PONTIFICIA UNIVERSIDAD
CATÓLICA DE CHILE
FACULTAD DE LETRAS

Web-based tools and resources for legal translators: the JudGENTT translation-oriented glossaries for criminal courts translators

Anabel Borja Albi
Universitat Jaume I
España

Isabel García-Izquierdo
Universitat Jaume I
España

Anabel Borja Albi: Departamento de Traducción y Comunicación, Facultad de Ciencias Humanas y Sociales, Universitat Jaume I, España. | Correo electrónico: borja@uji.es
Isabel García-Izquierdo: Departamento de Traducción y Comunicación, Facultad de Ciencias Humanas y Sociales, Universitat Jaume I, España. | Correo electrónico: igarcia@uji.es

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

Abstract

One of the requirements that most obviously characterises the work of translators in highly specialised fields is to have instant access to expert knowledge management systems that enable them to carry out their work with the precision and rigour demanded when working with specialist texts. However, in the field of professional private translation practice we do not yet have sufficiently powerful tools (of the kind available to international organisations) to enable translators to carry out their terminological work in a systematic way, combining the requirements for linguistic information with those for conceptual information and those related to the time pressure to which they are frequently subject in their work.

In this paper we reflect on the general features that such expert knowledge management systems for freelance legal translators should possess, and in particular on the terminological resources that they should feature. To this end we present the design and methodology used by the GENTT research group in creating highly specialised glossaries for translators of criminal court procedure documents from four legal systems: Spain, France, the United Kingdom and Germany. These glossaries are part of a broader web-based expert knowledge management system for translators which also encompasses conceptual (legal concepts and procedures) and textual resources (corpus).

Keywords: Legal translation; expert knowledge management systems for legal translation; glossaries for legal translators; corpus based terminology extraction; e-resources for criminal court documents translation.

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

1. Introduction[1]

One of the requirements that most obviously characterises the work of freelance translators in highly specialised fields is to have instant access to expert knowledge and terminology management systems that enable them to carry out their work with the precision and quality demanded when working with specialist texts. In international organisations, equipped with powerful translation and interpreting services, such systems currently operate on computer platforms which use various technologies for indexing, searching, retrieving and analysing linguistic occurrences. However, in the field of professional private translation practice we do not yet have sufficiently powerful tools (of the kind available to international organisations) to enable translators to carry out their terminological work in a systematic way, combining the requirements for linguistic information with those for conceptual information and those related to the time pressure to which they are frequently subject in their work.

In the following pages we reflect on the general features that an expert knowledge management system for criminal court documents translators should possess, and in particular on the terminological resources that it should feature. The design and methodology used by the GENTT research group in creating highly specialised glossaries for translators of criminal court documents from four legal systems (Spain, France, the United Kingdom and Germany) shall be discussed as part of a broader online expert knowledge system platform for legal translators, the JudGENTT platform (Borja, 2013).

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

2. New technologies and electronic resources to improve the quality and efficiency of legal translation

In the last few years there has been a proliferation of information resources for translators, but there is a major shortage of comprehensive and accurate multilingual resources that satisfy the needs of legal translators. Translation solutions for certain legal concepts and texts require a very high level of precision and faithfulness to safeguard the principle of legal safety and make the target text carry the same function and legal effect. Following Garzone (2000: 3) "…according to the principle of legal equivalence, the translation of a legal text will seek to achieve identity of meaning between original and translation, i.e. identity of propositional content as well as identity of legal effects, while at the same time pursuing the objective of reflecting the intents of the person or body which produced the ST". This statement, although accurate, would need further elaboration to take into account different translation *skopos*.

However, finding translation equivalents in the legal field is sometimes a complex task, requiring techniques and procedures that a lawyer would describe as "comparative law techniques" and a linguist as "comparative definitions methodology", and which, in any case, are not addressed by traditional glossaries.

> The distinctive nature of legal concepts works as a decisive factor against equivalence. We stated above that legal concepts are embedded in a specific working environment and in national legal systems, and that each national legal setting has its own principles for the application of concepts. There cannot be absolute equivalence, unless it is a consequence of complete identity of moral values, legal provisions, interpretation rules and forms of application of laws - but this again would mean the same legal framework […] A methodology should be developed to deal with the cases of partial equivalence or overlapping characteristics. To achieve this we have to abandon the concept of equivalence in favour of a more flexible comparative approach (Sandrini, 1996).

The high degree of expertise and comparative legal knowledge required in this domain obliges translators to seek new systems for acquiring and retrieving knowledge and data and to incorporate electronic systems for multilingual information management and

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

retrieval into their working routines. Legal translators scour the Internet for glossaries, reference texts, parallel documents and forums. There is indeed an incredible amount of information, much of it fairly "anonymous" and with no indication on the quality and relevance of the source or data. Certainly, legal translation may benefit from web-based electronic resources in the form of discussion fora, google searches, electronic termbases (e.g. IATE, TERMIUM Plus®, Juriterm Database, UNTERM Linguee…), translation memories (e.g. DGT's TM, MyMemory…), and translation-driven corpora (see Biel, 2010; Garofalo, 2009; Pontrandolfo, 2014). But what translators are looking for is not to lose themselves for hours in the tangled web of the Internet, but to find quick, reliable answers to real doubts and problems, as well as selected, systematised compilations with the most terminologically valuable and useful web addresses for specific domains of specialty.

Starting from a review of the literature on multilingual knowledge representation based on web technologies for legal translators (see among others: Dam et al., 2005; Borja, 2005), empirical studies on the documentary requirements and habits of legal translators (García-Izquierdo & Conde, 2012, and Pajares Nievas, 2013, among others) and our own experience and needs as specialised translators, our research team (GENTT research team) has kinds (legislation, case law and legal doctrine corpora, translation memories, glossaries of terms and phraseological units, style guides, etc.) in an optimised form, through a single interface adapted to the needs of target users. The translation-oriented glossaries (and other resources) included in this system are therefore aimed at a very well defined set of users with very specific needs: highly specialized translators working in criminal judicial settings.

As early as 1997, Mayoral (1997: 148) remarked that the sources of information for specialised translation must fulfil a series of requirements, which include reliability, accessibility, originality, specificity, comprehensiveness and alignment with user requirements. In the opinion of this author (1997: 139), the process of classifying information for translators is different from that which is performed for documentalists and even for terminologists, because translation is a communication process subject to different

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

needs and priorities, although among such sources we also find terminographic materials, such as dictionaries and glossaries.

And as we shall shortly see, this is precisely what our proposal offers: specific electronic resources for a very well defined set of legal translators (those working in/for criminal courts), easy to access and integrated into a web-based information system. The organisation of the information (contained in relational data bases) responds to the needs identified in various empirical studies on translators preferences and needs and is easily retrievable through a single user friendly interface.

3. An expert knowledge system for legal translation: the JudGENTT platform

The JudGENTT online platform offers legal translators a range of linguistic and legal resources (*ad hoc* monolingual and bilingual corpora, terminological and phraseological glossaries, subject-field information…) which facilitate the conceptual and terminological documentation process, making the comparative analysis of legal terms and concepts easier, and helping them to reuse terminology and existing translations more effectively. In its current phase it includes resources related to four legal systems: Germany, Spain, France and the United Kingdom.

The JudGENTT platform has been designed as an intelligent multilingual specialised information management system that will make it possible to automate the processes of retrieval, indexing, semi-controlled composition and assisted translation of professional texts, always taking account of the needs, habits and processes of the end users of the texts. This expert system for translators has been designed based on the concept of *textual genre*. Previous studies by the research group (Borja, 2005; Borja, 2013; Borja y García-Izquierdo, 2015; García-Izquierdo, 2005, ed.), have demonstrated the usefulness of the idea of textual genre, as a complex concept combining formal, communicative and cognitive considerations, for characterizing the types of language that serve as vehicles for specialized communication/translation, and it is precisely this concept that underlies and governs the construction of the GENTT corpus.

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

The contents collected or compiled so far are organised in five sections[2]:

- A *Catalogue of criminal court documents* which includes a list of the documents (*genres*) used in criminal judicial procedure in the four domestic legal systems studied accompanied by an explanatory genre matrix on each of them.

- A series of specialized *Glossaries of judicial terminology and phraseology* created by extracting terminology from a subset of the platform's *ad hoc* corpora.

- A section called *Criminal courts procedure,* in which conceptual/contextual information pertaining to the four national systems can be found: courts, judicial personnel, rules of procedure, crimes and penalties…

- A *Resources for translators* section, which includes a subsection for Bibliography and another for Links, with connections to carefully selected web pages and portals useful to legal translators.

- Perhaps the most important contribution this tool offers is the *ad hoc corpus* of criminal court documents the team has compiled. It is made up of monolingual corpora in four languages (Spanish, English, French and German) and bilingual corpora (in the different combinations of Spanish and the other three languages). It comprises a collection of genuine judicial documents (originals and translations: witness summons, court orders, judgements, case summaries, bail forms…) from which personal details have been removed, forms and precedents, legislation, case law, doctrine, legal opinions... all of them related to criminal procedure matters practice. This corpus is the database for the tool's *integrated search facility*, which covers all sections of the site and retrieves results of various types depending on the search criteria selected in the advanced search window. It is also the source corpus from which the extraction sub-corpus has been compiled to identify the glossary terms.

---

[2] A detailed explanation of these sections can be found in Borja, 2013.

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

Figure 1. JudGENTT platform interface

The JudGENTT web-based platform has been developed using Drupal, an open source content management system (CMS). For full-text indexing and search of the platform databases (*ad hoc* corpus of explanatory texts, *ad hoc* corpus of prototypical texts, html resources and glossaries) we use a Lucene and Solr text search engine library.

Figure 2. GENTT platforms web architecture

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…
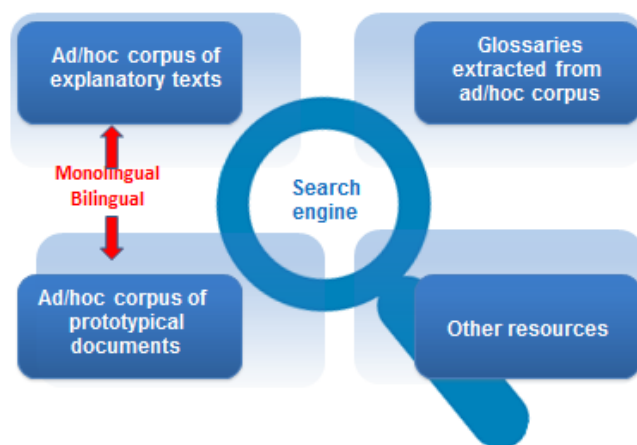
4. How can terminology help legal translators?

As we remarked in section 2, legal asymmetry or cultural anisomorphism (differences between legal systems and cultures) poses serious difficulties for legal translators due to the lack of conceptual and terminological equivalence, especially in those terms related to procedures, institutions and personnel. This shortcoming can only be addressed through expert knowledge of legal concepts in the two systems that the translation brings into contact and through accurate and powerful knowledge representation resources.

Multilingual knowledge representation is the object of study of various disciplines: knowledge engineering, terminology engineering, ontology-based systems, lexicographic approaches, computational linguistics, natural language processing and any other context that supports multilingual knowledge representation methods.

The object of terminology is to satisfy social needs related to optimizing communication between specialists and professionals; hence our interest in their working methods, which, in part, inspired the compilation of our glossaries. Specifically, we have drawn on concepts from *translation-oriented terminology*, which is based on the idea that specialized texts convey knowledge through speciality terms and aims to identify terminological equivalences in the various languages that can serve as a point of reference for the translator and contribute to the quality of the translated text.

In this connection, communicative approaches (Cabré, 1992, 2000; Gaudin, 2003) and cognitive approaches (Temmerman, 2001; Faber, Márquez & Vega, 2005) to terminology that involve using objective structures drawn from reality are highly productive for specialized translators. Our glossaries, however, do not conform to the canonical structure of projects for creating terminological resources. In the case of legal translation, the idea that there is a structure, a tree of predefined concepts regarded as "universal" in all languages which can be assigned a given linguistic expression in each language *a posteriori*, is highly debatable. The idea that these universals constitute the link

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

between terms in different languages, and that they convey information that is not specific to a legal culture, is not productive in the field of legal translation.

In contrast, in our knowledge system, conceptual information relating to four different systems is presented in a *systematic and contrastive* form in a specific section of the platform, in which a tree of conceptual fields is established, whose initial nodes (documents, procedures, actors, courts, crimes, penalties…) are gradually diversified and at the same time in all the systems in an exercise in conceptual contrastivity based on comparative law. When defining the information categories for each term in the glossary, we have included, wherever possible, a reference to the location of the term in the system of conceptual nodes for the corresponding legal system, by means of the system of hyperlinks used in the platform.

The translator thus has immediate access to a doctrinal or legislative explanation of the concept in both legal systems, to real contexts of use and to a range of proposed equivalents. With this information, the translator is ultimately responsible for making informed decisions suited to the *skopos* of each translation job: using the equivalent proposed in the glossary, keeping the term in the original language accompanied by an explanatory gloss, using a functional equivalent…

In short, our approach is inspired by the theories of translation-oriented terminology, but it is clearly geared towards translation studies and deliberately eclectic. Our project cannot therefore be regarded as a contribution to the discipline of translation-oriented terminology as such, but rather as the result of an interaction between terminology and legal translation applied to the actual work of *professional* translators.

5. JudGENTT on line glossaries for legal translators

Among the basic principles for compiling monolingual and multilingual glossaries, Cabré (2012) highlights the need to bring together the skills required to complete a translation job (cognitive, linguistic, socio-functional and methodological). Our team has

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

tried to address these requirements by having not only linguists and skilled legal translators participate in compiling the tool but also experts in the specialised field involved, who have contributed cognitive competence in the form of advice on the use of particular terminological units in the context of the genres being studied and on the systemic and semantic relations between legal concepts.

As Cabré remarks (2012: 1) terminological work, of whatever kind, needs to be founded on quality or "good practice", and for this purpose the concepts of adaptation to context, reliability (real usage), systematicity and use of new technologies are crucial[3]. Too often, however, in the field of translation, we encounter work in which these requirements are not obviously fulfilled. With reference to glossaries, Cabré (2000: 3), for example, remarks:

> Many existing glossaries do not satisfy translators, either because they are out of date, or because they lack necessary information, or because they lack criteria for evaluating their quality and reliability. A terminological glossary designed to solve problems that arise when translating a specialised text must be based on an analysis of the real needs of translators in their activity and must be formulated according to those needs. But only those familiar with these needs are capable of designing and producing a terminological application that is appropriate and sufficient in terms of information, and easy to handle in terms of format and presentation[4].

To develop the JudGENTT glossaries of criminal court procedure terms in four different legal systems we have followed a four-step methodology to guarantee high quality information: 1) Compilation of an *ad hoc* extraction corpus of texts representative of the criminal law and procedure domain; 2) Selection and extraction of terms from the corpus; 3) Creation of term data (concept oriented); 4) Publishing of data on a web-based platform including appropriate categories for translators needs and using interchangeable formats which can be integrated into translation memories.

---

[3] Our translation.

[4] Our translation.

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

5.1 . Compilation of an *ad hoc* extraction corpus of texts representative of the criminal law and procedure domain

The first step was to decide on the terms to be included in each of four glossaries (British, French, German and Spanish systems or jurisdictions). For this purpose we selected firstly the reference texts for this discipline in the four jurisdictions (not in the four languages, as law on a particular issue differs according to the specific jurisdiction). For all the systems we decided to include the statutory texts of substantive and adjective criminal law: the *Código Penal* (Criminal Code) and the *Ley de Enjuiciamiento Criminal* (Criminal Procedure Act) in Spain, for instance, and the equivalent legal instruments in the other systems where applicable. For the British glossary, for example, we also used Archbold's handbook on Criminal Law as an extraction text, in view of the differences between Common Law and Civil Law systems with respect to written law. In order to obtain an up-to-date selection of terms we expanded the source corpus with case law from the last ten years. The status, normative character and prominence of these texts make them the standard references for the study of criminal court proceedings, and we considered them sufficient to ensure the reliability and comprehensiveness of the selection of terms.

As we mentioned before, although our proposal is multilingual, as we work with different languages, the extraction corpora for the monolingual glossaries include only original texts from each jurisdiction, from which we generate independent monolingual glossaries for each system, and to these we subsequently add equivalent terms in the other working languages, as we shall explain later. The software we have used (Terminotix SynchroTerm) has given excellent results in the monolingual extraction process. Having completed the monolingual extraction we used the selected terms as keys for mining in TM or parallel corpora to find target language equivalents.

5.2 . Term extraction for court translation: selection and use of source corpus

The identification of terms for a glossary aimed at translators is one of the most frequently performed terminology management tasks. Manual extraction is very time-

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

consuming and error-prone, and therefore specific software needs to be used. We need fast, powerful software that allows automated batch processing and a well-defined manual "clean-up" system. As already mentioned, automatic extraction was performed using the SynchroTerm program from the Terminotix suite. The program generates an automatic initial selection of "candidate terms" which requires manual revision to discard non-terminological units belonging to the field being analysed. In a first phase, members of the team who are experts in translation (but not experts in law, though with some training in that field) were responsible for choosing the obvious terms, indicating doubtful terms. In the second phase, legal experts checked the first version to validate or reject the initial selection. It was a semi-automated processing routine to make the glossary more coherent and eliminate "noise".

Bearing in mind that the glossary is intended for specialised translators, the criteria used for the selection of terms were: domain specificity, significant frequency, terms that are difficult to translate, new or ambiguous terms, and terms for which different translations are possible. However, given that the glossary is highly specialised, we included low-frequency terms if they fulfilled the other requirements mentioned. We excluded general lexical items, terms containing typographical errors, proper names and generalist legal terms not related with the criminal procedure already present in existing glossaries.

Our glossary includes both single terms and phraseological units. The latter are expressions coined within the specialised field that exceed the maximum number of words for them to be regarded as terms. There exists a great deal of diversity regarding phraseological units in terms of length, boundaries, variations allowable and the methodologies for their identification in corpora. However, for practical reasons we applied an *ad hoc* length limitation (10 words) based on pilot trials on our extraction corpus. We completed the extraction process with terms that we felt should appear in our glossary, on the basis of our experience in the field or after consulting existing glossaries and dictionaries. The only terms or expressions we added were related to those found in the extraction process, to complete the semantic fields or ontologies conceptually. If we did not apply this restriction we could continue adding terms *ad infinitum* and our work would end

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

up as a mere reworking or amalgamation of existing glossaries. To date the Spanish glossary includes 2,300 units and the British one, 1.300. The French and German glossaries are still in their initial phases and we do not have yet the final number of entries they will encompass.

Finally, we would like to mention some of the methodological problems we have had to tackle in this phase, mainly related to the presentation of the results of the glossary. This determined how the search engine would proceed to interrogate the database; a series of standards therefore had to be agreed with respect to the following questions, mainly of a morphological nature:

a. In the initial sifting of the results obtained from the glossary, the number morphemes had to be standardised, and apart from cases where we were dealing with *plurale tantum* expressions (e.g., *voluntades anticipadas*; compare "damages") or those that have become lexicalised in the plural through usage (*usos forenses*, *posibles efectos adversos*; compare "interrogatories" or "jury instructions"), we opted to use the singular in all cases.

b. Regarding the category of descriptors, the main problem we encountered was the presence of both infinitives and nominalisation. For example, we found ourselves with examples such as *abrir juicio oral* ("to open oral proceedings") and *apertura de juicio oral* ("opening of oral proceedings"), which inevitably coexist in the system. In these cases we have cross-referenced the entries. Otherwise, unless the expression was coined using the infinitive (e.g., *dar traslado*, "to serve notice"), we have opted for nominalisation in all cases.

c. In the case of complex expressions consisting of verb + (art.) + noun, or verb/noun + prep. + art. + noun, our intention in principle was to omit the intermediate elements wherever possible: thus, for example, expressions such as *abrir juicio oral* are included in the glossary; however, this decision could not be applied across the board, since in cases like *admitir un recurso* ("to allow an appeal") or *apertura de/del juicio oral* ("opening of [the] oral proceedings") the omission is not appropriate, as they are structures that have become lexicalised in that form.

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

d. With phraseology or use of complex syntactic units, we had to decide whether to include the terms as single entries or break them down into several. In the case of phrases such as *adopción de medidas cautelares* ("issuing of an injunction") we decided to include both the complete phrase and *medidas cautelares*.

e. Finally, with regard to expressions that share the descriptor term, we decided to regard them as independent, provided that the descriptor was not a specialised term. So, for example, although *acto punible* ("punishable act/offence"), *acto pericial* ("expert act/operation") and *acto de comparecencia* ("[act of] appearance") are all types of *acto*, they are treated separately and not as sub-entries derived from the hyperordinate concept *acto* (since we are constructing a specialised glossary and this word is not regarded as a term); however, *acusación privada* (private action/prosecution brought by citizen(s) affected by a private offence), *acusación popular* (private action/prosecution brought by citizen(s) not affected by the offence) or *acusación particular* (private action/prosecution brought by citizen(s) affected by a public offence) will be regarded as being derived from the entry *acusación*, since the latter *is* a term in the field being studied.

5.3 . Creation of term data: contexts, definitions, equivalencies…

Following a review of the data categories proposed in ISO 12620, 2009 for creating Terminology Records and those of the Juriterm Database, as well as revising several recent university research projects on e-resources for legal translators: the LAW10n project (Bestúe and Torres, forthcoming) and the JCiCM project (Biel, 2014), we defined the following fields for our datase:

1) Administrative data of the glossary

- Glossary name
- Languages
- Term extraction corpus

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

- Context extraction corpus

- Authors and revisers

- Domain

- Subdomain

2) Term information

- Entry

- Author and reviser

- Language

- Frequency (occurrences in extraction corpus)

- Domain

- Subdomain

- Part of speech/Number

- Context

- Context source (particular document of the extraction corpus)

- Definition

- Hyperlink to legislation

- Definition source

- Remarks

- Conceptual relations (hyperlinks to the conceptual section of the system described in section 3)
- Synonyms

- Cross-references to related terms

3) Translation data (Equivalence into other legal systems: total, partial, zero)

- Equivalence

- Equivalence source

- Functional, partial and zero equivalence

- Functional, partial and zero equivalence source

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

- Functional, partial and zero equivalence explanation
- False friends

This internal data base structure is very wide-ranging and we will only comment on its most important features: context, definition and equivalences. In the following section on web publication we will highlight those we have selected as significant for the end user, which will be those that appear in the user interface.

With respect to the automatic markup of the selected terms, the program allows us to configure the system so as to make it possible to retrieve different kinds of information for each terminological unit selected. Our initial markup includes source, author of extraction, target text and author of target text (in the case of bilingual texts), domain, subdomain (in our case: criminal law and criminal procedure rules), context and frequency.

The SynchroTerm program allows the user to select one out of the various **contexts of use** in which each term appears in the source corpus when performing manual clean-up. In selecting the contexts we have tried to make them representative in terms of both usage and meaning, giving priority to contexts in which the terms are defined by legislative instruments. Having obtained the list of terms and of the context or contexts in which they appear, the next step was to introduce the definitions.

To ensure that the **definitions** were reliable, we took advice from legal experts, who recommended to us the most authoritative and up-to-date dictionaries best adapted to the legal system. This is undoubtedly proving to be the most laborious phase of the process, due to the fact that we are working with a number of sources for each language and have to rely on the legal experts for final checking. To select the dictionaries we used the following criteria: they had to be related to the specific legal system (we have not used US dictionary definitions for the UK, nor have we used dictionaries from Latin American countries for the Spanish glossary despite the high quality of some of them); and the date of publication had to be as recent as possible, as law is a rapidly changing, dynamic phenomenon which evolves with society.

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

We have taken special care to avoid the definitions being obsolescent, and in terms whose definition depends on the current law we have indicated this and have included diachronic information where necessary. The definitions have been carefully checked by a legal expert and a senior legal translator to make sure that they are accurate and satisfy the needs of the end user.

Equivalencies. In terms of the parameters normally used to classify terminological tools, ours belongs to the category that Cabré and Tebé (2005: 20) call "ad-hoc terminology work" (*trabajo terminológico puntual*); in other words, one that arises from the need to solve concrete terminological problems in a specialised translation context (as opposed to "systematic terminology work" carried out by standardising organisations).

With regard to the number of languages, moreover, our proposal is multilingual, since, with certain qualifications to be explained later, all the working languages are treated equally. However, given the specificity of the legal concepts in each system, our glossaries could also be described as monolingual glossaries, since the selection of terms, the contexts and the definitions correspond to the legal reality of a particular country, or rather of a particular legal system, and in many cases the proposals do not coincide. Court procedures are totally different in each country. There are differences in the types of proceedings, hierarchy of courts, trial documents, crimes, penalties… This lack of equivalence means that the level of correspondence between the glossaries for the different languages is sometimes very low, and that we can therefore regard them as independent monolingual glossaries directly linked to the organisation of criminal procedure in the country in question, in which partial equivalents, accompanied by comments on comparative law, are sometimes provided.

Hence, our glossaries can be considered monolingual, as they are based on the legal reality of the court system of the source language, but include equivalents in other languages. As Garofalo (2008: 251) says with reference to legal terminology:

> […] it is obvious that the definition of similar terminological units resists all attempts at conceptual standardisation, since its internal organisation is conditioned by the ideological bias of the legislators

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

in each country and by their variable interpretation of such invariable principles as the presumption of innocence, enshrined in most modern democratic constitutions[5].

As a result of this conceptual anisomorphism the degree of correspondence between the glossaries of the different languages is sometimes low, and we can therefore regard them as independent monolingual glossaries, directly linked to the organisation of criminal procedure in the country in question, and sometimes offering partial equivalents accompanied by comments on comparative law. Much has been written on equivalence in legal translation but lack of space prevents us from exploring this interesting issue here, and we shall confine ourselves to describing the solutions we have adopted.

To alleviate these difficulties we have relied on the most highly regarded bilingual dictionaries, and when necessary we have indicated that there is no equivalent or only a partial equivalent by adding explanatory notes written by experts in comparative law. It should be emphasised that the greatest difficulties are encountered in intra-category organisation; in other words, in the way each legal system organises types of offences and types of injunctions, for example. To demonstrate these differences the main requirement is to be able to draw on encyclopaedic information about the legal system of each country (Garofalo, 2008: 250).

Although the names of courts and judicial posts are in principle "do-not-translate" terms, our glossary includes various solutions for this kind of term so that translators can chose whichever they consider most suitable according to the *skopos* of the translation. For the term *Crown Court*, for example, several solutions are proposed: keeping it in its English form (Crown Court), using a word-for-word translation (*Tribunal de la Corona*: "court of the crown"), or a combination of methods, leaving the term in the original language and including a brief explanation in parentheses on the hierarchical level of the court and the types of cases it hears.

---

[5] Our translation.

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

5.4 . Publishing of data on a web-based platform. Versatility of presentation of the results in web-based format

With regard to format, one of the advantages of presenting glossaries in web-based format is the ability to access large amounts of information easily and intuitively, owing to the possibility of cross-referring and hyperlinking between the different parts of the platform and other sources of legal information. The glossaries user interface currently shows the term, definition (with links to legislation and doctrine if necessary), context, equivalence (and explanations on partial or zero definition with hyperlinks to encyclopaedic information), and context of the equivalence. However, the versatility of the system makes it possible to show different combinations of data in the user interface (see figure 3).
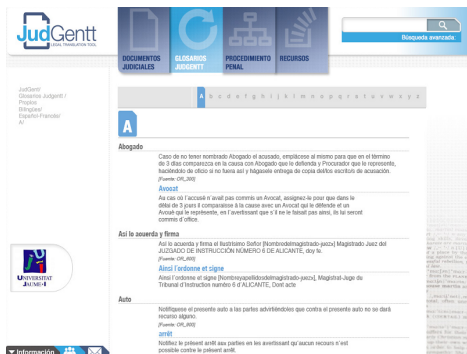

Figure 3. Glossaries user interface

Furthermore, web-based format, enables us to relate terms to each other, incorporate encyclopaedic information and legal analysis, view the contexts of production simultaneously or consult the relevant legislation by means of the search engine and the advanced search functionality, among other possibilities (see figure 4). The ability to obtain results through an advanced search engine and hypertext data enables use of the glossary to be adapted to the needs of each user. The advanced search function interrogates and

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

retrieves data both from the glossary and from the source corpus and the other resources included in the platform (full-text court genres, legal textbooks, case law, legal opinions, etc.). This enables users to look for a traditional definition of a term and its usage context or to conduct more far-reaching terminological research by consulting occurrences of the term in legislation, encyclopaedic explanations and contradictory interpretations of the same concept or point of comparative law according to different authors and schools.
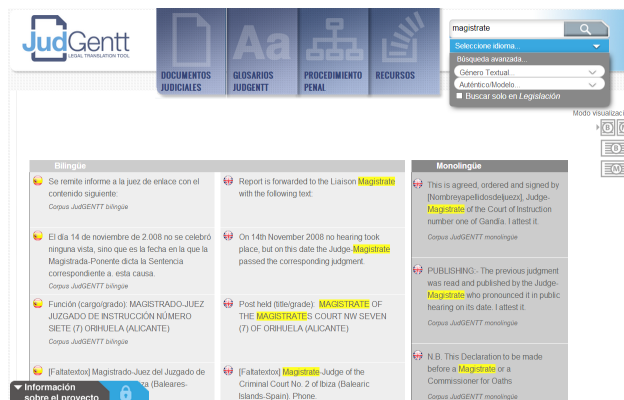

Figure 4. Search engine results page

6. Conclusion

Multilingual knowledge representation based on web technologies is an open research area calling for interdisciplinary approaches to creating, managing, and using information on the web. With this contribution we have tried to demonstrate that the JudGENTT platform and its translation-oriented glossaries meet all the requirements identified in the current literature: reliability, accessibility, originality, specificity, comprehensiveness and alignment with user requirements. It is a knowledge management system for translators designed and compiled by specialist legal translators and validated by experts in criminal and procedural law. The web design, development and delivery of the information allow translators to choose which data they would like to access and with what

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

degree of detail. In order to achieve this it is presented in a web-based format with hypertext and advanced search functions as part of a larger expert knowledge system.

Our proposal of translation-oriented glossaries offers reliability, guaranteed by the methodology employed and by the authority of the sources of extraction, the reference texts used to obtain definitions and usage contexts. The web-based format makes it possible to keep it constantly updated. A further guarantee of reliability is the fact that all the information is reviewed by experts (legal and linguistic) responsible for the final checking. As is apparent from the foregoing, the field delimitation is strict, which means that a high degree of specificity is achieved in the resources generated, since we are working from *ad hoc* source corpora that are very restricted in terms of subject matter and text genre.

The use of web-based technology and the inclusion of the glossary in an expert knowledge management system guarantee its accessibility and represents an innovative approach designed by translators working in highly specialised areas. There are many online resources available (websites, repositories, termbanks, glossaries, translations memories…) for legal translators but not so many that provide information of different kinds—linguistic and textual, conceptual, pragmatic, etc.—in an integrated way.

As regards alignment with user requirements, the JudGENTT glossary responds to the end user needs detected in a preliminary phase of the project devoted to identifying requirements (surveys and focus groups).

We are aware of the limitations of this study, as it is a first approach to a glossary for translators incorporated into an expert knowledge base which offers ample scope for expansion to refine criteria and improve the final product. However, our project is essentially dynamic and aims to be an interactive platform work in which expert users can make contributions and suggestions identify errors, etc., so as to foster collaborative expert knowledge building in the translation community.

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

7. References

BESTUÉ, Carmen & Olga TORRES HOSTENCH (forthcoming): "Technology and e-resources for legal translators: the Law10n project" in Pilar SANCHEZ-GIJÓN (ed.): *The Coming of Age of Translation Technologies in Translation Studies*, London: Peter Lang.

BIEL, Lucja, 2010: "Corpus-Based Studies of Legal Language for Translation Purposes: Methodological and Practical Potential" in Carmen HEINE & Jan ENGBERG (eds.): *Reconceptualizing LSP. Online proceedings of the XVII European LSP Symposium* 2009, Aarhus: Aarhus School of Business, Aarhus University [http://bcom.au.dk/research/publications/conferencepublications, accessed 15 March 2015].

BIEL, Lucja, 2014: "The JCiCM Platform Improving the quality and efficiency of legal translation in criminal matters via corpus-based tools", presentation to the final *Qualetra Conference,* held in Antwerp, on October 2014 [http://www.eulita.eu/qualetra-final-conference-presentations, accessed 15 March 2015].

BORJA ALBI, Anabel, 2005: "Organización del conocimiento para la traducción jurídica a través de sistemas expertos basados en el concepto de género textual" in Isabel GARCÍA-IZQUIERDO (ed.): *El género textual y la traducción. Reflexiones teóricas y aplicaciones pedagógicas*, Bern: Peter Lang, 37-68.

BORJA ALBI, Anabel, 2013: "A genre analysis approach to the study of the translation of court documents", *Linguistica Antverpiensia* 12. *New Series – Themes in Translation Studies, Research models and methods in legal translation*, Lucja BIEL & Jan ENGBERG (eds.), 33-53.

BORJA ALBI, Anabel & Isabel GARCÍA-IZQUIERDO, 2015: "Corpus-based knowledge management systems for specialized translators: bridging the gap between theory and practice" in M. Teresa SÁNCHEZ NIETO (ed.): *Corpus-Based Translation and Interpreting Studies. From Description to Appplication*, Berlin: Frank & Timme, 191-209.

CABRÉ, M. Teresa, 1992: *La terminologia. La teoria, els mètodes, les aplicacions*, Barcelona: Empúries.

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

CABRÉ, M. Teresa, 2000: "El traductor y la terminología: necesidad y compromiso", *Panace@*, vol. 1, 2 [http://www.medtrad.org/panacea/PanaceaPDFs/Panacea2_Diciembre2000.pdf, accessed 15 March 2015].

CABRÉ, M. Teresa, 2012: "Buenas prácticas en terminología", *Puntoycoma* 130, Brussels/Luxemburg: Dirección General de Traducción de la Comisión Europea, 25-33.

CABRÉ, M. Teresa and Carles TEBÉ, 2005: "El trabajo terminológico multilingüe de enriquecimiento lingüístico: una propuesta metodológica", *Revista española de lingüística aplicada* 1, 19-41.

DAM, Helle V., Jan ENGBERG and Heidrun GERZYMISCH-ARBOGAST (eds.), 2005: *Knowledge Systems and Translation*, Berlin: Mouton de Gruyter.

FABER, Pamela, Carlos MÁRQUEZ LINARES & Miguel VEGA EXPÓSITO, 2005: "Framing Terminology: A Process-Oriented Approach", *META* 50 (4). CD-ROM.

GARCÍA-IZQUIERDO, Isabel (ed.), 2005: *El género textual y la traducción: reflexiones teóricas y aplicaciones pedagógicas*, Bern: Peter Lang.

GARCÍA IZQUIERDO, Isabel & Tomás CONDE, 2012: "Investigating Specialized Translators. Corpus and Documentary Sources", *Ibérica* 23, 131-157.

GAROFALO, Giovanni, 2008: "Unidades de comprensión del dominio judicial español e italiano: el caso de las medidas cautelares 'prisión provisional'/*custodia cautelare*" in Carmen NAVARRO et al. (eds.): *La comunicación especializada*, Bern: Peter Lang, 243-260.

GAROFALO, Giovanni, 2009: *Géneros Discursivos de la Justicia Penal. Un análisis contrastivo español-italiano orientado a la traducción*, Milano: Franco Angeli.

GARZONE, Giuliana, 2000: "Legal Translation and Functionalist Approaches: a Contradiction in Terms?" in *ASTTI/ETI*, 395-414.

GAUDIN, François, 2003: *Socioterminologie: Une approche sociolinguistique de la terminologie*, Brussels: Duculot.

MAYORAL, Roberto, 1997: "La traducción especializada como operación de documentación", *Sendebar* vols. 8/9, 137-153.

ONOMÁZEIN 33 (junio de 2016): 226-250
Anabel Borja Albi e Isabel García-Izquierdo:
Web-based tools and resources for legal translator: the JudGENTT translation-oriented…

PAJARES NIEVAS, José S., 2013: *La traducción jurada u «oficial» de textos médico-jurídicos: análisis de la práctica socioprofesional y de los recursos de documentación disponibles.* Unpublished master's thesis. Universitat Jaume I, Castellón (Spain).

PONTRANDOLFO, Gianluca, 2014: "Marcadores argumentativos del contraste y discurso judicial: un estudio propedéutico para la traducción", *Hermes – Journal of Language and Communication in Business* 53, 99-124.

SANDRINI, Peter, 1996: "Comparative Analysis of Legal Terms: Equivalence Revisited" in Charles GALINSKI & Klaus D. SCHMITZ (Hg): *TKE '96*, Frankfurt: Indeks, 342-351.

TEMMERMAN, Rita, 2001: "Sociocognitive terminology theory" in M. Teresa CABRÉ & Judit FELIU (eds.): *Terminología y Cognición*, Barcelona: Universitat Pompeu Fabra, 75-92.