This chapter describes the results of an empirical study of LSP phraseological units in a specific domain (criminal law) and type of legal genre (criminal judgments). The final goal of the research is to provide legal translators with a multifunctional resource having a positive impact on the translation process and product. More specifically, it aims at assisting translators – as well as legal experts – to develop their phraseological competence through exposure to real, authentic (con)texts in which these phraseological units are used. Based on cospe, a 6-million trilingual, comparable corpus of criminal judgments, this study approaches phraseology from a contrastive (Spanish-Italian-English), quantitative and qualitative perspective. Corpus analysis and term extraction have been carried out by means of concordancers (mainly WordSmith Tools v. 5.0). From a methodological point of view, the study combines corpus-based and corpus-driven approaches, as well as traditional approaches applied to Language for General Purposes (LGP) phraseology, and more recent distributional studies of Language for Specific Purposes (LSP) and legal phraseology. Emphasis is placed on four categories of phraseological units frequently found in judicial discourse: complex prepositions, lexical doublets and triplets, lexical collocations and routine formulae.

1 Introduction

Legal translation is not only a question of terminology – which is indeed one of the major obstacles legal translators have to face in their daily activity – but also a question of phraseological conventions. Beyond lexical and terminological equivalence, translators have to tackle the additional difficulty of acquiring
familiarity with the genre structures – or “generic” structures in Hasan’s (1978) terms – through which legal institutions conduct their affairs. Hatim & Mason (1990: 190) use the term “routines” to describe “those conventions which translators either know or simply do not know: frozen patterns of a formulaic nature which are typical of legal texts and which can be translated only resorting to parallel routines in the target language”. As a matter of fact, even the most skilled translator may run the risk of producing a translation that is inaccurate from the standpoint of the “register choices”, all other aspects of the target text being perfectly acceptable (grammar, content, etc.) (see Garzone 2007: 218–219).

The current studies of phraseology in specialised registers acknowledge the need for corpus-based studies of the prototypical lexico-grammatical patternings and discourse functions of lexical phrases across disciplines. Gaining control of a new language or register requires, following Hyland (2008: 5), a sensitivity to expert users’ preferences for certain sequences of words over others that might seem equally possible.

In line with these preliminary remarks, the study stems from three main considerations:

1. Judgments represent a fertile ground for the study of phraseology: the frequent use of phraseological units is one of the most striking features of these judicial texts, a real “trademark” of legal texts (Mortara Garavelli 2001: 154);

2. Phraseology is one of the main obstacles legal translators have to tackle in their professional activity (see Garzone 2007; Kjær 2007);

3. Confronted with the task of translating legal texts, professional translators have few phraseological resources at their disposal.

The relation between “phraseology”, “judicial texts” and “translation” appears to have been scarcely investigated so far. The research has represented a first, tentative step towards filling such gap.¹

¹ This chapter is based on a PhD research project conducted on specialised phraseologies employed in criminal judgments (Pontrandolfo 2013b). The PhD thesis entitled “La fraseología en las sentencias penales: un estudio contrastivo español, italiano, inglés basado en corpus” (Supervisor: Helena Lozano Miralles; Co-supervisors: Emilio Ortega Arjonilla, Mitja Gialuz) was defended by the author on 12/04/2013 at the University of Trieste within the XXV PhD cycle in Interpreting and Translation Studies (coordinator: Federica Scarpa). The contribution is also based on a conference paper given during the 19th European Symposium on Languages for Special Purposes, 8–10 July 2013 held at the University of Vienna. It is part of the research project.
2 Theoretical background

Phraseology in legal and judicial language is a rather unexplored field of study. The following literature survey of the relevant sources, concepts and definitions is structured into three main parts:

1. Phraseology. A number of influential classifications of phraseological units have been analysed in LGP (from the seminal work of Benson, Benson & Ilson 1997; Corpas Pastor 1996; Gläser 1994/1995; 1998; Ruiz Gurillo 1997; Cowie 1988; 2001; Mel’čuk 1998; Moon 1998; Burger 1998; Granger & Paquot 2008), LSP (L’Homme 2000; Lorente 2001; Tercedor Sánchez 1999; Montero Martínez 2002; Bevilacqua 2004; Aguado de Cea 2007) and legal phraseology (Kjær 1990a; 1990b);

2. Corpora for the study of legal and judicial language (see Pontrandolfo 2012). After presenting the main definitions and concepts, the review focuses on the main corpora built in Spain (e.g. JUD-GENTT, the IULA’s CORPUS, CLUVI), Italy (e.g. BoLC, CORIS/CODIS, CADIS), England and Wales (e.g. Cambridge Corpus of Legal English, HOLJ Corpus, Proceedings of the Old Bailey), as well as in the European Union and the rest of the world;

3. Studies carried out by researchers from different areas and schools dealing with the topic of the present research, as the subject or a side aspect of their investigations.

As far as the third part is concerned, research in this area can be classified into four different subareas, according to the methodological approach, the types of phraseological units investigated as well as the focus of the analysis:

a) Studies that analyse lexico-syntactic combinations in legal language, with a preference for specialized collocations, based on the traditional notion of phraseology (Benson, Benson & Ilson 1997; Hausmann 1989; Corpas Pastor 1996; Berdychowska 1999; Nardón-Schmid 2002; Lombardi 2004; Rovere 1999; Nystedt 2000; Martínez & Soledad 2002; Giráldez Ceballos-Escalera 2007; Anderson 2006; Assunção & Raquel 2007; Biel 2011; Bhatia 2004);
b) Studies that focus on the formulaic nature of legal language in terms of routine formulae (Rega 2000; Bachmann 2000; Monzó 2001; Carvalho Fonseca 2007; Giurizzatto 2008);

c) Lexicographic studies aimed at building specialised legal dictionaries (De Groot 1999; François & Grass 1997; Gisbert & Joaquina 2008; Fernández Bello 2008);

d) Studies that adopt a wider notion of phraseology and are based on large corpora of legal texts aimed at analysing co-occurrence patterns (Mazzi 2005; 2010; Goźdź-Roszkowski 2011).

The survey highlighted a significant gap in the literature on translation-oriented studies of the phraseological nature of legal or judicial discourse, as there are still only few studies dealing with the role of phraseology in judicial discourse from an empirical, contrastive, corpus-based perspective.

3 Aim, scope and objectives of the research

The study deals with the complex universe of phraseology, in its broader sense (see Gries 2008: 6), from a contrastive (Spanish–Italian–English), quantitative and qualitative perspective. Emphasis has been placed on a specific genre, criminal judgments, i.e. “courts’ final determination of the rights and obligations of the parties in a case” (see Bryan 2009: 918).

Judgments epitomise the nature of judicial discourse, as they are the most important acts in criminal trials, and represent one of the most striking examples of “living law” or “law in action” to refer to the 1910’s pioneer paper by the distinguished legal scholar Roscoe Pound (see Garavelli 2010: 154; Cadoppi 1999: 253). Studying judgments means exploring the language of the discourse community composed by judges. Narrowing down the huge normative subjects the courts are asked to rule on has allowed to focus on a coherent and consistent share of case-law. Furthermore, this has left room for a long-term study that could delve into the correlation between a specific field of law (e.g. civil, labour law, etc.) and the type of phraseological patterns used by legal experts.

The research questions lying at the basis of the study can be summarised as follows:

1. Phraseology is a key stylistic feature of criminal judgments, a real “trademark” in judges’ writing conventions. What is the quantitative and qualitative relevance of this typological trait in criminal judgments?
2. Due to the different legal traditions (common law vs. civil law) characterising the three cultures involved in the present study (Spain, Italy, and England and Wales), does the weight of phraseological units change depending on the respective source country?

3. Once a selected number of phraseologisms have been extracted, will it be possible to establish a comparability between them?

The main goal of this empirical study of specialised phraseological units in a specific type of legal genre, i.e. criminal judgments, is that of providing legal translators dealing with criminal procedure with a multifunctional resource having a positive impact on the translation process and product. More specifically, it aims at assisting legal translators (as well as legal experts) in developing phraseological competence, guiding them to achieve “naturalness” in writing through exposure to real, authentic (con)texts in which phraseological units are used.

4 Material

In order to answer the research questions, a trilingual, comparable corpus of judicial texts has been built, i.e. the Corpus of Criminal Judgments (CORpus de Sentencias PENales, COSPE). The focus has been placed on a single genre (criminal judgments) for a number of reasons (see Pontrandolfo 2013b: 171–181). Among them, the importance of this specific genre in judicial discourse – see the importance of the judicial “precedent” in the common-law as well as civil-law traditions – and, from a practical point of view, the need to find a shared ground across legal cultures to allow for a full comparative analysis.

As Hunston (2008: 156–157) put it, “all corpora are a compromise between what is desirable, that is, what the corpus designer has planned, and what is possible”. Table 1 shows the result of a number of strategic decisions which have been taken and challenges which have been tackled to compile a balanced legal corpus. As shown in Table 1, COSPE is made of two subcorpora: COSPE-Sup which gathers 380 criminal judgments delivered between 2005 and 2012 by the Supreme Courts (courts of last instance) in the three judicial systems, and COSPE-Ap which contains 402 criminal judgments delivered in the same period by various courts of appeal (courts of second instance) in Spain, Italy, and England and Wales. These courts have been chosen for being comparable in terms of role and functions.

With a view to obtaining a representative sample of the genre (see Biber 1993: 243), a number of variables have been established to guarantee heterogeneity and balance in the process of storing and categorising the judgments, as well as
Table 1: Composition of cospe (Corpus of Criminal Judgments)

<table>
<thead>
<tr>
<th>Type of corpus:</th>
<th>Trilingual, comparable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Languages:</td>
<td>ES-IT-EN</td>
</tr>
<tr>
<td>Size:</td>
<td>Tot. 782 txt (6,036,915 tokens)</td>
</tr>
<tr>
<td>Genre:</td>
<td>Criminal Judgments</td>
</tr>
<tr>
<td>Period:</td>
<td>2005 - 2012</td>
</tr>
<tr>
<td>Purposes:</td>
<td>Pratice, Research, Training</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ES</th>
<th>Court</th>
<th>txt</th>
<th>tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>COSPE-Sup</td>
<td>Tribunal Supremo</td>
<td>100</td>
<td>1,088,770</td>
</tr>
<tr>
<td>COSPE-Ap</td>
<td>Audiencia Provincial</td>
<td>127</td>
<td>722,177</td>
</tr>
<tr>
<td></td>
<td>Tribunal Superior de Justicia</td>
<td>35</td>
<td>208,619</td>
</tr>
<tr>
<td></td>
<td>tot</td>
<td>162</td>
<td>930,796</td>
</tr>
<tr>
<td>COSPE-ES (tot.)</td>
<td></td>
<td>262</td>
<td>2,019,566</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IT</th>
<th>Court</th>
<th>txt</th>
<th>tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>COSPE-Sup</td>
<td>Corte Suprema di Cassazione</td>
<td>230</td>
<td>1,014,224</td>
</tr>
<tr>
<td>COSPE-Ap</td>
<td>Corte d’Appello</td>
<td>95</td>
<td>357,057</td>
</tr>
<tr>
<td></td>
<td>Corte d’Assise d’Appello</td>
<td>40</td>
<td>629,905</td>
</tr>
<tr>
<td></td>
<td>tot</td>
<td>135</td>
<td>968,962</td>
</tr>
<tr>
<td>COSPE-IT (tot.)</td>
<td></td>
<td>365</td>
<td>2,001,186</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EN</th>
<th>Court</th>
<th>txt</th>
<th>tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>COSPE-Sup</td>
<td>Supreme Court</td>
<td>20</td>
<td>428,529</td>
</tr>
<tr>
<td>COSPE-Sup</td>
<td>House of Lords</td>
<td>30</td>
<td>455,468</td>
</tr>
<tr>
<td></td>
<td>tot</td>
<td>50</td>
<td>883,997</td>
</tr>
<tr>
<td>COSPE-Ap</td>
<td>Court of Appeal</td>
<td>105</td>
<td>1,132,166</td>
</tr>
<tr>
<td>COSPE-EN (tot.)</td>
<td></td>
<td>155</td>
<td>2,016,163</td>
</tr>
</tbody>
</table>
to ease its consultation and queries: ID number, division of the court, region/city (to guard against diatopic usage), date of the hearing, subject matter (to vary the relationship between phraseology and specialised contents), type of proceedings, reporting judge (to guard against idiosyncratic usage), notes (e.g. outcome of the appeal, final decision of the court, etc.).

Following Zanettin (2012: 105–107), cospe is a “translation-driven corpus” in that it has been created with applied (translation) purposes in mind and it does not include translated texts, but texts produced in the three languages under similar circumstances and within the same domain. It is also a “web corpus” (“corpus virtual” according to Corpas Pastor 2004: 227) in that all the texts have been collected from the web (from cendoj, DeJure, Bailii databases) and were therefore already available in electronic format. cospe is currently being POS-tagged.

The corpus has represented the test bed for the investigation based on the research questions which have been tackled adopting a corpus-based methodology.

5 Methodology

The study is a descriptive, empirical research which has fully adopted the corpus linguistics paradigm (see McEnery, Xiao & Tono 2006). Phraseology in criminal judgments has therefore been approached through “real judicial life” examples. Extraction and analysis of relevant phraseologisms have been performed by means of concordancers (mainly WordSmith Tools, but also AntConc and ConcGram).

Querying a corpus of large dimensions like cospe inevitably requires the adoption and integration of different methods, according to the different types of phraseological unit. Methodologies for phraseology extraction vary along a continuum having the manual analysis on one side and the automatic one on the other. Such dichotomy is also reflected in the corpus-driven vs. corpus-based approaches to phraseology, or, to put it in Granger’s (2005:3) terms, between the bottom-up approach/corpus-driven (an inductive approach generates a wide range of word combinations, which do not all fit predefined linguistic categories) and the top-down approach/corpus-based (which identifies phraseological units on the basis of linguistic criteria).

Table 2 shows the methodological moves adopted to extract the phraseological units around the four types object of the investigation, along the continuum corpus-based vs. corpus-driven.
Table 2: Methods of extraction along the continuum corpus-based vs. corpus-driven

<table>
<thead>
<tr>
<th>+ corpus-based</th>
<th>+ corpus-driven</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Complex prepositions</strong></td>
<td><strong>Lexical doublets/triplets</strong></td>
</tr>
<tr>
<td>Semi-manual extraction (e.g. in + * + with)</td>
<td>Semi-automatic extraction (e.g. * + and + *)</td>
</tr>
<tr>
<td><strong>Lexical collocations</strong></td>
<td><strong>Routine/Standardised formulae</strong></td>
</tr>
<tr>
<td><strong>Semi-automatic extraction</strong></td>
<td><strong>Automatic extraction</strong></td>
</tr>
<tr>
<td><strong>Semi-automatic extraction</strong></td>
<td>(ws ConcGram and ConcGram 1.0)</td>
</tr>
</tbody>
</table>

*top-down approach* | *bottom-up approach*

---

*a “A measure of how strongly two words seem to associate in a corpus, based on the independent relative frequency of two words” (Church & Hanks 1990).*

*b To identify the nodes of the collocations an innovative method has been followed based on Schank and Abelson’s notion of “script” – “a structure that describes appropriate sequences of events in a particular context” Schank & Abelson (1977: 141) – adapted to the context of criminal judgments of second or last instance (e.g. During a trial the Court/judge issues a judgment against a person accused of a crime, i.e. a defendant who committed an offence. The appellant contests the court’s decision adducing his/her arguments. The Court can allow or dismiss the appeal (acquitting or convicting him/her), (re)determining the sentence. The judge explains his/her opinion]. Starting from the script, nine key terms have been identified – es: juicio, tribunal/juez, acusado, delito, sentencia, motivo, recurso, apelante/recurrente, pena; it: giudizio, giudice/corte/tribunale, imputato, reato/delitto, sentenza, motivo, ricorso/appello, appellante/ricorrente, pena; en: trial, court/judge, defendant, offence, judgment/decision/opinion, argument, appeal, appellant, sentence – and later scrutinised to discover phraseological patterns.*
To retain methodological rigour, a cut-off point of 5 occurrences per 2,000,000 words has been fixed, combined with the multiple-text requirement whereby a given phraseological unit had to appear in at least 5 different judgments to guard against judges’ idiosyncrasies (Goźdż-Roszkowski 2011: 110). The extraction has yielded a significant numbers of specialised phraseological units which will be dealt with in the following sections.

6 Results

A quantitative and qualitative analysis of the four types of recurrent phraseological units mentioned above has been carried out. The following sections contain a summary of the findings that, for reasons of space, cannot be presented exhaustively in this paper.

6.1 Complex prepositions

Following Biber et al. (1999: 75), “complex prepositions are multi-word sequences that function semantically and syntactically as single preposition”, i.e. “grammaticalised combinations of two simple prepositions with an intervening noun, adverb or adjective” (Granger & Paquot 2008: 44). There can be two types of complex prepositions: N + P (e.g. ES: encima de; IT: innanzi a; EN: owing to) and P + N + P (e.g. ES: con arreglo a; IT: in ordine a; EN: in accordance with).

These phraseological units, especially the second type, are highly frequent in Spanish, Italian and English legal language (see Pontrandolfo 2013a), as can be seen from the examples taken from COSPE:

- **ES**: al amparo de, a juicio de, en aras de, en concepto de, a instancia(s) de, etc.
- **IT**: in relazione a, in ordine a, a titolo di, in conformità a, in deroga a, a pena di, etc.
- **EN**: on behalf of, by reason of, without prejudice to, by virtue of, on the ground(s) of, etc.

Figure 1 shows the quantitative results in terms of total number, total number of patterns (e.g. as + * + as; by + * + of; etc.) vs. number of types (single different combinations such as “with reference to”, “without prejudice to”, etc.).

The Spanish subcorpus (CospES) presents a wider variety of complex prepositions (159 different types generated by 14 patterns), compared with the Italian
one (CospIT) (129 vs. 12) and the English one (CospEN) (84 vs. 17). This seems to suggest that, although the Italian judgments contain the highest number of complex prepositions (21,012), they tend to be much more repetitive than their Spanish and English counterparts. Indeed, the 21,012 complex prepositions are always made of the same patterns (12) which compose 129 different types. CospES displays a lower number of instances (13,854), although the types seem to generate a wider number of phraseological units (159). CospEN shows the lowest number of occurrences (10,343), even though the range of prepositional patterns is much wider (17), but less significant quantitatively (84 different types of complex prepositions stemming from 17 formal structures).

As far as the qualitative analysis of the results is concerned, obviously not all the complex prepositions detected are typical of judicial language. In order to uncover those phraseological units which are used with a certain preference by judges, a comparison between the relative frequency of these patterns in CospE and their frequency in reference corpora (CREA and Corpus del Español for Spanish, Coris/Codis for Italian, BNC for English) has been conducted. “By virtue of”, for example, is used with a raw frequency of 26 in CospEN (normalised frequency of 12.90 per million words), whereas in the BNC it has a frequency of 19 (normalised frequency 0.19 per million words). It is therefore much more used in legal and judicial language. The same applies to “in furtherance of” which has no occurrences in the BNC (vs. 37 instances in 13 different texts in CospEN). The full list of complex prepositions used much more frequently in judicial language can be found in Pontrandolfo (2013a: 200–cd205).
6.2 Lexical doublets and triplets

Following Bhatia (1984: 90), “binomial or multinomial expressions are sequences of two or more words or phrases belonging to the same category having some semantic relationship and joined by some syntactic device such as and or or”. The following examples are all taken from COSPE:

- **ES**: pronunciamos, mandamos y firmamos, [debo] absolver y absuelvo, real y efectivo, natural y vecino, etc.
- **IT**: illogica e contraddittoria, penale e processuale, rigetta e condanna, previsto e punito, connesso e collegato, etc.
- **EN**: adequate and proper, fair and public, reasoning and conclusions, stop and search, etc.

Figure 2 shows the quantitative results of the doublets extracted from COSPE (triplets do not play a crucial role in the genre under investigations).

![Figure 2: Lexical doublets (quantitative findings)]

The analysis showed a proportionality between the total number of instances and the types generated by the 7 patterns identified in the three subcorpora. CosPES contains the highest number of types (202) and tokens (5,059), followed by CosPEN (3,459 distributed over 116 types) and CosPI (1,887 vs. 82).

The quantitative analysis revealed that the most frequent doublets are those made of two nouns (45% in CosPES, 56% in CosPI and 34% in CosPEN) – e.g. violencia e intimidación, contraddittorietà e dilogicità, the prosecution and the defence – followed by the patterns made of two verbs (22% in CosPES, 9% in CosPI and 11% in CosPEN) – e.g. previsto y penado, rappresentato e difeso, aiding and abetting – and those made of two adjectives (16% in CosPES, 16% in CosPI and 25% in CosPEN), e.g. oral y público, connesso e collegato, noble and learned. Lexical
doublets made of prepositions (e.g. unless and until), articles, pronouns (e.g. he or she) and adverbs (e.g. before and during) seem to be used much more frequently in LGP rather than in legal language.

The comparison with the reference corpora confirms that lexical doublets are used much more frequently in judicial language. “Adequate and proper”, for example, is used with a raw frequency of 15 in CospEN (normalised frequency of 7.43 per million words), whereas in the BNC it has a frequency of 3 (normalised frequency 0.03 per million words).

6.3 Lexical collocations

Following Granger & Paquot (2008: 43), “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 language user for its independent meaning. The second element, i.e. the collocate/collocator, is selected by and semantically dependent on the base”.

The analysis conducted on COSPE has been based on nine key terms (see note b) as base of the collocation. As a matter of fact, specialised terminology tends to cluster around terms. Phraseology acts as a link between the term and the text. In particular, the analysis has focused on four types of collocations, exemplified as follows:

1. N [subject] + V
   - ES: valorar una sentencia, carecer un motivo, celebrar un juicio, entender un tribunal, etc.
   - IT: sussistere un reato, ritenere la corte, osservare il giudice, etc.
   - EN: plead guilty/not guilty an appellant, hold the court, conclude the judge, etc.

2. V + N [object]
   - ES: esgrimir un motivo, aducir un motivo, cometer un delito, condenar al acusado, etc.
   - IT: irrigare una pena, proporre un ricorso, accogliere un ricorso, adire la corte, etc.
   - EN: to convict/acquit a defendant, to allow/dismiss an appeal, to await trial, to impose a sentence, etc.
3. **N + ADJ**
   - **es:** motivo impugnatorio, motivo decisorio, pena accesoria, etc.
   - **it:** giudice a quo, imputato contumace, sentenza contraddittoria, etc.
   - **EN:** appropriate sentence, reduced sentence, leading judgment, honest opinion, etc.

4. **N + PREP + N**
   - **es:** celebración del juicio, desestimación del recurso, anulación de la sentencia, etc.
   - **it:** rigetto del ricorso, entità della pena, accoglimento del motivo, etc.
   - **EN:** fairness of the trial, decision of the court, commission of the offence, seriousness of the offence, etc.

Figure 3 shows the frequency of co-occurrence of each collocational pattern, adopting the same categories used for the previous phraseological units (total number of instances vs. patterns).

Overall, the analysis of **COSP** has revealed a balanced picture: **COSPIT** contains the highest number of collocations with a nominal base, followed by **COSPES** and **COSPIT**. The only exception is the **N + V** pattern which shows a higher number of collocations in **COSPEN** (2,084 vs. 112 types).

The most frequent lexical combination in the three subcorpora is **N + ADJ** (**COSPIT**: 4,718, **COSPES**: 3,884, **COSPEN**: 3,160). However, the total number of types is higher in the English subcorpus (125) which seems to point to a higher lexical variation in English and Welsh criminal judgments. Also the **V + N** pattern, where the **N** functions as direct object of the sentence, displays a high number of collocations: 3,774 in **COSPIT**, 3,160 in **COSPES** and 1,202 in **COSPEN**. Yet, a relatively lower number of types of collocations has emerged (**COSPES**: 56, **COSPIT**: 63, **COSPEN**: 43) which could be interpreted as a higher level of repetition or lower lexical variation. Finally, the **N + PREP + N** pattern shows a trend which is similar to **N + ADJ**: **COSPIT** displays the highest number of collocations (3,519 vs. 144 types), followed by **COSPES** (2,717 vs. 102) and **COSPEN** (1,074 vs. 42).

A general trend can be outlined: the English subcorpus contains a lower frequency of lexical collocations, which is indeed an important quantitative result (see §7).

The quality analysis of these phraseological units revealed that lexical collocations play a key role in the genre under analysis and significantly contribute to the “taste” of judicial style which translators have to recreate in their target
texts. The comparison with the reference corpora showed that such phraseological units are much more frequent in judicial language than in general language. This is also due to the presence of the judicial term as node of the collocations extracted (e.g. “appropriate sentence” 12.39 in CospEN vs. 0.12 in the BNC; “fairness of the trial” 15.4 in CospEN vs. 0.04 in the BNC; “allow the appeal” 96.2 in CospEN vs. 1.07 in the BNC). For an in-depth analysis of the qualitative results, see (241–252 Pontrandolfo 2013b).

6.3.1 Routine formulae

Routine formulae or phrases are “recurring lexical sequences, of different length, that develop in the case-law tradition and are usually collected in formularies” (Kjær 1990a: 28–29). The following examples have been extracted from COSPE:
These standardised formulae have been treated combining the insights provided by the genre analysis (Swales 1990, Bhatia 1993). In other terms, routine formulae have been clustered into the five main “moves” of criminal judgments: heading (EN), facts (H), legal background (D), operative part (F), final provisions (DIL). Figure 4 shows the quantitative results.

In general terms, the Spanish criminal judgments present the highest number of routine formulae (1,386), compared to the Italian (740) and English (693) ones. As far as the rhetorical sections (moves) are concerned, the most standardised section of the judgment is the operative part (decision) (CosPE: 674, CosPEN: 399, CosPIT: 324), followed by the heading (CosPIT: 353, CosPEN: 126, CosPE: 53). A high frequency of routine formulae is also found in the facts section of the Spanish judgments (503 instances distributed along 12 types), compared with the Italian (48 vs. 1) and English (18 vs. 1) ones. The legal sections (CosPEN: 117, CosPE: 87 and CosPIT: 4) and the final provisions (CosPE: 70, CosPEN: 33 and CosPIT: 11) are the moves which display the lower number of standardised sequences.

The quality analysis of these phraseological patterns (Pontrandolfo 2013a: 255–261) confirmed that the genre under examination contains a high degree of standardisation and homogeneity, although these two traits do not seem to characterise the English and Welsh judgments. As far as the comparison with the general language is concerned, routine formulae are hardly present in reference corpora.

The following final section attempts to interpret these results in the light of the two different legal traditions characterising the three systems: common law on the one hand (England and Wales), and civil law on the other hand (Spain and Italy).
Results thus obtained have confirmed the three initial hypotheses. As far as the first research question is concerned, criminal judgments display a high percentage of phraseological units in the three subcorpora. The comparison between frequency of co-occurrence of the extracted phraseologisms and their frequency in reference corpora (e.g. **CREA** for Spanish, **CORIS/CODIS** for Italian and **BNC** for English) confirms that phraseology is indeed a key lexico-syntactic feature of this genre and it is part of judges’ idiosyncratic drafting conventions.

As far as the second research question is concerned, the English subcorpus shows a lower degree of standardisation and, consequently, a lower percentage of phraseological units, compared to the Spanish and the Italian subcorpora. As illustrated extensively elsewhere (see, in particular, Pontrandolfo 2013b: 144–145), although criminal judgments have the same function in the three judicial sys-
tems, their content and their textual realisation differ significantly. English and Welsh judgments are the results of a long oral tradition. Common-law judges have their personal style, that they use to justify their decisions in a personal, subjective way. There are no constrictions in the outline or content of the texts they produce, also because, unlike the other two, the English and Welsh judicial system lacks important reference texts, such as the Spanish or Italian Codes of Criminal Procedure. This affects the standardisation and the different phraseological weight between English and Welsh judgments on the one hand, and Spanish and Italian on the other.

As far as the third research question is concerned, the results of the analysis interestingly highlight the comparability of phraseologisms found in the Spanish, Italian and English criminal judgments. The presence of “parallel phraseologims” is of crucial importance in terms of legal translation. Functional equivalence (see Tognini-Bonelli 1996) can be achieved in most cases, as can be demonstrated by applying the “translation by collocation” approach (see Tognini-Bonelli & Manca 2004; Pontrandolfo 2013b). From the perspective of judicial reasoning, the results seem to point to the existence of a “legal/judicial grammar”, especially in the case of Spain and Italy, namely a series of phraseological, idiosyncratic conventions that typically recur in judicial discourse.

8 Conclusions

As far as the applications of the present study are concerned, the specialised phraseologisms yielded by the research can serve, first and foremost, phraseographic purposes, providing legal translators with a practical guide containing useful information on the contexts of use and, above all, the frequency of some expressions (see Lombardi 2004). Such tool will help translators in the stylistic rendering of their target texts and “reassure” them about the appropriate linguistic and legal use of specialised phraseological combinations.

The extracted phraseological units can also be used for lexicographic purposes, integrating already existing legal databases or dictionaries, or constituting the basis for new phraseological resources specifically designed for legal language.

However, the most valuable application of the study is in the training of legal translators. Familiarising with the “routines” of the genre (Hatim & Mason 1997), as well as mastering their use (both at receptive and productive level) are crucial factors in legal translators’ training (see Garzone 2007). Phraseology is also a fundamental way for trainees to understand the conceptual relation between the different elements of a specialised text. While terms map out the legal system
and therefore pertain to the knowledge (discipline) space in each judicial system, phraseology structures the texts of the legal domain. Getting familiar with the specific phraseology of the register of a discourse community will therefore bring about not only a better knowledge of the genre, but also an enhanced competence in the process of writing and reading specialised registers (see Williams 2002). A tool like cospe can help legal translators improve their phraseological competence, showing them how to produce texts that fit the stylistic conventions of the target language original texts.

One of the future challenges will be that of enlarging cospe to include other legal genres, as well as a parallel corpus. This would allow a replication of the study with different legal texts, focusing on a comparison between phraseological behaviours across different genres. Furthermore, a new hypothesis will be tested: phraseology as a quality-enhancing factor in legal translation.

References


Future studies will attempt to answer the following research question: What is the relationship between phraseology and the quality of the target text? In other words, is there a relationship between the translators’ phraseological competence and the final quality of the text? Can phraseology contribute to improving the quality of translated texts, and if so, how?


Granger, Sylviane. 2005. Pushing back the limits of phraseology: How far can we go? In Christelle Cosme, Céline Gouverneur, Fanny Meunier & Magali Paquot


Gianluca Pontrandolfo


Pontrandolfo, Gianluca. 2013a. La fraseologia come estilema del lenguaje judicial: el caso de las locuciones prepositivas desde una perspectiva contrastiva. In
Luisa Chierichetti & Giovanni Garofalo (eds.), *Discurso profesional y lingüística de corpus. Perspectivas de investigación*. Bergamo: CELSB.


