PHRASEOLOGY IN CRIMINAL JUDGMENTS:
A CORPUS STUDY OF ORIGINAL VS. TRANSLATED ITALIAN

Gianluca Pontrandolfo
Universidad de Trieste
gianluca.pontrandolfo@phd.units.it

Resumen

Este artículo presenta la metodología y los resultados preliminares de un estudio piloto monolingüe que se inserta en el marco de una tesis de doctorado interdisciplinar en curso sobre un subgénero de textos judiciales, es decir, sentencias penales en inglés, español e italiano. El estudio pretende comparar el uso de los fraseologismos en un corpus monolingüe comparable de sentencias italianas dictadas por el Tribunal de última instancia nacional ("Corte di Cassazione") y sentencias emitidas a nivel supranacional (por el "Tribunal de Justicia de la Unión Europea" y el "Tribunal Europeo de Derechos Humanos") traducidas al italiano.

Palabras clave: Análisis de géneros jurídicos/judiciales; análisis de corpus de textos jurídico/judiciales; traducción judicial; fraseología; sentencias penales.

Abstract

This paper presents the methodology and the preliminary results of a monolingual pilot study conducted within the framework of an ongoing interdisciplinary doctoral thesis focused on a subgenre of judicial texts, namely criminal judgments, in English, Spanish and Italian. The study compares the use of phraseologisms in a monolingual comparable corpus of Italian judgments delivered by the national Supreme Court ("Corte di Cassazione"), and judgments delivered at a supranational level (the "Court of Justice of the European Union" and the "European Court of Human Rights") translated into Italian.

Keywords: Genre analysis of legal/judicial texts; corpus analysis of legal/judicial texts; judicial translation; phraseology; criminal judgments.

1. Introduction

The present study aims to analyse the preliminary results of a pilot study conducted within the framework of an ongoing Ph.D. research project dealing with a translation-oriented contrastive study of the phraseologisms in use in a subgenre of judicial texts, i.e. criminal judgments, in a multilingual and multicultural environment

(England and Wales, Spain and Italy). From a methodological point of view, the study combines a corpus and a genre perspective, the former lying in the collection and exploitation of a multilingual comparable corpus of criminal judgments, the latter stemming from this investigation in terms of *genre analysis* (Swales 1990; Bhatia 1993). The main objective of the project[1] is very similar to Lombardi's research objectives (2004: 173): to provide legal translators dealing with criminal procedure with a multifunctional tool that has a positive impact on the translation process. The classification of the specialised phraseologisms used in the genre "criminal judgment" will provide legal translators with a practical guide containing useful information on the contexts of use and, above all, the frequency of some expressions. Such a tool will help them in the stylistic rendering of their target texts and "reassure" them about the appropriate linguistic and legal use of specialised phraseological combinations.

Throughout this study reference will be made to the wider definition of phraseology provided by Granger (2005: 167), who refers to the "whole range of co-occurrence patterns", and to the rather exhaustive definition of phraseologism provided by Gries (2008: 6):

> the co-occurrence of a form or a lemma of a lexical item and one or more additional linguistic elements of various kinds which functions as one semantic unit in a clause or sentence and whose frequency of co-occurrence is larger than expected on the basis of chance.

Indeed, to this day, there is no standardised term to refer to phraseology: this is mainly due to the inevitable difficulties linked to the polysemy of the term (cf. Cowie 1998; Musacchio 2002; Corpas Pastor 2003; Goźdź-Roszkowski 2006; Granger & Meunier 2008) and to the limited attention phraseology was given in early LSP studies (Musacchio & Palumbo 2008: 71).

As Guadec (2007: 23) points out, "phraseography" – phraseology mining and management – is becoming an important component task of the translation process. In fact, it is now becoming quite clear that customers and translation users want "phraseological conformity", meaning the translator uses the collocations, set phrases, sentence patterns and paragraph organisations which are particular to the domain area of the type of document concerned or used only by a particular profession group. Translators are required to produce target language (TL) texts that read naturally, i.e. that sound like TL original texts.

This is particularly true for legal translators: phraseology is indeed one of the most striking features of legal texts and judgments in particular, and it becomes

---

1. The project also follows the research lines of JUD-GENTT (Borja Albi et al., forthcoming), an action-research project aiming at improving the socio-professional conditions of legal translators working in judicial settings and their productive processes, through the use of an electronic knowledge management tool (JUD-Workstation).

particularly important for both the professional activity of translators and their training (Garzone 2007). Phraseology is also one of the key factors in evaluating the quality of a translation (Colson 2008): by using the appropriate phraseology, legal translators may produce "better" translations that conform to a specific genre. Indeed, even the most skilled translators may run the risk of producing a translation that is inaccurate from the viewpoint of its "register choices" (cf. Sinclair's idiom principle, 1991: 110-115), all other aspects of the target text being perfectly acceptable, such as grammar, content, etc. (Garzone 2007: 218-219).

In line with the use of pilot studies in social science research[2] (cf. van Teijlingen & Hundley 2001), the aim of this preliminary study was to identify potential practical problems in following the research procedure, hence the idea of a "trial run". In particular, the objectives were twofold: on the one hand, testing and assessing different modalities of phraseology extraction to choose a suitable one to be adopted in the Ph.D. research project; on the other hand, comparing the use of phraseologisms – in terms of frequencies and typologies – in judgments originally written in Italian and judgments translated into Italian.

The results of such preliminary study are still being analysed and, therefore, the objective of this paper is to comment on the very first data obtained, while reflecting on the implications of the methodology adopted and the difficulties encountered in the process. Reflecting on the challenges posed by the approach is a fundamental step towards the refining of the methodology.

The paper is divided into three main sections: the first part addresses issues concerning the methodology of the study and the material adopted; the second section is dedicated to the analysis of the preliminary results obtained with a focus on three types of phraseologisms detected, namely routine formulae, lexical collocations and prepositional phrases; finally, some conclusions are drawn on the basis of the challenges posed by the pilot study and the future steps that will be taken in the research project.

---

2. The term "pilot study" is used in two different ways in social science research. It can refer to so-called feasibility studies which are "small scale version[s], or trial run[s], done in preparation for the major study" (Polit et al., 2001: 467 in van Teijlingen & Hundley 2001). However, a pilot study can also be the pre-testing or "trying out" of a particular research instrument (Baker 1994: 182-3 in van Teijlingen & Hundley 2001). One of the advantages of conducting a pilot study is that it might give advance warning about where the main research project could fail, where research protocols may not be followed, or whether proposed methods or instruments are inappropriate or too complicated. There can be important reasons for undertaking a pilot study, such as developing and testing adequacy of research instruments, assessing the feasibility of a full-scale study, designing a research protocol, identifying logistical problems which might occur using proposed methods, etc.

2. Corpus and methodology

2.1. The comparable corpus

The monolingual comparable corpus for the analysis, made up of two sub-corpora (translated vs. non-translated judgments), has approximately 480,000 tokens[3] distributed among a hundred texts. It covers the period from 2005 and 2010. Its composition is summarised in Table 1:

Table 1. Composition of the monolingual comparable corpus

Genre	Source	Tokens	Documents
Original Italian criminal judgments	Corte di Cassazione	239,303	61
Translated Italian criminal judgments	Court of Justice of the European Union	89,259	17
	European Court of Human Rights	150,603	17
TOTAL		**479,165**	**95**

The disproportion between the two sub-corpora in terms of number of documents lies in the fact that the judgments of the *Corte di Cassazione* are shorter than those of the Court of Justice of the European Union and the European Court of Human Rights.

Both original and translated judgments were selected according to the subject field: all of them deal with criminal cases. Moreover, most of the translated judgments were selected according to their source language, i.e. French, which is the main language of both EU institutions; in this way, attempts were made to avoid problems of "directionality".

In order to define the methodology adopted in this study, it is important to identify some important features of the criminal judgments delivered by the three courts.

2.1.1. The judgments of the *Corte di Cassazione*

The original sub-corpus is made up by judgments delivered by the Italian Supreme Court (*Corte Suprema di Cassazione*), the highest court in the Italian criminal judicial system. Among its major functions there is the duty "to ensure

---

3. "While the number of *tokens* in a corpus refers to the total number of words, the number of *types* refers to the total number of unique words. For example, the word 'ship' may occur 177 times in a corpus, but it only counts as one type of word. Types are used in calculating the type/token ratio (a measure of lexical repetition) of a text or corpus" (Baker et al. 2006: 162).

the correct application of the law and its uniform interpretation". One of the main features of its task which is essentially unifying and "nomofilattica", i.e. aiming at providing certainty in the construction of law (in addition to the decision issued as third-instance judge), is that the existing rules allow the Supreme Court to investigate the facts of a case only when they have already been dealt with in the previous part of the proceedings and only if necessary to estimate the grounds allowed by the law to support a petition to the Supreme Court.

The application lodged before the Supreme Court (the so-called "ricorso in cassazione") can challenge the decisions issued by the ordinary judges of lower courts: the grounds of the application can be, in the criminal field, the infringement of provisions of substantive law ("errores in iudicando") or of procedure ("errores in procedendo") or defects (lacking, insufficient or contradictory grounds) concerning the grounds of the decision or questions relating to jurisdiction. Therefore, the judgments delivered by the Supreme Court concern the *quaestio iuris* (the court checks that laws have been applied correctly by the lower courts) and not the *quaestio facti* (the "merits" of the facts). Indeed, in the Italian criminal judgments, the reconstruction of the facts, i.e. the detailed and realistic account of the happenings, tends to lose importance "as we go up the steps of the ruling ladder" (Ondelli 2006: 304).

The judgments of the Cassazione are the product of a long historical tradition which is still perceivable in these texts. This is the main reason why the judgments of the Cassazione still represent a point of reference also for the judgments of the lower courts that tend to be shaped on the basis of the acts produced by the highest court (Taruffo 1988: 198).

For the purposes of the study, all the selected judgments were delivered by the criminal sections of the Supreme Court. They have been collected from the "case-law" section of the online De-Jure database[4].

2.1.2. The judgments of the Court of Justice of the European Union

The Court of Justice of the European Union has been given clearly defined jurisdiction which it exercises in various categories of proceedings (references for preliminary rulings, actions for failure to fulfil obligations, actions for annulment, actions for failure to act, appeals and reviews).

The first subsection of the translated sub-corpus contains fifteen judgments dealing with a "preliminary ruling procedure" and two judgments envisaging "actions for annulment". In the former case, the national courts in each EU country are responsible for ensuring that EU law is properly applied in that country. But there is a risk that courts in different countries might interpret EU law in different ways. To

_____

4. Edited by Giuffrè publishing house, De Jure is a source of legal information that contains real-time updated documents that pertain to different categories (law, case-law, legislation and law procedures, etc.).

prevent this happening, there is a "preliminary ruling procedure". This means that if a national court is in any doubt about the interpretation or validity of an EU law it may, and sometimes must, ask the Court of Justice for advice. This advice is given in the form of a "preliminary ruling". As for the latter case, if any of the member states, the Council, the Commission or – under certain conditions – Parliament believes that a particular EU law is illegal they may ask the Court to annul it. These "actions for annulment" can also be used by private individuals who want the Court to cancel a particular law because it directly and adversely affects them as individuals. If the Court finds that the law in question was not correctly adopted or is not correctly based on the Treaties, it may declare the law null and void.

All the judgments were chosen according to the field, namely "Police and judicial cooperation in criminal matters", and were collected from the Case-Law database available on the website of the EU Court of Justice.

As far as the translated texts are concerned, multilingualism characterises the activities of the Court of Justice. A language of the case is determined for each action brought before the Court. The starting point is that in direct actions the applicant has the right to choose one of the official languages to be used in proceedings (the so-called *lingua processuale*). In preliminary ruling cases, the language of the case is always the language used in the respective national proceedings in which the language has been determined according to national provisions. The language of the case is used throughout the proceedings, and as for the judgment, only the version written in the language of the case is authentic. In order to meet these linguistic requirements, the DG Translation of the Court, the Institution's largest service comprising almost half of the Court's staff, is composed of lawyer-linguists each of whom has a law degree and a thorough knowledge of several official languages. Theoretically, all official languages are used as source languages as well as target languages and, therefore, translations cover all possible combinations of the official languages. However, during the final stage of the proceedings, the Court prefers not to use translations, thus avoiding the disclosure of secret information (Capotorti 1988: 243-244). This is the reason why the Court resorts to a "lingua franca", a language presumably known by all the members of the Court – traditionally French – in which all the acts produced in the language of the case are later translated in all the other 23 official languages (Cosmai 2007: 83-85). The judgments of the Court are therefore written in French, irrespective of the language of the proceedings. Consequently, the official text of many judgments is a translation from the original French, adopted at the end of the decisional stage.

2.1.3. The judgments of the European Court of Human Rights

The European Court of Human Rights is an international court set up in 1959. It rules on individual or State applications alleging violations of the civil and political

rights set out in the European Convention on Human Rights. Since 1998 it has sat as a full-time court and individuals can apply to it directly. In almost fifty years the Court has delivered more than 10,000 judgments. These are binding on the countries concerned and have led governments to alter their legislation and administrative practice in a wide range of areas. The Court's case-law makes the Convention a powerful living instrument for meeting new challenges and consolidating the rule of law and democracy in Europe.

It has jurisdiction to hear allegations of violations of the European Convention on Human Rights (e.g. life, torture, servitude, fair trial, discrimination, crime and family, etc.) and does so on receiving individual or inter-State applications. The Convention makes a distinction between two types of application: "individual applications" lodged by any person, group of individuals, company or NGO having a complaint about a violation of their rights, and "inter-State applications" brought by one State against another. Since the Court was established, almost all applications have been lodged by individuals who have brought their cases directly to the Court alleging one or more violations of the Convention.

All the judgments delivered by the European Court of Human Rights can be accessed on the HUDOC database, available on the website of the ECHR. Documents may be in only one or both of the official languages (English and French). Translations of judgments into languages other than English and French, commissioned by the Council of Europe, have been integrated into the HUDOC database. Italian translations of the judgments included in the corpus were not available in the HUDOC database. They were collected from the website of the *Osservatorio della giurisprudenza CEDU*, an observatory of the ECHR case-law managed by the "Unione Forense per la Difesa dei Diritti dell'Uomo", under the patronage of the Italian National Forensic Council. The aim of this body is to improve the accessibility of the ECHR case-law[5] by providing unofficial translation into Italian of the most important judgments delivered by the ECHR[6]. Also in this case, translations are carried out by law experts – and not by

---

5. As far as the interest for criminal case-law is concerned, the Decisions No. 348 and 349/2007 of the Italian Constitutional Court represented a milestone. Not only has the Italian Constitutional Court clarified, through Article 117 paragraph 1 of the Italian Constitution, the European Court of Human Rights' relevant role in the domestic legal system, but it has also interpreted international obligations as an "interposed" standard of review, on the basis of which the constitutionality of domestic law must be assessed. Community case-law becomes a point of reference for domestic case-law.

6. This is clearly mentioned in the presentation of the *Osservatorio* on its website "La finalità dell'Osservatorio è quella di migliorare il grado di accessibilità della giurisprudenza della Corte europea, tuttora resa oggettivamente difficile dagli ostacoli di carattere linguistico. Le sentenze e le decisioni della Corte europea, infatti, sono redatte in inglese e/o francese e non esiste allo stato un regolare sistema di traduzione in lingua italiana. L'Osservatorio rappresenta, quindi, un valido strumento di consultazione a disposizione di avvocati, magistrati, pubbliche amministrazioni, forze dell'ordine e semplici cittadini i quali possono accedere ai suoi contenuti in modo agevole, semplificato, gratuito e, soprattutto, in lingua italiana […]. Le traduzioni in lingua italiana dai testi originali in lingua inglese e/o francese delle sentenze e decisioni della Corte europea dei diritti dell'uomo nonché gli atti del Comitato dei Ministri, pubblicati

trained translators – mastering one or more languages. However, the function of the translated documents differ significantly: while the lawyer-linguists of the Court of Justice have the obligation to produce official documents because of the multilingualism characterising the EU body, the "translators" of the *Osservatorio* produce texts that have merely an informative function, and such activity is not compulsory.

2.2. Method: stages of the analysis

From a methodological point of view, the present monolingual study combines a genre and a corpus perspective. Following the move analysis pioneered by Swales (1990) and further applied to legal cases by Bhatia (1993), judgments have been divided into three "moves", on the basis of their content and stylistic features: "Introduction" (*Introduzione*), "Facts" (*Motivazione*), "Decision" (*Dispositivo*), with the primary aim of comparing the frequency of use of phraseologisms in the different sections of the judicial acts.

The first section is an introductory paragraph in which we learn about the "main characters" of the case (composition of the Court, some personal data of the persons involved in the case, etc.). This part usually contains the "heading" of the judgment and is characterised by a high level of standardisation. The second section corresponds to Bhatia's "Arguing the case" (1993: 129-135): it is actually the most complex section of the judgment. It tends to have several "sub-moves", depending on the nature of the case and the length: in the corpus analysed, judgments present a short history of the case, the legal background (*Diritto*) and the part dedicated to the real facts (*Fatto*), ending with the *ratio decidendi* of the case. The third section corresponds to Bhatia's "Pronouncing judgment" (1993: 135-136) and contains the final decision of the Court.

Corpus analysis was carried out using WordSmith Tools[7] (version 5.0). Based on the methods of extraction adopted by a number of scholars dealing with phraseology mining through corpora (e.g. Lombardi 2004; Goźdź-Roszkowski 2006; Musacchio

---

nel presente sito, sono realizzate a cura dell'Unione forense per la tutela dei diritti dell'uomo e, pertanto, non impegnano in alcun modo il Consiglio d'Europa".

[The aim of the Observatory is to improve the level of accessibility to the case-law of the ECHR, which is now difficult because of linguistic obstacles. Judgments and decisions of the ECHR are written in English and/or French and, to this day, it does not provide any regular system of translation into Italian. This is the reason why the Observatory represents a useful tool for lawyers, magistrates, public administrations, police personnel and citizens who can easily access a simple and free content, available in their own mother tongue […]. The translations into Italian of the original judgments or decisions of the ECHR in English and/or French, and of the acts of the Committee of Ministers, published on this website, are carried out by the staff of the *Unione forense per la tutela dei diritti dell'uomo* and, therefore, they do not bind at all the Council of Europe (my translation)].

7. WordSmith Tools proved to be the most flexible tool for the purposes of the study. Preliminary analysis was carried out with TaLTaC2 and AntConc 3.2.1.

2008), wordlists of each section of the two sub-corpora were used to obtain six sets of keywords, three for the original judgments and three for the translated ones. In both cases, keywords were extracted using each sub-corpus as a reference corpus for the other, based on the assumption that the aim of the study was comparing the frequency and typologies of phraseologisms in the original and in the translated corpus.

After the extraction of those keywords that could point to candidate phraseologisms, KWIC concordances were run to investigate their use in greater detail. Each item from the keyword lists was scrutinised by means of the CONCORD feature of WordSmith Tools in order to identify its function in context and use: the functions "Collocates", "Clusters" and "Patterns" proved to be extremely important in this phase of the analysis. For fear of missing out important phraseologisms, concordances were run also from the wordlists. An inevitable dose of subjectivity was necessary while selecting the "key terms" in wordlists containing more than 1,000 words (e.g. the Facts sections).

The final phase of the analysis is currently underway. Based on the phraseologisms extracted, having distinguished them from compound terms, an attempt is being made to group the identified phraseological units according to some criteria.

In what follows, some trends detected in the use of phraseologisms in original vs. translated judgments will be exemplified.

3. Analysis

3.1 Corpus description: statistics

Table 2 shows the composition of the corpus in terms of types and tokens:

Table 2. Standardised type/token ratios in the comparable corpus

		ORIGINAL IT	TRANSLATED IT		
				CJ	ECHR
I	Tokens	6,787	13,881	7,261	6,620
	Types	682	1,700	884	1,133
	STTR	24.22	33.89	12.95	34.87
F	Tokens	226,955	219,364	78,061	141,303
	Types	14,446	11,112	4,653	9,619
	STTR	44.16	39.16	36.56	40.71
D	Tokens	3,583	6,257	3,328	2,929
	Types	369	966	686	420
	STTR	20.57	31.08	35.60	23.95
Tot.	Tokens	239,303	239,862	89,259	150,603
	Types	14,760	11,769	4,990	10,005
	STTR	43.87	39.09	36.24	40.74

A cursory glance at the figures immediately suggests that in the original judgments and in the translated ones, the "Facts" are the longest section of the judicial act; it is actually the "freest" part of the text in which judges explain what happened, beginning with the "birth" of the case ending with the reasons that drove them to the final decision, followed by the "Introduction" and the "Decision". However, the interesting statistical feature shown in the table is the type/token ratio (TTR), an indicator of lexical variation as found on the surface of a text (number of types/ number of tokens *100). It is obvious that, the more types there are in comparison to the number of tokens, the more varied is the vocabulary (the higher the percentage, the more varied the vocabulary). In fact it was found that the TTR is too sensitive to corpus size since it varies widely in accordance with the length of the text (the longer the text, the higher the TTR). Thus, aware of the fact that the TTR figure is a relative number and should be used cautiously, an average and more accurate measure is indicated in Table 2, namely the standardised type/token ratio (STTR)[8], provided by WordSmith Tools.

The findings of the statistical analysis indicate two contrasting trends. Firstly, that translations of the comparable corpus, considered as entire texts without any distinction in terms of "moves", show a lower percentage standardised type/token ratio than non-translations[9] (39.09 vs. 43.87), which could point to the conclusion that vocabulary used in the translated texts is less varied than that of the non-translated texts. Secondly, moving to the single parts of the judgments, some sections of the translated judgments show a higher standardised type/token ratio than original judgments, tokens being almost the same (e.g. I-Translated ECHR: 34.87 vs. I-Original: 24.22; D-Translated CJ: 35.60/D-Translated ECHR: 23.95 vs. D-Original: 20.57, etc.). Such a trend seems to be more marked in the translated judgments of the European Court of Human Rights. According to these figures, translators of supranational judgments would use synonymy and a richer vocabulary compared with the original writers of the judgments produced at national level. Empirical evidence of this supposed greater lexical variation should be seen, for the purpose of the study, against the background of the phraseological patterns used in the texts.

---

8. The STTR is computed every n words as Wordlist goes through each file. By default, n = 1,000. In other words, the ratio is calculated for the first 1,000 running words, then calculated afresh for the next 1,000, and so on until the end of the text or corpus. A running average is computed, which means that an average type/token ratio is obtained based on consecutive 1,000-word chunks of text. Texts with less than 1,000 words (or whatever n they are set to) will get a standardised type/token ratio of 0.

9. This is actually a rather generalised finding in the literature on translated vs. non translated-texts. Indeed, it may represent what Chesterman (2004: 40) calls a "T-universal", namely a universal difference between translations and comparable non-translated texts, i.e. a characteristic of the way translators use the target language. The hypothesis in question is the so-called "simplification" (cf. Baker 1993: 244; Laviosa 2002: 58-64): "the range of vocabulary used in the translational texts is narrower than the range of vocabulary in the non-translational texts and this difference is independent of the source language variable" (Laviosa 2002: 60). Therefore, a text which exhibits less lexical variation than another text can be considered lexically simpler, hence the simplification hypothesis.

3.2. Trends in phraseological patterns

This part of the analysis was inevitably corpus-based: dwelling on the phraseologisms extracted, different trends in the use of phraseological units in the original judgments compared with the translated ones were detected.

In what follows, three kinds of phraseological patterns which emerged from corpus browsing will be analysed:

1) Routine formulae
2) The lexical collocation Noun/noun phrase Subject + Verb/verb phrase
3) Prepositional phrases

The three typologies identified reflect three important macro-categories of phraseological units extensively used in literature[10]. An example is Corpas Pastor's classification (1996) in collocations (*colocaciones*), phrases (*locuciones*) and phraseological statements (*enunciados fraseológicos*) or Granger and Paquot's (2008) categorisation – based on Cowie (1981), Mel'čuk (1998) and Burger (1998) – in "referential phrasemes" (to which lexical collocations belong), "textual phrasemes" (to which prepositional phrases belong) and "communicative phrasemes" (to which routine formulae belong).

3.2.1. Routine formulae

The first category is evidently the easiest to recognise since it represents one of the most striking features of judgments.

Different terms have been used in literature on legal texts to refer to them (see Bachmann 2000: 51-67):

---

10. Cf. Granger & Meunier (2008: 35): "As word combinations are highly heterogeneous, linguists have quite naturally felt the need to subcategorise them. Typologies abound in the literature: some are designed for lexicological or lexicographic purposes (Gläser 1986; Cowie 1988; Moon 1998), others are pedagogically-oriented (Nattinger & DeCarrico 1992; Lewis 1993) or take a psycholinguistic perspective (Wray & Perkins 2000; Wray 2002). Several "ad hoc" descriptions have also been proposed within the field of natural language processing […]. Differences in the typologies largely correspond to differences in the selected features used to categorise multi-word units and the prioritisation of selected features." A comprehensive survey of the existing typologies of phraseological units that have been proposed in different fields is not within the scope of the present pilot study, although it will necessarily be considered for the final Ph.D. project.

Table 3. Routine formulae in literature (adapted from Bachmann 2000)

Scholar	*Term*
Fonagny (1982: 143 in Bachmann 2000)	Enunciati legati
Snell-Trampus (1989: 59 in Bachmann 2000)	Formule rituali
Stolze (1992: 190-191 in Bachmann 2000) Koutsivitis	Standardformeln Formules standardiséès
Hohnhold (1992: 255 in Bachmann 2000)	weitergehende Standardformulierung
Kjær (1992: 52-53 in Bachmann 2000)	normbedingte Wortverbindungen
Sandrini (1996: 256 in Bachmann 2000)	standardisierte Formeln
Göpferich (1995: 17 in Bachmann 2000)	textsortenspezifische Textversatzstücke
Wilss (1997: 37 in Bachmann 2000) Mandler	Textbausteine "building blocks of cognition"
Wright (1997: 16 in Bachmann 2000)	Standard texts
Rega (1997: 122, 2000: 449 in Bachmann 2000)	Formule di rito[11]
Bachmann (2000)	Formule standard
Alcaraz Varó & Hughes (2002:25-26)	Fórmulas estereotipadas
Corpas Pastor (2003: 141-147) Rojo (2009: 146-151)	Fórmulas rutinarias Routine formulae

Phraseological units, like any recurring linguistic pattern, are strongly associated with the context of situation in which they occur. However, some of them have become distinctively fossilised as situation-specific formulae that allow no changes in their structure and perform a prominent pragmatic function. According to Wright (1997: 16):

> Standard texts, often called "boilerplate" in US English, are more extended identical or nearly identical chunks of text that recur under specific circumstances. For instance, the preamble to a contract or a power of attorney is an example of a standard text. These chunks of text range from incomplete sentences to full sentences to text passages that go beyond single sentences to form paragraphs or even longer segments. Regardless of their scope, these components can be documented in and retrieved from terminological entries analogous to terms, multiword terms and phrases.

---

11. According to Rega (2000: 451), "ritual formulae" should satisfy the following requirements: a) they should be repetitive; b) they should be written; c) they should belong to the category of "performative and/or constative utterances"; d) they should have an illocutionary force and a perlocutionary effect.

For the purposes of this study, "routine formulae" are defined as recurrent expressions used exclusively in the genre "criminal judgment" (collocations and terms occurring in legal language but not in our genre are not included in this category); such formulae are generally quoted in the so-called "formulari" (formularies), collections of standardised formulae used in particular text-types (e.g. judicial acts, contractual acts, etc.), because of their crystallised use. Sometimes they can reach the length of an entire text, as they become "macro-formulae" (*formularformen*), that contain, within their boundaries, micro-formulae or other phraseologisms.

Three criteria were adopted to identify routine formulae in the corpus: firstly, their frequency of use (their repetitiveness is a crucial factor); secondly, the "clusters" they form, since they tend to occur in bundles of more than three-word sequences (this is easily manageable with the WordSmith Tools' function "Clusters" and "Patterns"), although even a single word or two-word sequence might be considered a standardised formula should they meet the last criterion, namely their position in the text: routine formulae tend to occupy a fixed position in the judgments.

In general terms, standardisation affects the judgments' sections at different levels both in original and translated texts. Corpus browsing revealed that the Introduction is the more standardised section of the judgment, and this is true not only for original judgments, but also for translated ones. The heading of the judgment is characterised by a macro-formula, whose recurring patterns appear in almost every single document of the corpus, together with some isolated routine formulae. The Decision section is also a very standardised part of the judgment, although it is less crystallised than the Introduction. Finally, the Facts section is the "freest" section: standardised formulae are limited, apart from the title of each "sub-move" (*In fatto*, *In diritto*/*Contesto normativo*, *Sulle questioni pregiudiziali*, etc.).

As far as the comparison between original and translated judgments is concerned, a higher percentage of routine formulae was found in the judgments originally written in Italian. Moreover, among the translated texts, the judgments of the Court of Justice of the European Union tend to adopt a greater number of routine formulae compared with those employed in the judgments of the European Court of Human Rights, notwithstanding the fact that both are translated by lawyer-linguists. This trend could be in line with the aforementioned lexical variation that seems to be greater in the translated judgments (cf. 3.1).

In what follows, examples will be provided of the most important routine formulae adopted in the Introduction section of original vs. translated judgments.

Text boxes 1, 2 and 3 show the heading of the original and translated judgments. Each routine formula is accompanied by the frequency of occurrence, calculated on the total number of documents included in each sub-corpus:

Text box 1. Heading of the *Corte di Cassazione*'s judgments

> *La Corte Suprema di Cassazione* [61/61]
> *Sezione X Penale* [56/61]
>
> *Composta dagli Ill.mi Sigg.ri Magistrati* [61/61]
>
> *ha pronunciato la seguente: sentenza* [61/61]
>
> *sul ricorso proposto da*: [59/61]
>
> *avverso* [58/61]
> *sentenza* [35/61] / *ordinanza* [22/61] / *decreto* [1/61]…

Text box 2. Heading of the *Court of Justice of the European Union*'s judgments

> *Nel procedimento C-X* [15/17],
> *avente ad oggetto* [15/17] *la domanda di pronuncia pregiudiziale proposta alla Corte, ai sensi dell'art. 35 UE* [15/17] da […] *con decisione* [x/y/year] *pervenuta in cancelleria* il […] [15/17], *nel procedimento penale* a carico di [13/17]
>
> *La Corte (X Sezione) composta dal sig.* [17/17]…
>
> *vista la fase scritta del procedimento* [16/17]
> *considerate le osservazioni presentate:* [14/17]
> *per il sig. X, da* Y [10/17]
> *per il governo* X, *dal sig.* Y, [13/17] *in qualità di agente* [17/17];
>
> *sentite le conclusioni dell'avvocato generale, presentate all'udienza del* […] [12/17]
>
> *ha pronunciato la seguente sentenza* [17/17]…

Text box 3. Heading of the *European Court of Human Rights*'s judgments

> *Nel caso* X *c.* Y [16/17],
> *La Corte europea dei diritti dell'uomo (sezione X), riunita in una camera composta da* [16/17]:
>
> *dopo aver deliberato in camera di consiglio* [16/17]
>
> *rende* [10/17] / *emette* [3/17] / *pronuncia* [4/17]
> *la seguente sentenza* [15/17]
> *adottata* [in tale ultima data] [16/17]…

The text boxes show a high level of standardisation in both translated and non-translated texts. However, a closer analysis of the formulae reveals an untypical

trend: the translated judgments of the Court of Justice exhibit a higher degree of phraseological uniformity compared with those of the European Court of Human Rights. This is well demonstrated by the main clause of the headings: the original judgments adopt the formula "*[La Corte] ha pronunciato la seguente sentenza*" (literally, "[The Court] has delivered the following judgment"), that occur in a well-established and fixed position, in all of the 61 documents that make up the original sub-corpus; the same phraseologism is used in all of the 17 translated judgments of the Court of Justice "*[La Corte, X sezione] ha pronunciato la seguente sentenza*", but in the translated judgments of the European Court of Human Rights, the formulae are not 100% standardised. Indeed, translators adopt synonyms of the Italian verb "pronunciare" (to deliver), namely "rendere" (to give) and "emettere" (to issue), in a context that is extremely standardised. To exclude the hypothesis of a lexical variation in their source texts, a check was carried out on the respective original judgments in both versions (French and English), which confirmed the deliberate choice of the translators to resort to synonymy (in the original texts, "*Rend l'arrêt que voici, adopté*" and "*Delivers the following judgment*" appear in 17 out of 17 judgments).

The preliminary results of the pilot study confirm this generalised trend, exemplified above, i.e. a greater lexical variation in the translated Italian that can be interpreted as a greater freedom, on the part of translators, in translating and using phraseologisms. From the point of view of the researcher dealing with this corpus, the phraseological patterns of the translated documents are much less predictable compared with the original ones.

In particular, the judgments of the European Court of Human Rights, translated by the members of the Osservatorio CEDU, show a lower degree of "phraseological uniformity" and standardisation, whereas the translations carried out at the Court of Justice reveal phraseological patterns that are more regular.

However, the following example might represent an exception to the outlined trend. The analysis of the Introduction sections of the translated judgments of the Court of Justice showed an interesting example. In the macrostructure of the judgment, after the statement "ha pronunciato la seguente sentenza", a list made up of two points in which the judge explains briefly, in a sort of summary, the remarkable points of the preliminary ruling is presented. The first point explains the subject matter of the reference, mentioning the articles under "interpretation", whereas the second one indicates the context in which the reference has been made.

The introductory part of the two points should be the same in every judgment, due to the fact that it is a routine formula[12]. Text box 4 shows the phraseological patterns identified:

12. Indeed, the French – presumably the source language of these texts – version of the judgment states:
    1. La demande de décision préjudicielle porte sur [l'interprétation des articles]
    2. Cette demande a été présentée dans le cadre de [d'une procédure pénale engagée à l'encontre de]

Text box 4. Routine formulae in the *Court of Justice
of the European Union*'s judgments

---

1. *La domanda di pronuncia pregiudiziale verte su* [15/17]

2. *Tale domana è stata presentata nell'ambito di* [8/17]
Detta domanda è stata presentata nell'ambito di [1/17]
Tale domanda è stata presentata nel contesto di [2/11]

Tale domanda è stata sollevata nell'ambito di [1/17]
La domanda è stata sollevata nell'ambito di [1/17]

Tale domanda è stata proposta nell'ambito di [1/17]

La questione è sorta nell'ambito di [1/17]

Even though this is a clear routine formula, translators of the Court of Justice decided to resort to synonyms: the reference for a preliminary ruling "has been presented in the course of" (11/17), "has been raised in the course of" (2/17), "has been proposed in the course of" (1/17), "has risen in the course of" (1/17). Lexical variation affects not only collocations (verb + noun object), but also adjectives ("tale", "detta") and prepositional phrases ("nell'ambito di", "nel contesto di").

Corpus findings revealed other instances of lexical variation in translated texts. In this context, translators' deliberate choice not to use routine formulae might be interpreted and judged in different ways: from the viewpoint of the jurists, that could be a "bad" strategy, because the use of synonyms may hinder the interpretation of the texts and increase the risk of possible misunderstandings that may have legal consequences.

However, criticisms towards translators' strategies is not the aim of the present study, since the focus is on the objective observation of the data. Translators' different behaviours need to be considered within the institutional framework in which they work: each body has its own "culture" which might justify the existence of "translation cultures", a local approach to translations that is the result of a number of factors: function of the target text, future reader of the translation, etc. As Koskinen (2009) points out, translators adapt to their institutional settings. In order to make the shape and profile of the translators discernible, they have to be seen against the background of the institutional and material context of their work. It is this particular institutional context that moulds the translation processes and it is in this environment that they negotiate their role and professional identity. Text documents, according to Koskinen, should be perceived as being artifacts of a particular culture and interpreted against the background of the cultural framework where they were originated (Koskinen 2009: 6).

3.2.2. Lexical collocations: the Noun (subject) + Verb pattern

Lexical collocations[13] are an example of "associated patterns" (Biber et al. 1998) and are based on Sinclair's "idiom principle" or co-selection, i.e. on the fact that users have available to them a large number of semi-preconstructed phrases (cf. Sinclair 1991).

Among the different typologies of phraseologisms, collocations are perhaps the easiest category to recognise, although the term "collocation" is used in literature with a large number of different meanings (Granger & Paquot 2008: 35).

Combining research in general language phraseology and evidence from phraseology in terminological databases, Musacchio (2002: 143-146) identifies the following patterns (the examples in brackets are taken from the comparable corpus):

- *Adjective + noun or noun phrase* (e.g. apparente motivazione, motivazione logica, dichiarazioni mendaci, etc.)
- *Noun or noun phrase as subject + verb or verb phrase* (e.g. il difensore oppone, il difensore lamenta, la Corte ritiene, la Corte rileva, la Corte osserva, etc.)
- *Verb or verb phrase + noun or noun phrase as object* (e.g. commettere un reato, adottare un provvedimento, accogliere il ricorso, rigettare il ricorso, etc.)
- *Noun or noun phrase + preposition + noun or noun phrase* (e.g. ricevibilità del ricorso, merito del caso, assunzione della prova, etc.)
- *Verb or verb phrase + preposition + noun or noun phrase* (e.g. concludere per l'annullamento, conoscere di un reato, etc.)
- *Verb or verb phrase + adverb or adverb phrase* (e.g. esaminare congiuntamente [la ricevibilità e il merito di un caso], procedere penalmente, etc.)
- *Adverb or adverb phrase + adjective* (e.g. manifestatamente infondato, parzialmente ricevibile, etc.)

In this section of the paper attention will be focused on the second type of collocations identified by Musacchio (noun SUBJECT + verb) which will be exemplified by the analysis of the collocational network of the term "*ricorrente*" (applicant) that plays a crucial role in both original judgments and the translated ones, especially those delivered by the European Court of Human Rights.

---

13. "(Lexical) collocations are usage-determined or preferred syntagmatic relations between two lexemes in a specific syntactic pattern. Both lexemes make an isolable semantic contribution to the word combination but they do not have the same status. Semantically autonomous, the 'base' of a collocation is selected first by a language user for its independent meaning. The second element, i.e. the 'collocate' or 'collocator', is selected by and semantically dependent on the 'base'. Examples: *heavy rain*, *closely linked*, *apologise profusely*" (Granger & Paquot 2008: 43).

Table 4 shows the relative frequency and the total occurrence of the term – base of the collocation – "ricorrente" in the two sub-corpora:

Table 4. Frequency of the term *ricorrente** in the comparable corpus

Word/Term	Original Italian	Translated Italian	
ricorrente*	RF: 501 [0.20%]	RF: 1,327 [0.55%]	
	INT: 0.14% [10] FAC: 0.20% [454] DEC: 1.03% [37]	INT: [2] FAC: [1] DEC: [Ø]	INT: 0.98% [65] FAC: 0.86% [1,224] DEC: 1.22% [36]

It is above all in the Facts section of the judgment that the term "ricorrente" occurs because it is the section that contains the arguments presented by the applicants, their points of view and personal stance concerning the case[14] (Bhatia 1993: 130-131). Hence, the interest in studying the collocation noun (noun – *ricorrente* – as subject) + verb (usually verba *dicendi* or *sentiendi*, i.e. verbs of thinking and opinions), a bundle that reports personal states in the third person.

Tables 5 and 6 show the *ricorrente* (subject) + verb collocation in the original sub-corpus and in the translated sub-corpus:

Table 5. *Ricorrente* (subject) + verb collocation in the original sub-corpus

Original IT		
Il ricorrente *lamenta*	9	
Eccepisce il ricorrente	4	
Il ricorrente *deduce*	10+4	R: 10; L: 4
Premette il ricorrente	3	
Sostiene il ricorrente	6+2	L: 6; R: 2
Ha rilevato il ricorrente	1	
Assume il ricorrente	3+2	L: 3; R: 2

Table 6. *Ricorrente* (subject) + verb collocation in the translated sub-corpus

Translated IT	
Il ricorrente *ritiene*	21
Il ricorrente *si lamenta*	12
Il ricorrente *lamenta*	39

(Cont.)

---

14. As far as the sub-corpus of original judgments is concerned, the term occurs in the Introduction mainly in association with "difensori" (counsels for the defence), while the judge presents the "characters" of the case, whereas in the Decision the reference to the applicants is connected with their conviction ([La Corte] rigetta il ricorso e condanna il ricorrente…].

Translated IT	
Il ricorrente *sostiene*	12
Il ricorrente *ha affermato*	5
Il ricorrente *contesta*	4
Il ricorrente *ammette*	4
Il ricorrente *adduce*	4
Il ricorrente *ha contestato*	2
Il ricorrente *si duole*	2
Il ricorrente *fa presente*	2

The tables show the verbs that most frequently collocate with the noun "ricorrente": to complain, to argue, to contend, to submit, to assert, to request, to claim, to allege, to reckon, to opine, to object, to point out, to notice, etc, i.e. verbs that Rovere (2000) defines as "illocutivi espositivi" (having an illocutive and expositive function).

An interesting stylistic feature of original judgments which can be detected from the corpus findings (see Table 5) is the systematic subject-verb inversion (cf. Rega 1997: 122-123; Scarpa & Riley 2000: 258; Rovere 2000; Mortara Garavelli 2000; Santulli 2008: 222-226). Indeed, the postposition of the subject is marked as a stylistic characteristic of legal texts, especially judicial texts. Such a tendency is usually associated with other marked word orders, such as the preposition of the adjective or the adverb (*detta decisione*, *medesimi fatti*, *fondato motivo*, *appellata sentenza*, etc.; *correttamente osservando*, *fondatamente sostenere*, etc.). The verb + subject pattern produces a "presentational effect" (Rovere 2000: 30) aimed at underlining the formal and solemn character of the "appearance on the scene", in this case the intervention of the applicant. Table 7 shows some concordances of the term *ricorrente* in the original judgments:

Table 7. A concordance of the term *ricorrente* in the original sub-corpus

LH co(n)text	Term	RH co(n)text
n. 69, art. 6, comma 4, lett. a). *Eccepisce* il	ricorrente	che al momento dell'arresto non era stata allegata - così come, invece,
artt. 6 e 16. *Deduce* il	ricorrente	che il D. era stato condannato in con-tumacia e
art. 18, lett. p). *Premette* il	ricorrente	che ai sensi della L. n. 69 del 2005, art. 18, lett. p) la Corte di Appello deve
Stazzanti, rv 233454). 4. *Lamenta*, infine, il	ricorrente	che non sussisterebbero i presupposti per l'emissione della misura cautelare

As shown in Table 7, the subject-verb inversion has also the function of assigning to the verb a thematic position in the informative structure of the sentence

(in Table 7: the appellant objects, opines, states beforehand, complains). The reasons adduced by the appellants are reported at the beginning of the reasoning of the judge, giving communicative salience to the legal significant element (the verb). Obviously, there is no trace at all of such inversion in the translated judgments: translating a judgment into Italian does not mean reproducing every stylistic feature according to the discourse conventions of the specialised register, including unnecessary ones. In this case, translators opted for the unmarked word order [S+V] instead of the inverted word order frequently employed in the original texts [V+S].

3.2.3. Prepositional phrases

A large number of lexical bundles in the corpus analysed consists of prepositional phrases that are commonly used in legal language, especially in judgments.

If collocations are relatively easy to classify, this group of phraseologisms creates certain difficulties of categorisation. According to Granger & Paquot (2008: 44), these are "*complex prepositions*", belonging to the category: "textual phrasemes", that is to say "grammaticalised combinations of two simple prepositions with an intervening noun, adverb or adjective (e.g. with respect to, in addition to, apart from, irrespective of, etc.)". Biber et al. (1999: 1017) define these phraseological units as "prepositional phrases with or without embedded of-phrase fragment"[15], functioning as post-modifier of the noun. Corpas Pastor (1996: 105-106) defines them as "*locuciones prepositivas*"[16], the same term adopted by Serianni (2003: 116-117): "locuzioni preposizionali". According to Serianni, these phraseologisms are a stylistic feature of legal language, since they are preferred to the simple one-word preposition most commonly used in standard Italian; e.g. "*a carico di*" instead of "*contro*" (against), "*a seguito di*" vs. "*per, dopo*" (after), "*a titolo di*" vs. "*per, come*" (as), "*ai fini di*" vs. "*per*" (for), "*ai sensi di*" vs. "*per, secondo*" (in accordance with), "*in danno di*" vs. "*su, contro*" (against), "*in ordine a*" vs. "*su*" (about), "*per via di*" vs. "*per, con, attraverso*" (for).

---

15. As for the prepositional phrases with embedded of-phrase fragment, the scholars distinguish "four-word bundles" (e.g. as a result of, at the level of, in view of the, with the exception of, etc.), "five-word bundles" (e.g. as in the case of, from the point of view, in the context of the, etc.), "six-word bundles" (e.g. from the point of view of). As for the prepositional phrases without embedded of-phrase fragment, they identify "four-word bundles" (e.g. as in the case, by the fact that, on the grounds that, with respect to the, etc.), "five-word bundles" (e.g. in such a way that, on the one hand and, etc.), "six-word bundles" (e.g. in such a way as to).

16. "Las locuciones prepositivas constituyen el núcleo de un sintagma prepositivo. Salvo el hecho de no formar sintagmas por sí mismas, satisfacen los demás criterios identificativos (ej. español: gracias a, delante de, con vistas a, en pos de; ej. inglés: by force of, as from, due to, instead of)" (Corpas Pastor 2003: 141).

In the comparable corpus, prepositional phrases are present in each section of the judgment, in both original and translated documents.

Table 8 shows the most important prepositional phrases detected in both sub-corpora, together with their relative frequency (asterisks indicate very rare frequencies: one or two single occurrences).

Table 8. Prepositional phrases in the comparable corpus

	ORIGINAL IT	**TRANSLATED IT**
INTRODUCTION	*in persona del* [0.49] *nella persona del* [0.10] *ai sensi di* [0.07]	*in qualità di* [0.99] *a carico di* [0.07] *dinanzi a* [0.05]
FACTS	*in ordine a* [0.09] *in relazione a* [0.12] *in punto di* [0.01] *in tema di* [0.02] *con riferimento a* [0.04] *in riferimento a* [0.02] *a carico di* [0.008] *ai sensi di* [0.06]	*in caso di* [0.03] *nell'ambito di* [0.06] *nel contesto di* [0.004] *ai sensi di* [0.16] *in conformità a* [0.07] *in conformità di* [0.006] *in conformità con* [0.003] *senza pregiudizio di* [1*] *senza pregiudizio per* [2*] *in relazione con* [0.003] *con riferimento a* [0.01] *in riferimento a* [0.004] *in ordine a* [0.003]
DECISION	*ai sensi di* [0.25] *per gli adempimenti di cui a* [0.44] *nei confronti di* [0.44]	*alla luce di* [1*] *ai sensi di* [0.31] *in conformità con* [1*] *a norma di* [0.04] *a titolo di* [0.23] *pari a* [0.17] *in relazione a* [2*] *in ragione di* [1*]

The prepositional phrases detected in the translated sub-corpus are characterised by a certain degree of variability (e.g. "con/in riferimento a"; "in relazione a/con"; "in conformità a/di/con"), whereas those employed in the original sub-corpus exhibit more

phraseological uniformity. Phraseological variation seems to be greater in translated texts, which could be in line with the tendency identified in 3.1.

Moreover, the analysis of this third type of phraseologism, can provide interesting insights into the "reasoning of the judge". The Decision section of the original judgments, for example, contains a high percentage (0.44%) of the prepositional phrase "nei confronti di" (against), in line with the move in question (pronouncing of the judgment): it is in this section that the decision is taken "against" the defendants/appellants and stated in its brief, formulaic and highly standardised form. In the Decision section of the translated judgments the focus seems to be different, that is to say the supranational Courts tend to stress the fact that their decisions comply with the existing laws ("alla luce di" – in view of, "ai sensi di" – within the meaning of, "a titolo di" – in respect of, etc.).

Evidently, these are only preliminary results: not only do they need to be analysed more exhaustively, but they also need to be classified and categorised according to uniform and coherent criteria, which will be assessed according to the objectives of the final research project.

4. Lessons learned and future steps

The aim of this final paragraph of the paper is to take stock of the pilot study, identifying the main challenges ahead.

One of the first aspects that will be addressed in the Ph.D. research relates to the vexed question of objectivity in phraseology extraction. Patterns may be easy to identify typologically, but tracing them in texts using objective criteria for selection is a complex activity (Musacchio & Palumbo 2008: 72), mainly because, most of the time, frequency is not enough. A qualitative approach to the selection of candidate phraseologisms is a crucial aspect. To increase the level of objectivity in phraseology mining, a threshold of occurrence (Biber et al. 1999: 992-993) will be set in the Ph.D. multilingual comparable corpus. To qualify as such a phraseologism, a word combination must frequently recur in a register. As suggested by Biber et al., lexical sequences should be counted as "recurrent" only if they occur at least a number of times per million words in a register (Biber et al.'s threshold is 10 times per million words). Moreover, these occurrences must be spread across at least a certain number of different texts in the register (5 in Biber et al.'s corpus analysis), to exclude individual writer idiosyncrasies.

Another challenge for the future research project concerns the categorisation of the judgments of the multilingual comparable corpus. This pilot study confirmed that the structural division into three moves (Introduction, Facts and Decision) is very useful for the purposes of the study, because it represents the bridge between the corpus and the genre perspectives. In addition to this, other variables will be included: delivering bodies, section of the Court issuing the judicial text, year (to account for potential diachronic variations), city (for geographical variations), etc.

Corpus analysis is undoubtedly a time-consuming activity, but it can be extremely rewarding and fascinating, mainly because of the inevitable element of serendipity that is implicit in every corpus study. A certain phraseologism found in a concordance could suggest a new corpus search, which sometimes reveals a phraseological pattern that had not been anticipated, so that the researcher can start following a new and unbeaten path.

Bibliography

A.A. V.V. (1988). *La sentenza in Europa. Metodo, tecnica e stile*. Padova: Cedam.

Alcaraz Varó, Enrique and Hughes, Brian (2002). *El español jurídico*. Barcelona: Ariel.

Bachmann, Verena (2001). *Le formule standard nelle istanze penali di primo grado in Italia, Germania e Austria: un'analisi contrastiva*. Unpublished MA thesis, SSLMIT, University of Trieste.

Baker, Mona (1993). Corpus Linguistics and Translation Studies. Implications and Applications. In *Text and Technology. In Honour of John Sinclair*, Mona Baker, Gill Francis and Elena Tognini-Bonelli (eds.), 233-250. Amsterdam-Philadelphia: John Benjamins.

Baker, Paul; Hardie, Andrew and McEnery, Tony (2006). *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Bhatia, Vijay K. (1993). *Analysing Genre. Language Use in Professional Settings*. London and New York: Longman.

Biber, Douglas; Conrad, Susan; Finegan, Edward; Johansson, Stig and Leech, Geoffrey (1999). *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education Limited.

Borja Albi, Anabel; Gil Sayas, Marta; Juste, Natividad and Martínez-Carrasco, Roberto (forthcoming). Gestión electrónica del conocimiento para la traducción en entornos judiciales: el proyecto JUD-GENTT. In *Proceedings of the IV International Conference on Public Service Interpreting and Translation*, Universidad Complutense de Alcalá, 2011.

Burger, Harald (1998). *Phraseologie. Eine Einführung am Beispiel des Deutschen*. Berlin: Erich Schmidt.

Capotorti, Francesco (1988). Le sentenze della Corte di Giustizia delle Comunità Europee. In *La sentenza in Europa. Metodo, tecnica e stile*, A.A. V.V., 230-247. Padova: Cedam.

Chesterman, Andrew (2004). Beyond The Particular. In *Translation Universals. Do They Exist?*, Anna Mauranen and Pekka Kujamäki (eds.), 33-49. Amsterdam-Philadelphia: John Benjamins.

Colson, Jean-Paul (2008). Cross-linguistic Phraseological Studies: An Overview. In *Phraseology: An Interdisciplinary Perspective*, Sylviane Granger and Fanny Meunier (eds.), 191-206. Amsterdam-Philadelphia: John Benjamins.

Corpas Pastor, Gloria (1996). *Manual de fraseología española*. Madrid: Gredos.

— (2003). *Diez años de investigación en fraseología: análisis sintáctico-semánticos, contrastivos y traductológicos*. Vervuert: Iberoamericana.

Cosmai, Domenico (2007). *Tradurre per l'Unione Europea*. Seconda edizione. Milano: Hoepli.

Cowie, Anthony Paul (ed.) (1998). *Phraseology. Theory, Analysis and Applications*. Oxford: Clarendon Press.

Garzone, Giuliana (2007). Osservazioni sulla didattica della traduzione giuridica. In *Tradurre le microlingue scientifico-professionali. Riflessioni teoriche e proposte didattiche*, Paola Mazzotta and Laura Salmon (eds.), 194-238. Torino: UTET.

Gouadec, Daniel (2007). *Translation as a Profession*. Amsterdam-Philadelphia: John Benjamins.

Goźdź-Roszkowski, Stanislaw (2006). Recurrent Word Combinations in Judicial Argumentation. A Corpus-Based Study. In *Langue, Droit, Société. Language, Law, Society*, Danuta Bartol, Anna Duszak, Hubert Izdebski and Jean-Marie Pierrel (eds.), 139-152. Nancy: Presses universitaires.

Granger, Sylviane (2005). Pushing Back The Limits of Phraseology: How Far Can We Go? In *Phraseology 2005: The Many Faces of Phraseology*, Christelle Cosme, Céline Gouverneur, Fanny Meunier and Magali Paquot, (eds), 165-168. Louvain-la-Neuve: CECL.

— and Meunier, Fanny (eds.) (2008). In *Phraseology: An Interdisciplinary Perspective*. Amsterdam-Philadelphia: John Benjamins.

— and Paquot, Magali (2008). Disentangling The Phraseological Web. In *Phraseology: An Interdisciplinary Perspective*, Sylviane Granger and Fanny Meunier (eds.), 27-49. Amsterdam-Philadelphia: John Benjamins.

Gries, Stefan Th. (2008). Phraseology and Linguistic Theory: A Brief Survey. In *Phraseology: An Interdisciplinary Perspective*, Sylviane Granger and Fanny Meunier (eds.), 3-25. Amsterdam-Philadelphia: John Benjamins.

Koskinen, Kaisa (2008). *Translating Institutions. An Ethnographic Study of EU Translation*. Manchester-Kinderhook: St. Jerome Publishing.

Laviosa, Sara (2002). *Corpus-based Translation Studies. Theory, Findings, Applications*. Amsterdam-New York: Rodopi.

Lombardi, Alessandra (2004). *Collocazioni e linguaggio giuridico. Proposte per un'analisi semi-automatica delle unità complesse in testi del diritto penale italiano e tedesco*. Milan: ISU Università Cattolica.

Mel'čuk, Igor (1998). Collocations and Lexical Function. In *Phraseology. Theory, Analysis and Applications*, Anthony Paul Cowie (ed.), 25-53. Oxford: Clarendon Press.

Mortara Garavelli, Bice (2001). *Le parole e la giustizia. Divagazioni grammaticali e retoriche su testi giuridici italiani*. Torino: Einaudi.

Musacchio, Maria Teresa (2002). I tecnicismi collaterali. In *Manuale di Terminologia. Aspetti teorici, metodologici e applicativi*, Marella Magris, Maria Teresa Musacchio, Lorenza Rega and Federica Scarpa (eds.), 135-150. Milan: Hoepli.

— (2008). Universal or Culture-specific Metaphors in Economics? A Corpus Study of Original vs Translated Italian. In *LREC 2008: Proceedings of the 6th International Conference on Language Resources and Evaluation. Marrakech, Morocco, 26 May – 1 June*, 62-69. European Language Resources Association: Paris.

— and Palumbo, Giuseppe (2008). Shades of Grey: A Corpus-driven Analysis of LSP Phraseology for Translation Purposes. In *Corpora for University Language Teachers*, Carol Taylor Torsello, Katherine Ackerley and Erik Castello (eds.), 69-79. Bern: Peter Lang.

Ondelli, Stefano (2006). Il genere testuale della sentenza penale in Italia. In *Studi in ricordo di Carmen Sánchez Montero, vol. 1*, Graziano Benelli and Giampaolo Tonini (eds.), 295-309. Trieste: EUT.

Rega, Lorenza (1997). La sentenza italiana e tedesca nell'ottica della traduzione. In *La lingua del diritto. Difficoltà traduttive. Applicazioni didattiche*. Atti del primo convegno internazionale, Milano, 5-6 ottobre 1995, Centro Linguistico dell'Università Bocconi, 117-126. Roma: CISU.

— (2000). Aspetti e problemi della traduzione delle formule di rito nell'ambito giuridico italo-tedesco. In *Linguistica giuridica italiana e tedesca/Rechtslinguistik des Deutschen und Italienischen*, Daniela Veronesi (ed.), 449-457. Padova: Unipress.

Rojo, Ana (2009). *Step by Step. A Course in Contrastive Linguistics and Translation*. Bern: Peter Lang.

Rovere, Giovanni (2000). L'avverbiale strumentale nel linguaggio giuridico. In *Traduttori e giuristi a confronto. Interpretazione traducente e comparazione del discorso giuridico*, vol. I, Leandro Schena and Rita D. Snel Trampus (eds.), 25-35. Bologna: CLUEB.

Santulli, Francesca (2008). Lingua giuridica e generi testuali. In *Il linguaggio giuridico. Prospettive interdisciplinari*. Giuliana Garzone and Francesca Santulli (eds.), 207-238. Milan: Giuffrè.

Scarpa, Federica and Riley, Alison (2000). La fisionomia della sentenza in Inghilterra e in Italia: un'analisi orientata alla traduzione. In *Traduttori e giuristi a confronto. Interpretazione traducente e comparazione del discorso giuridico*, vol. I, Leandro Schena and Rita D. Snel Trampus (eds.), 227-291. Bologna: CLUEB.

Serianni, Luca (2003). *Italiani scritti*. Bologna: Il Mulino.

Sinclair, John (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Swales, John M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.

Taruffo, Michele (1988). La fisionomia della sentenza in Italia. In *La sentenza in Europa. Metodo, tecnica e stile*, A.A. V.V., 180-214. Padova: Cedam.

van Teijlingen, Edwin R. and Hundley, Vanora (2001). The Importance of Pilot Studies. In *Social Research Update* [on line], University of Surrey, http://sru.soc.surrey.ac.uk/SRU35.html [Accessed April 2011]

Wright, Sue Ellen (1997). Term Selection: The Initial Phase of Terminology Management. In *Handbook of Terminology Management*. Sue Ellen Wright and Gerhard Budin (eds.), 13-24. Amsterdam-Philadelphia: John Benjamins.

Online Resources

De-Jure database (available under subscription):

http://dejure.giuffre.it/psixsite/PaginePubbliche/default.aspx [Accessed April 2011]

CURIA: Court of Justice of the European Union's case-law:

http://curia.europa.eu/jurisp/cgi-bin/form.pl?lang=it [Accessed April 2011]

HUDOC database – European Court of Human Rights Document Collections:

http://cmiskp.echr.coe.int/tkp197/search.asp [Accessed April 2011]

Osservatorio della Giurisprudenza CEDU – Observatory of the case-law of the European Court of Human Rights (in Italian):

http://www.osservatoriocedu.it/ [Accessed April 2011]