

# 32 Using corpus management tools in public service translator training: an example of its application in the translation of judgments

María Del Mar Sánchez Ramos<sup>1</sup>  
and Francisco J. Vigier Moreno<sup>2</sup>

---

## Abstract

As stated by Valero-Garcés (2006, p. 38), the new scenario including public service providers and users who are not fluent in the language used by the former has opened up new ways of linguistic and cultural mediation in current multicultural and multilingual societies. As a consequence, there is an ever increasing need for translators and interpreters in different public service environments (hospitals, police stations, administration offices, etc.) and successful communication is a must in these contexts. In this context, Translation Studies has seen the emergence of a new academic branch called Public Service Interpreting and Translation (henceforth PSIT), which is present in a wide range of environments where communication (and mediation) is, as stated above, essential, such as healthcare, education and justice to name a few. In PSIT, legal translation principally involves the documents most commonly used in criminal proceedings, as in Spain legal aid is usually provided in criminal cases. Hence, PSIT legal translation training is intended to help trainees to develop their legal translation competence and focuses mainly on legal asymmetry, terminological incongruence, legal discourse, comparative textology and, fundamentally, on the rendering of a text which is both valid

---

1. FITISPos Research Group - Universidad de Alcalá, Alcalá de Henares, Madrid, Spain; mar.sanchezr@uah.es

2. FITISPos Research Group - Universidad de Alcalá, Alcalá de Henares, Madrid, Spain; francisco.vigier@uah.es

**How to cite this chapter:** Sánchez Ramos, M. d. M., & Vigier Moreno, F. J. (2016). Using corpus management tools in public service translator training: an example of its application in the translation of judgments. In A. Pareja-Lora, C. Calle-Martínez, & P. Rodríguez-Arancón (Eds), *New perspectives on teaching and working with languages in the digital era* (pp. 375-384). Dublin: Research-publishing.net. <http://dx.doi.org/10.14705/rpnet.2016.tislid2014.449>

in legal terms and comprehensible to the final reader (Prieto, 2011, pp. 12-13). Our paper highlights how corpus management tools can be utilised in the translation of judgments within criminal proceedings in order to develop trainees' technological competence and to help them to acquire expertise in this specific language domain. We describe how monolingual virtual corpora and concordance software can be used as tools for translator training within a PSIT syllabus to engender a better understanding of specialised text types as well as phraseological and terminological information.

---

**Keywords:** legal translation, specialised corpora, concordance programs.

## 1. Legal translation training in PSIT training

The ever-increasing mobility of people across boundaries, be it for economic, political or educational reasons, has led to the creation of multilingual and multicultural societies where the need for language and cultural mediation is also ever growing. Even if this is a worldwide phenomenon, it is most conspicuous in countries which have been traditionally considered as countries of emigration and have become countries of immigration in the last 20 years, thus evolving into complex multilingual and multicultural societies. This is also the case of Spain, a country where the high influx of immigrants and tourists poses challenges which require adequate responses to ensure a balanced coexistence (Valero-Garcés, 2006, p. 36). This need for translators and interpreters is even greater in public services like schools, hospitals, police stations, courts... where users who do not command the official language of the institution must be catered for up to the point where it has fostered the creation of a new professional activity and, subsequently, a new academic branch within Translation Studies, commonly referred to as PSIT. Hence, PSIT has a very wide scope, including healthcare, educational, administrative and legal settings. PSIT legal translation is mostly concerned with the documents which are most commonly used in criminal proceedings, such as summonses, indictments and judgments; probably because it is in criminal cases that legal-

aid translation and interpreting services are provided (Aldea, Arróniz, Ortega, & Plaza, 2004, p. 89).

In an attempt to provide the education required in competent professionals, the University of Alcalá offers a program specifically designed for PSIT training, namely a Master's Degree in Intercultural Communication, Public Service Interpreting and Translation, which is part of the European Commission's European Master's in Translation network. This programme, which is offered in a wide variety of language pairs including English-Spanish, comprises a specific module on legal and administrative translation into both working languages. In line with the so-called competence-based training (Hurtado, 2007), this module is mainly intended to equip the students with the skills, abilities, knowledge and values required in a competent translator of legal texts. Based on previous multicomponent models and his own professional practice as a legal translator, Prieto (2011, pp. 11-13) offers a very interesting model for legal translation competence which encompasses the following sub-competences: (1) strategic or methodological competence (which controls the application of all other sub-competences and includes, among others, the identification of translation problems and implementation of translation strategies and procedures); (2) communicative and textual competence (linguistic knowledge, including variants, registers and genre conventions); (3) thematic and cultural competence (including but not limited to knowledge of law and awareness of legal asymmetry between source and target legal systems); (4) instrumental competence (documentation and technology); and (5) interpersonal and professional competence (for instance, teamwork and ethics). According to our experience in PSIT legal translation training, it is precisely in communicative and textual competence (especially as regards terminological and phraseological use of legal discourse in the target language) that many of our trainees show weaknesses, chiefly when translating into their non-mother tongue (in our case, English). As we firmly agree with the view that "being able to translate highly specialised documents is becoming less a question of knowledge and more one of having the right tools" (Martin, 2011, p. 5) and that we must ensure that our students "move beyond their passive knowledge of basic legal phraseology and terminology and take a more proactive stance in the development of their legal language proficiency" (Monzó, 2008,

p. 224), we designed the activity explained below to make our students aware of the usefulness of computer tools when applied to legal translation to overcome many of the shortcomings they face when translating legal texts.

## 2. Corpora in PSIT training

The pedagogical implications of using corpora in specialised translator training have been shown by various researchers (Bowker & Pearson, 2002, p. 10; Corpas & Seghiri, 2009, p. 102; Lee & Swales, 2006, p. 74), and also specifically in legal translator training (Biel, 2010; Monzó, 2008). Some of the main advantages identified are related to the development of instrumental sub-competence (PACTE Group, 2003, p. 53), or so-called information mining competence (EMT Expert Group, 2009). The need to know and use different electronic corpora and concordancing tools is also illustrated by Rodríguez (2010), who identifies a further sub-competence within the instrumental sub-competence of the PACTE model, namely “the ability to meet a number of learning outcomes: identifying the principles that lie at the basis of the use of corpora; creating corpora; using corpus-related software; and solving translation problems by using corpora” (p. 253).

Development of instrumental competence, including the use of documentation sources and electronic tools, is particularly relevant in PSIT, where translators need to manage different information sources in order to acquire sufficient understanding of the subject of a text and thus enable the accurate transfer of information. Given the importance of documentation in PSIT training to ensure production of a functionally adequate and acceptable target language text, we designed an activity focused on compiling and analysing monolingual virtual corpora to translate judgments issued in criminal proceedings. A virtual corpus is a collection of texts developed from electronic resources by the translator and compiled “for the sole purpose of providing information – either factual, linguistic or field-specific – for use in completing a translation task (Sánchez, 2009, p. 115). The compilation process would also help to develop our translation trainees’ technical skills.

Of the different major corpus types (Bernardini, Stewart, & Zanettin, 2003, p. 6), we found monolingual corpora especially useful for our task as the students needed to compile a corpus containing texts produced in the target language. The final monolingual corpus would thus provide them with information about idiomatic use of specific terms, collocations, and other syntactic and genre conventions of the legal language.

Our students attended two training sessions of six hours in total. In the first session, they were introduced to the main theoretical concepts in Corpus-based Translation Studies, the main documentation resources for PSIT (lexicographical databases, specialised lexicographical resources and specialised portals) and different word search strategies needed to take advantage of search engines and Boolean operators. In the second session, they learned the differences between the so-called *Web for Corpus* (WfC) and *Web as Corpus* (WaC) approaches and were shown how to use retrieval information software, such as *SketchEngine* and *AntConc*. They also learned the basic functions of both software programs (generating and sorting concordancing, identifying language patterns, retrieving collocations and collocation clusters, etc.). After this training session, the students were each asked to compile a monolingual corpus (British English) as part of the module on Legal Translation, to translate a judgment issued in Spanish criminal proceedings into English. They were also asked to investigate genre and lexical conventions and to use their ad hoc corpus to solve terminological and phraseological problems when translating.

### **3. Compiling an ad hoc corpus in PSIT**

The need for an initial determination of criteria for selection and inclusion is the starting point when designing and compiling a corpus. Our methodology was divided into three stages: source-text documentation, the compilation process and corpus analysis. In the first stage, we encouraged our students to read texts similar to the source text (in this case, a judgment passed by Spain's Supreme Court), which we provided to help them learn about the nature of this type of text and to familiarise them with the main linguistic and genre conventions. In

the second stage, the students needed to be able to locate different Internet-based texts to be included in their own corpus. To do so, they needed to put into practice what they had learned about Boolean operators in previous sessions, that is, to search for information using keywords (e.g. ‘appeal’, ‘constitutional rights’, ‘presumption of innocence’). It is of paramount importance at this stage to use very precise keywords – *seed words* – as filters, in order to exclude irrelevant information or ‘noise’. Institutional web pages, such as that of the British and Irish Legal Information Institute (BAILII), can be used to download and save complete texts, namely UK Supreme Court’s judgments. Students were also encouraged to use free software (i.e. *HTTrack*, *GNU Wget* or *Jdownloader*) so that they could automate the downloading process. As previously stated, “[o]nce the documents had been found and downloaded, the texts had to be converted to .txt files in order to be processed by corpus analysis software [like *AntConc*]. This task is especially necessary in the case of texts retrieved in .pdf format” (Lázaro Gutiérrez & Sánchez Ramos, 2015, p. 285). Finally, all documents were stored and the students were able to initiate an analysis of their materials.

The ad hoc corpora compiled by our students were highly useful in terms of all the terminological and idiomatic information they offered to aid the completion of the translation task. The students appreciated the immediate solutions their ad hoc corpora provided to different translation problems. For instance, they used the collocations and collocation cluster functions to identify the frequency of appearance of ‘direct evidence’ or ‘direct proof’ for the translation of ‘prueba indiciaria’. The cluster/N-gram function was particularly useful for checking the collocational patterns of the most problematic words, such as those followed by a preposition (e.g. *judgment on/in*), where students positively evaluated the contextual information their ad hoc corpora offered (see [Figure 1](#)).

The concordance function was also a very attractive resource for our students. A simple query generated concordance lines listed in KeyWord In Context (KWIC) format. For instance, students looked up the appropriate English term for ‘infracción de precepto constitucional’. The ad hoc corpus they had compiled offered a number of alternatives, such as ‘breach’, ‘infringement’, and ‘violation’, with ‘violation’ being the most frequent (see [Figure 2](#)).

Figure 1. Collocation cluster/N-gram function

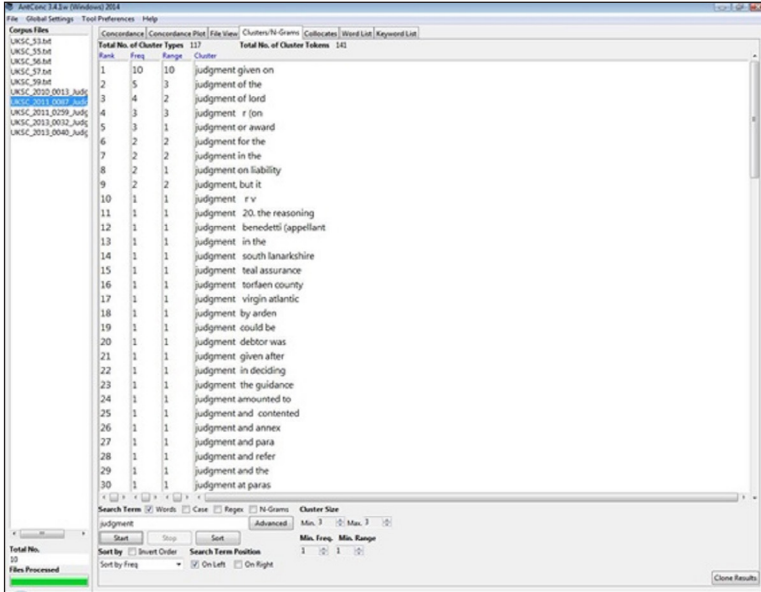
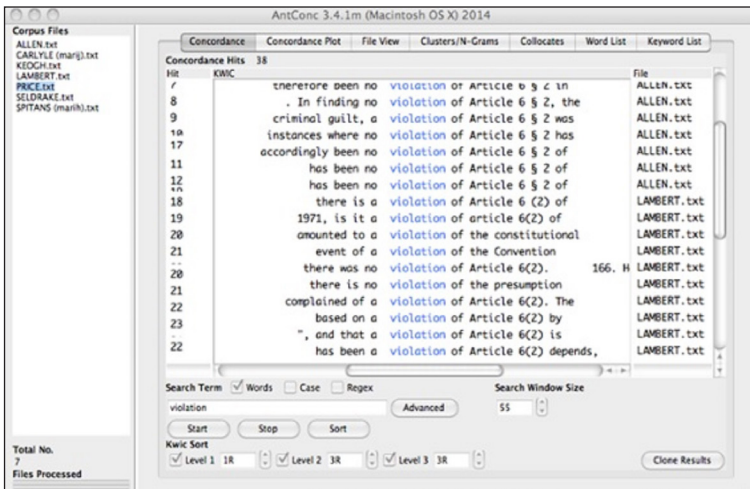


Figure 2. Concordance function



In general terms, the students' feedback was largely positive. They appreciated the usefulness of the different functions that a software tool such as *AntConc* can provide (e.g. frequency lists, collocates, clusters/N-grams, or concordancers). Compiling and using corpora made the students feel more confident in their technical skills and translation solutions. Altogether, corpus use was evaluated by our students as a valuable tool for developing their instrumental competence.

## 4. Conclusions

We have shown how we developed and exploited monolingual virtual corpora as a resource in the PSIT training environment. A corpus can be a valuable aid for specialised translation students, who can consult the corpus to acquire both subject field knowledge and linguistic knowledge, including information about appropriate (and inappropriate) terminology, collocations, phraseology, style and register. However, training is essential when compiling and using corpora, as this requires a variety of competences, both linguistic and technological. As we have commented, well-planned training on corpora and compiling methodology can contribute to the development of these competences, essential in the world of professional translation.

## References

- Aldea, P., Arróniz, P., Ortega, J. M., & Plaza, S. (2002). Situación actual de la práctica de la traducción y de la interpretación en la Administración de Justicia. In S. Cruces & A. Luna (Eds.), *La traducción en el ámbito institucional: autonómico, estatal y europeo* (pp. 85-126). Vigo: Universidade de Vigo.
- Bernardini, S., Stewart, D., & Zanettin, F. (2003). (Eds.). *Corpora in translator education: an introduction*. Manchester: St. Jerome.
- Biel, L. (2010). Corpus-based studies of legal language for translation purposes: methodological and practical potential. In C. Heine & J. Engberg (Eds.), *Reconceptualizing LSP. Online proceedings of the XVII European LSP Symposium 2009* (pp. 1-15).



- Bowker, J., & Pearson, J. (2002). *Working with specialized language: a practical guide to using corpora*. London: Routledge. Retrieved from <http://dx.doi.org/10.4324/9780203469255>
- Corpas, G., & Seghiri, M. (2009). Virtual corpora as documentation resources: translating travel insurance documents (English–Spanish). In A. Beeby, P. Rodríguez Inés, & P. Sánchez-Gijón (Eds.), *Corpus use and translating. Corpus use for learning to translate and learning corpus use to translate* (pp. 75-107). Amsterdam: John Benjamins. Retrieved from <http://dx.doi.org/10.1075/btl.82.07cor>
- EMT Expert Group. (2009). *Competences for professional translators, experts in multilingual and multimedia communication*. European Master's in Translation Website of the DG Translation of the European Commission. Retrieved from [http://ec.europa.eu/dgs/translation/programmes/emt/key\\_documents/emt\\_competences\\_translators\\_en.pdf](http://ec.europa.eu/dgs/translation/programmes/emt/key_documents/emt_competences_translators_en.pdf)
- Hurtado, A. (2007). Competence-based curriculum design for training translators. *The Interpreter and Translator Trainer*, 1(2), 163-195. Retrieved from <http://dx.doi.org/10.1080/1750399X.2007.10798757>
- Lázaro Gutiérrez, R., & Sánchez Ramos, M. d. M. (2015). Corpus-based interpreting studies and public service interpreting and translation training programs: the case of interpreters working in gender violence contexts. In J. Romero-Trillo (Ed.), *Yearbook of Corpus Linguistics and Pragmatics 3* (pp. 275-292). Springer International Publishing Switzerland. Retrieved from [http://dx.doi.org/10.1007/978-3-319-17948-3\\_12](http://dx.doi.org/10.1007/978-3-319-17948-3_12)
- Lee, D., & Swales, J. (2006). A corpus-based EAP course for NNS doctoral students: moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes*, 25(1), 56-75. Retrieved from <http://dx.doi.org/10.1016/j.esp.2005.02.010>
- Martin, C. (2011). Specialization in translation – Myths and realities. *Translation Journal*. Retrieved from <http://www.bokorlang.com/journal/56specialist.htm>
- Monzó, E. (2008). Corpus-based activities in legal translator training. *The Interpreter and Translator Trainer*, 2(2), 221-252. Retrieved from <http://dx.doi.org/10.1080/1750399X.2008.10798775>
- PACTE Group. (2003). Building a translation competence model. In F. Alves (Ed.), *Triangulating translation: perspectives in process oriented research* (pp. 43-66). Amsterdam: John Benjamins. Retrieved from <http://dx.doi.org/10.1075/btl.45>
- Prieto, F. (2011). Developing legal translation competence: an integrative process-oriented approach. *Comparative Legilinguistics – International Journal for Legal Communication*, 5, 7-21.

- Rodríguez, P. (2010). Electronic corpora and other information and communication technology tools. An integrated approach to translation teaching. *The Interpreter and Translator Trainer*, 4(2), 251-282. Retrieved from <http://dx.doi.org/10.1080/13556509.2010.10798806>
- Sánchez, P. (2009). Developing documentation skills to build do-it-yourself corpora in the specialized translation course. In A. Beeby, P. Rodríguez-Inés, & P. Sánchez-Gijón (Eds.), *Corpus use and translating. Corpus use for learning to translate and learning corpus use to translate* (pp. 109-127). Amsterdam: John Benjamins. Retrieved from <http://dx.doi.org/10.1075/btl.82.08san>
- Valero-Garcés, C. (2006). *Formas de mediación intercultural: traducción e Interpretación en los Servicios Públicos*. Granada: Comares.



Published by Research-publishing.net, not-for-profit association  
Dublin, Ireland; Voillans, France, [info@research-publishing.net](mailto:info@research-publishing.net)

© 2016 by Antonio Pareja-Lora, Cristina Calle-Martínez, and Pilar Rodríguez-Arancón (collective work)  
© 2016 by Authors (individual work)

New perspectives on teaching and working with languages in the digital era  
Edited by Antonio Pareja-Lora, Cristina Calle-Martínez, Pilar Rodríguez-Arancón

**Rights:** All articles in this collection are published under the Attribution-NonCommercial -NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Under this licence, the contents are freely available online as PDF files (<http://dx.doi.org/10.14705/rpnet.2016.tislid2014.9781908416353>) for anybody to read, download, copy, and redistribute provided that the author(s), editorial team, and publisher are properly cited. Commercial use and derivative works are, however, not permitted.



**Disclaimer:** Research-publishing.net does not take any responsibility for the content of the pages written by the authors of this book. The authors have recognised that the work described was not published before, or that it was not under consideration for publication elsewhere. While the information in this book are believed to be true and accurate on the date of its going to press, neither the editorial team, nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, expressed or implied, with respect to the material contained herein. While Research-publishing.net is committed to publishing works of integrity, the words are the authors' alone.

**Trademark notice:** product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

**Copyrighted material:** every effort has been made by the editorial team to trace copyright holders and to obtain their permission for the use of copyrighted material in this book. In the event of errors or omissions, please notify the publisher of any corrections that will need to be incorporated in future editions of this book.

Typeset by Research-publishing.net

Cover design and frog picture by © Raphaël Savina ([raphael@savina.net](mailto:raphael@savina.net))

ISBN13: 978-1-908416-34-6 (Paperback - Print on demand, black and white)

Print on demand technology is a high-quality, innovative and ecological printing method, with which the book is never 'out of stock' or 'out of print'.

ISBN13: 978-1-908416-35-3 (Ebook, PDF, colour)

ISBN13: 978-1-908416-36-0 (Ebook, EPUB, colour)

**Legal deposit, Ireland:** The National Library of Ireland, The Library of Trinity College, The Library of the University of Limerick, The Library of Dublin City University, The Library of NUI Cork, The Library of NUI Maynooth, The Library of University College Dublin, The Library of NUI Galway.

**Legal deposit, United Kingdom:** The British Library.  
British Library Cataloguing-in-Publication Data.

A cataloguing record for this book is available from the British Library.

**Legal deposit, France:** Bibliothèque Nationale de France - Dépôt légal: mai 2016.