of this Treaty.	<i>En vertu du present Trait.</i>	<i>In virtù del presente Trattato.</i>
By virtue of article 32 (3) (c) of regulation (EEC) no 222/77.	<i>En vertu de l'article 32 paragraphe 3 sous c) du règlement (CEE) no 222/77.</i>	<i>In virtù dell'articolo 32, paragrafo 3, lettera c) del regolamento (CEE) n. 222/77.</i>
By virtue of article 28 of regulation (EEC) no 3330/74.	<i>En vertu de l'article 28 du règlement (CEE) no 3330/74.</i>	<i>A norma dell'articolo 28 del regolamento (CEE) n. 3330/74.</i>

Table 6. Usages in context of “by virtue of” and some of its French and Italian equivalents.

The multi-word “as enshrined in” encapsulates the content of articles, treaties and/or conventions. It is generally mentioned in documents signed by multinational parties, such as those of the United Nations or the European Union. This formula generally applies to highly valued principles, Charters, Constitutions, or (human) rights¹. The LawInsider contract platform, in fact, reports phrases such as “as enshrined in the UN Universal Declaration of Human Rights” or “the peoples’ fundamental rights and freedoms as enshrined in chapter 4 of the Constitution”. By contrast, the English corpus of Terms and Service implemented by Giampieri (Giampieri, 2024), produces no instances of “as enshrined”. As a matter of fact, “as enshrined in” is not employed to refer to articles or clauses in private legal documents such as contracts or agreements (Giampieri, 2024 and 2025). However, irrespective of the aim and scope of the phrase “as enshrined in”, analysing its possible equivalents is interesting and relevant, especially in international settings where LELF is used.

¹ The The 2005 UN Convention on the Protection and Promotion of the diversity of cultural expressions, for example, mentions “as enshrined”. See <https://www.unesco.org/creativity/en/2005-convention>.

Table 7 illustrates equivalents sourced from dictionaries and the EUR-Lex corpora.

English Formula	Source	French equivalent	Italian equivalent
As enshrined in	EUR-Lex	<i>Être consacré par Être sanctionné par Tel que sanctionné / consacré par / dans</i>	<i>Qual è sancito da Come sancito da Quale formulato da</i>
	Cambridge	-	-
	Collins	<i>Être garanti(e) par</i>	(only results for “enshrine”) <i>Custodire</i>
	WordReference	(only results for “enshrine”) <i>Conserver</i>	(only results for “enshrine”) <i>Custodire Conservare Racchiudere</i>

Table 7. French and Italian equivalents of “as enshrined in”

As evident, bilingual dictionaries provide either limited translations (e.g., entries for the single word “enshrine”) or no translation at all (see the Cambridge dictionary). In addition, dictionary translation options are very different from the equivalents obtained from the parallel corpora. It is noteworthy that some dictionaries propose literal renditions of “enshrine” (especially Collins and WordReference in their translations into Italian), whereas legal documents refer to its pragmatic (or functional) meaning.

Sample phrases sourced from the EUR-Lex corpora are the following ones: “cohesion policy, as enshrined in the Treaties”, rendered as “*la politique de cohésion, telle que consacrée dans les Traités*” and “*la politica di coesione, così come sancita dai trattati*”. Table 8 contains more instances in the three languages.

English	French	Italian
Cohesion policy, as enshrined in the Treaties.	<i>La politique de cohésion, telle que consacrée dans les Traités.</i>	<i>La politica di coesione, così come sancita dai trattati.</i>
The freedom of expression as enshrined in Article 10 (1) of the Convention for the Protection of Human Rights.	<i>La liberté d'expression telle qu'elle est consacrée par l'article 10 paragraphe 1 de la convention de sauvegarde des droits de l'homme.</i>	<i>Qual è sancito dall'articolo 10, paragrafo 1 della Convenzione sulla salvaguardia dei diritti dell'uomo.</i>
The principle of non-discrimination on the grounds of nationality between Community citizens, as enshrined in the Treaty.	<i>Le principe de la non-discrimination entre les citoyens communautaires sur la base de la nationalité, principe consacré par le Traité.</i>	<i>Il principio della non discriminazione tra i cittadini comunitari in base alla nazionalità, principio sancito dal Trattato.</i>
A violation of freedom of expression as enshrined in that article of the Convention.	<i>Une violation de la liberté d'expression sanctionnée par l'article précité de la convention.</i>	<i>Libertà di espressione qual è sancito dall'articolo 10, paragrafo 1 della Convenzione.</i>

Table 8. Usages in context of “as enshrined in” and some of its French and Italian equivalents.

The next formula focused on is “by way of derogation from”. This expression is applied when making derogations, or exceptions, to articles or provisions. For this reason, this multi-word is generally followed by the article, comma, or paragraph it is intended to derogate from. Table 9 illustrates the equivalents retrieved from bilingual dictionaries and the EUR-Lex corpora.

English Formula	Source	French equivalent	Italian equivalent
By way of derogation from	EUR-Lex	<i>Par dérogation à</i>	<i>In deroga a</i>
	Cambridge	(only results for “derogation”)	-

		<i>Dérogation</i>	
	Collins	(only results for “derogation”) <i>Dérogation</i>	-
	WordReference	(only results for “derogation”) <i>Dérogation</i>	(only results for “derogation”) <i>Privazione</i> <i>Immunità</i>

Table 9. French and Italian equivalents of “by way of derogation from”

As can be understood, “by way of derogation from” does not appear in all dictionaries, which do not show any results. Alternatively, they propose translations of the single word “derogation”. Also, the WordReference dictionary suggests options imbued with very different meanings from the one of the source phrase (such as “*privazione*” and “*immunità*”, meaning “deprivation” and “immunity”, respectively). Therefore, dictionary results may be misleading if words are not applied in context and/or are not fully understood in the source and target language.

A sample phrase retrieved from the EUR-Lex corpora is “by way of derogation from the provisions of Article 9 of this Convention”, whose French and Italian equivalents are “*par dérogation aux dispositions du paragraphe 9 de la présente Convention*” and “*in deroga alle disposizioni del paragrafo 9 della presente Convenzione*”, respectively. Table 10 shows these and other examples in context.

English	French	Italian
By way of derogation from the provisions of Article 9 of this Convention.	<i>Par dérogation aux dispositions du paragraphe 9 de la présente Convention.</i>	<i>In deroga alle disposizioni del paragrafo 9 della presente Convenzione.</i>
By way of derogation from the rules laid down in paragraph 3, the same procedures shall apply.	<i>Les mêmes procédures sont applicables, par dérogation aux règles du paragraphe 3.</i>	<i>In deroga alle norme previste dal paragrafo 3, si applicano le stesse procedure alle</i>

		<i>decisioni.</i>
By way of derogation from paragraph 2.	<i>Par dérogation aux dispositions du paragraphe 2.</i>	<i>In deroga al paragrafo 2.</i>

Table 10. Usages in context of “by way of derogation from” and some of its French and Italian equivalents.

An interesting formula that, to some extent, is the opposite of the previous one is “under no circumstances”. This expression, in fact, excludes derogations. Table 5 below lists its possible renderings, as appearing in the EUR-Lex corpora and in bilingual dictionaries.

English Formula	Source	French equivalent	Italian equivalent
Under no circumstances	EUR-Lex	<i>D’aucune manière</i> <i>En aucun cas</i>	<i>In nessun caso</i> <i>In nessun modo</i>
	Cambridge	<i>En aucun cas</i>	(only results for “circumstances”) <i>Circostanze</i>
	Collins	<i>En aucun cas</i>	(only results for “circumstances”) <i>Circostanze</i>
	WordReference	<i>En aucun cas</i> <i>Sous aucun prétexte</i>	<i>In nessun caso</i> <i>In alcun caso</i>

Table 11. French and Italian equivalents of “under no circumstances”

The French dictionary-based translations of “under no circumstances” are satisfactory, as results indicate “*en aucun cas*”. This prepositional phrase is also found in the French EUR-Lex sub-corpus, along with “*d’aucune manière*”. By contrast, the Italian dictionary-driven translations are not suitable, given that two sources only provide the translation of the single word “circumstances”, i.e., “*circostanze*”. The Italian EUR-Lex sub-corpus proposes “*in nessun*

caso” and “in nessun modo”. As can be seen, there is no mention of the word “circostanza” (back-translation: “circumstance”).

A sample phrase with “under no circumstances” sourced from the EUR-Lex corpora is the following: “tax due under that provision may, according to article 17 (2), under no circumstances be deducted” whose French and Italian equivalents are as follows: “*la taxe due en vertu de cette disposition ne pourrait, en aucun cas, d’après l’article 17, paragraphe 2, être déduite*” and “*l’imposta dovuta in forza di questa disposizione non può, in nessun caso, secondo l’art. 17, n. 2, essere detratta*”. Table 12 below lists these and other corpus-driven sample sentences.

English	French	Italian
Refusing to agree to a period which could under no circumstances be regarded as reasonable and bona fide.	<i>Refuser, compte tenu des nécessités du service, un délai qui ne pouvait d’aucune manière être considéré comme raisonnable et sérieux.</i>	<i>Rifiutare la proposta di un termine che in nessun modo poteva considerarsi ragionevole e serio.</i>
Whereas it seems appropriate also to specify that the period of validity of licences issued in 1983 shall under no circumstances be valid beyond 31 December 1983.	<i>Considérant qu’il semble en outre approprié de préciser que la durée de validité des certificats délivrés en 1983 ne pourra en aucun cas dépasser le 31 décembre de cette année.</i>	<i>Considerando che è inoltre opportuno precisare che la durata di validità dei titoli rilasciati nel 1983 non potrà in nessun caso superare il 31 dicembre di quest’anno.</i>
Tax due under that provision may, according to article 17 (2), under no circumstances be deducted.	<i>La taxe due en vertu de cette disposition ne pourrait, en aucun cas, d’après l’article 17, paragraphe 2, être déduite.</i>	<i>L’imposta dovuta in forza di questa disposizione non può, in nessun caso, secondo l’art. 17, n. 2, essere detratta.</i>

Table 12. Usages in context of “under no circumstances” and some of its French and Italian equivalents.

The last formula tackled is “in witness whereof”. This is a

Patrizia Giampieri: How EUR-lex parallel corpora...

closing formulaic expression stating the parties' willingness to comply with the provisions of a deed, an agreement, or a convention. The phrase is generally followed by the parties' signatures. Table 13 exhibits its French and Italian equivalents.

English Formula	Source	French equivalent	Italian equivalent
In witness whereof	EUR-Lex	<i>En foi de quoi</i>	<i>In fede di che</i>
	Cambridge	-	-
	Collins	-	-
	WordReference	<i>En foi de quoi</i>	<i>In fede</i>

Table 13. French and Italian equivalents of “in witness whereof”

Table 13 clearly illustrates that not all dictionaries are helpful, whereas the EUR-Lex corpora produce the following equivalences: “*en foi de quoi*” (French) and “*in fede di che*” (Italian). Table 14 reports some instances.

English	French	Italian
IN WITNESS WHEREOF, the undersigned, being duly authorized for this purpose, have signed this Agreement.	<i>En foi de quoi, les soussignés, dûment habilités à cet effet, ont signé le présent accord.</i>	<i>In fede di che, i sottoscritti, debitamente autorizzati a tal fine, hanno firmato il presente accordo.</i>
In witness whereof, the undersigned representatives of the Federal Council, of the Governments of the Member States of the Community, and of the Commission of the European Communities, duly authorized, have signed	<i>En foi de quoi, les représentants soussignés du Conseil fédéral, des gouvernements des Etats Membres de la Communauté et de la Commission des Communautés Européennes dûment autorisés, ont apposé leurs signatures au bas du</i>	<i>In fede di che i sottoscritti rappresentanti del Consiglio federale, dei governi degli Stati membri della Comunità e della Commissione delle Comunità Europee, debitamente autorizzati, hanno apposto le loro firme in</i>

this protocol.	<i>présent protocole.</i>	<i>calce al presente protocollo.</i>
----------------	---------------------------	--

Table 14. Usages in context of “in witness whereof” and some of its French and Italian equivalents.

5. Discussion

The analysis carried out in the sections above reveals that general dictionaries may not be accurate resources to tap into when addressing and translating legal texts. Although it was not the aim of this paper to verify and explore the effectiveness or usefulness of (general) dictionaries in legal translation, this is a result that corroborates literature findings (Biel, 2008). As shown in the tables above, in fact, lexicographical tools do not always provide the most adequate or suitable translation options. Conversely, they tend to propose phrase fragments and partial renditions, as in the case of “by virtue of” or “by way of derogation from”, where the suggested translations only revolve around the term “virtue” or “derogation”, respectively.

Sometimes dictionary entries are misleading, as in the case of “by way of derogation from”, where the Italian translation of “derogation” is “*privazione*” (back-translation: “deprivation”) or “*immunità*” (back-translation: “immunity”) (see the WordReference entries in Table 9). Such translation options do not correspond to the source phrase “by way of derogation from”.

In some other cases, dictionaries propose words that are not featured in the EUR-Lex corpora. An example is the expressions “hereinafter (referred to as)”, rendered as “*in futuro*” in Italian (back-translation: “in the future”). If “*in futuro*” is queried in the Italian sub-corpus of the EUR-Lex (or the BoLC), no occurrences are obtained where “*in futuro*” means “hereinafter (referred to as)”. Another case in point is “by virtue of”, translated as “*grâce à*” in French (back-translation: “thanks to”). If “*grâce à l’article*” is searched for in the Legifrance portal (which is the public service for the dissemination of French law), almost no hits are retrieved.

To provide a summary of the analysis, Table 15 outlines the

utility and performance of the lexicographical tools against the EUR-Lex corpora across the seven formulae.

Legal Formula (EN)	EUR-Lex Corpora	Cambridge Dictionary	Collins Dictionary
1. having regard to the opinion of	Full Equivalence	Phrase Fragment	Phrase Fragment
2. hereinafter referred to as	Full Equivalence	Phrase Fragment	Phrase Fragment
3. by virtue of	Full Equivalence	Zero Output	Phrase Fragment
4. as enshrined in	Full Equivalence	Zero Output	Phrase Fragment
5. by way of derogation from	Full Equivalence	Phrase Fragment	Phrase Fragment
6. under no circumstances	Full Equivalence	Phrase Fragment	Phrase Fragment
7. in witness whereof	Full Equivalence	Zero Output	Zero Output

Table 15. Performances: dictionaries vs EUR-Lex corpora

Legend

Full Equivalence = Phrase-level match corresponding to institutional usage

Phrase Fragment = Isolated headword entry only

Misleading = Inaccurate meanings

Zero Output = No entry

As can be guessed by analysing Table 15, it can be argued that when addressing LELF, general dictionaries should be complemented by other tools.

The results reported in Table 15 also show that such a cautious approach does not apply uniformly; for formulae employed in different genres such as “under no circumstances” or “by virtue of”, bilingual dictionaries can occasionally provide acceptable phrase-level outputs (such as “*en aucun cas*” or “*en vertu de*”). Hence, the primary failure of traditional lexicography lies in its inability to disambiguate multi-word legal items, thus offering phrase fragments.

6. Conclusions

This paper aimed at exploring whether and to what extent the French and Italian EUR-Lex corpora can provide acceptable equivalents of some legal formulae of English as a *lingua franca*. For this purpose, the paper focused on general legal multi-words. The results highlight the usefulness of the EUR-Lex corpora. On the one hand, such datasets contain and provide acceptable (or full) equivalents of the formulae queried; on the other hand, they show usages in context and, hence, they help notice patterns of language. Therefore, it can be claimed that the EUR-Lex corpora are valid language resources that can be used to understand and translate neutral legal formulae in international settings (namely, when addressing LELF). It is evident that this claim is limited to the analysis of English as a *lingua franca*. Legal formulaic expressions, in fact, are system-bound and each legal jurisdiction (and country) may use different ones.

Unexpectedly, this paper brought to the fore the limits of non-technical bilingual dictionaries, which may not always provide full or reliable translations of legal formulae. At times, they could even frame false equivalences.

Future analyses may consider legal dictionaries. For example, the lexicographical resources proposed by De Groot and Van Laer (2008) could be taken into account. Researchers may explore to what extent such dictionaries are satisfactory, especially when their entries are compared with the translation options appearing in the EUR-Lex parallel corpora.

The weakness of this paper lies in the limited number of dictionaries and formulaic expressions considered. Also, this study did not address other EU languages, or other language resources, such as multilingual platforms (e.g., Reverso Context or Linguee). These interfaces could be investigated in order to corroborate or confute the findings of this initial study.

Conflict of Interest Statement

The authors declare no conflicts of interest to disclose related to this manuscript. If any conflicts arise in the future, the Authors will promptly inform the journal.

AI Use Statement

The authors confirm that artificial intelligence (AI) tools were not used in any aspects of this research.

Reference List

- Abate, S. C. (1998). *Il documento legale anglosassone*. Hoepli.
- Anesa, P. (2019). Towards a conceptualization of Legal English as a Lingua Franca? *International Journal Of English Linguistics*, 9(6), 14-21.
- Aston, G. (2001). *Learning with corpora*. Athelstan.
- Bhatia, V. K. (2010). *Textbook on legal language and legal writing*. Universal law publishing Co. Pvt. Ltd.
- Bhatia, V. K., Langton, N. M, & Lung, J. (2004). Legal discourse: opportunities and threats for corpus linguistics. In U. Connor & T. A. Upton (Eds), *Discourse in the professions. perspectives from corpus linguistics* (pp. 203-231). John Benjamins.
- Biel, Ł. (2008). Legal terminology in translation practice: Dictionaries, googling or discussion forums. *Skase Journal of Translation and Interpretation*, 3(1), 22-38.
- Biel, Ł. (2017). Enhancing the communicative dimension of legal translation: comparable corpora in the research-informed classroom. *The Interpreter and Translator Trainer*, 11(4), 316-336.
- Bowker, L., & Pearson, J. (2002). *Working with specialized language: s practical guide to using corpora*. Routledge.

- Campos-Pardillos, M. Á. (2010). Going beyond the obvious in English for legal purposes: a few remarks on international legal English as a lingua franca in Europe. In Á. L. López & R. Crespo Jiménez (Eds.), *Professional English in the European context: the EHEA challenge* (pp. 175-194). Peter Lang.
- Civil Justice Council. 2013. *A guide to bringing and defending a small claim*. <https://www.judiciary.uk/wp-content/uploads/JCO/Documents/CJC/Publications/Other+pape rs/Small+Claims+Guide+for+web+FINAL.pdf> [Accessed on 14/09/2026]
- Coulthard, M., & Johnson, A. (2010). *The Routledge handbook of forensic linguistics*. Routledge.
- De Groot, G. R., & Van Laer, C. J. P. (2008). *The quality of legal dictionaries: an assessment*. University of Maastricht.
- Engberg, J. (2020). Comparative law for legal translation: through multiple perspectives to multidimensional knowledge. *International Journal for the Semiotics of Law - Revue Internationale de Sémiotique Juridique*, 33(2), 263-282.
- Fantinuoli, C., & Zanettin, F. (2015). Creating and using multilingual corpora in translation studies. In C. Fantinuoli & F. Zanettin (Eds.), *New directions in corpus-based translation studies* (pp. 1-10). Language Science Press.
- Farr, F., & O’Keeffe, A. (2019). Using corpora to analyse language. In S. Walsh & S. Mann (Eds.), *Routledge handbook of English language teacher education* (pp. 268-282). Routledge.
- Garner, B. A. (1995). *A dictionary of modern legal usage*. Oxford University Press.
- Garzone, G. (2007). Osservazioni sulla didattica della traduzione giuridica. In P. Mazzotta & L. Salmon (Eds.), *Tradurre le microlingue scientifico professionali. Riflessioni teoriche e proposte didattiche* (pp. 194-238). UTET.
- Giampieri, P. (2021). *Legal English and contract law – contrattualistica comparata*. Giuffrè-Lefebvre.
- Giampieri, P. (2024). *Corpus-based translation of private legal documents*. John Benjamins.
- Giampieri, P. (2025). *Legal formulae. Exploring legal multi-words in English, Italian and French*. Palgrave Macmillan.

- Gotti, M. (2012). Text and genre. In P. Tiersma & L. M. Solan (Eds.), *The Oxford handbook of language and law* (pp. 52-66). Oxford University Press.
- Jacometti, V., & Pozzo, B. (2018). *Traduttologia e linguaggio giuridico*. Wolters Kluwer.
- Kilgarrieff, A., Rychlý, P., Smrž, P., & Tugwell, D. (2004). Itri-04-08 the sketch engine. *Information Technology*. <http://www.sketchengine.eu>
- Priya, M., Nesi, H., & Benet, V. (2018). Corpus from scratch: collecting and processing a sizeable EAP corpus in a (relatively) resource-poor context. *Proceedings of the 2017 BALEAP Conference*, 10, 93-102.
- O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom*. Cambridge University Press.
- Orts Llopis, M. Á. (2009). Legal genres in English and Spanish: some attempts of analysis. *Ibérica*, 18, 109–130.
- Prieto Ramos, F. (2020). The use of resources for legal terminological decision-making: patterns and profile variations among institutional translators. *Perspectives*, 29(2), 278-310.
- Rossini Favretti, R., Tamburini, F., & Martelli, E. (2007). Words from Bononia legal corpus. In W. Tuebert (Ed.), *Text corpora and multilingual lexicography* (pp. 5-30). John Benjamins.
- Šarčević, S. (1997). *New approach to legal translation*. Kluwer Law International.
- Stein, D. (2009). Introduction: forensic translation — practical issues and their theoretical underpinnings. In O. Frances, L. Alexander & St. Dieter (Eds.), *Translation issues in language and law* (pp. 1-4). Palgrave Macmillan.
- Tiersma, P. (1999). *Legal language*. The University of Chicago Press.
- Tiersma, P., & Solan, L. (2012). *The Oxford handbook of language and law*. Oxford University Press.
- Vigier Moreno, F. J., & Sánchez Ramos, M. M. (2017). Using parallel corpora to study the translation of legal system-bound terms: the case of names of English and Spanish courts. In R. Mitkov (Ed.), *Computational and corpus-based phraseology EUROPHRAS 2017, lecture notes in computer science*, 10596 (pp. 206-273). Springer.

- Williams, C. (2004). Legal English and plain language: an introduction. *ESP Across Culture*, 1, 51-124.
- Williams, C. (2005). Vagueness in legal texts: is there a future for shall? In M. Gotti, V. Bhatia, J. Engberg & D. Heller (Eds.), *Vagueness in Normative Texts* (pp. 201-204), Peter Lang.
- Zanettin, F. (2002). DIY corpora: the WWW and the translator. In B. Maia, J. Haller & M. Ulrych (Eds.), *Training the language services provider for the new millennium* (pp. 239-248). Universidade do Porto.
- Zanettin, F. (2014). Corpora in translation. In J. House (Ed.), *Translation: a multidisciplinary approach* (pp. 178-199). Palgrave Macmillan.

Online Resources

Cambridge Dictionary: <https://dictionary.cambridge.org/>

Collins Dictionary: <https://www.collinsdictionary.com/>

Digital Library of the United Nations:
<https://www.digitallibrary.un.org>

EUR-Lex: <https://eur-lex.europa.eu/>

The Casemine: <https://www.casemine.com>

The BoLC: <https://corpora.ficlit.unibo.it/BOLCCorpQuery.html>

The LawInsider: <https://www.lawinsider.com>

The Legifrance: <https://www.legifrance.gouv.fr/>

The Sketch Engine platform: <https://app.sketchengine.eu>

WordReference Dictionary: <https://www.WordReference.com/>