



7th International Conference on Corpus Linguistics: Current Work in Corpus Linguistics:
Working with Traditionally-conceived Corpora and Beyond (CILC 2015)

E-learning practices in translation and interpretation: corpora as training platforms

Nataliya Godinho Soares Vieira*

CHAM-FCSH-NOVA University of Lisbon, Avenida de Berna, 26, 1069-061 Lisbon, Portugal

Abstract

To study different languages to be professional translator or interpreter is a challenging mission. It demands high-quality foreign language competences in a wide range of sectors. The abilities to immediately select from a variety of linguistic structures, accurately and precisely transmit specific terminology, cultural markers, complex or symbolical concepts, and constantly enlarge linguistic resources for expressing ideas – all these elements depend on the experience and require continuing enhancement. Along with the recent developments of methodology applied at translation courses, there have been spread positive responses on the corpus-based approach. This refers to verifying the hypothesis to implement corpora as training electronic platforms in translation and interpreting studies (TIS). The paper aims at presenting a set of didactic models that assume different types of English, Portuguese, Spanish, and Russian electronic corpora as required and effective means in the teaching-learning process. This recognizes the advantages of the corpus-informed support to translation trainings as well as expanding the professional interest in this domain.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of Universidad de Valladolid, Facultad de Comercio.

Keywords: corpus-based interpretation/translation trainings; English, Portuguese, Spanish, Russian corpora; corpus-informed activities.

1. Introduction

The present age of digital technology brings fundamental changes to translation studies: “the working practices of translators have been changed beyond recognition in terms of the access to many different kinds of knowledge that are afforded by the infrastructure of the internet” (Cronin, 2012: 8). Integrating high-tech resources in the teaching-

* Corresponding author. Tel.: +351968373270.
E-mail address: nataligsv@fcsh.unl.pt.

learning process, modern didactics significantly compensates the lack of materials for trainee translators. In this context appeared interest on electronic corpora in translation studies has caused by the significant reasons related to the advantages of large multifunctional and free available databases. Corpora application in the professional sphere and corpora use for learning to translate/interpretate are still gradually become visible in the academic curriculum of educational institutions (Corpas Pastor, 2008; Beeby, Rodríguez-Inés, and Sánchez-Gijón, 2009; Gallego Hernández, 2012). The main obstacle is that online corpora are in the process of development in various languages and they are very different from one another. Electronic corpora could be better involved in work, if they were more evolved. However, answering to the queries like “can translation be taught? If so, can it be taught through precepts, rules, principles? Or can it only be “taught” through doing it and getting feedback?” (Robinson, 2012: 87), the authors of corpus-informed didactic resources provide the opportunities to make the teaching-learning process active and natural.

Corpus-informed activities may be useful in a variety of teaching contexts in order to 1) provide authentic written and audio/visual contents, 2) carry out comparative analyses of original and translated texts; 3) familiarize learners with diachronic, synchronic and pragmatic aspects as integral parts of linguistic phenomena; 4) categorize the rates and frequencies of translation equivalents, etc. Analyzing corpus data for translation purposes, it is essential to compose some personal glossaries of frequent and relevant or infrequent but relevant patterns in the branch of interest. There are various types of corpora specialized in literary, scientific, technical texts, periodicals, transcribed audio and video interviews on specific topics. For instance, one of the most suitable sources for translation trainings is multidirectional parallel corpora that consist of a large number of sections that allow simultaneously search one term for each of the languages of translation. However, any corpora may serve for creation of training activities.

Corpora application in translation studies supports knowledge improvements and facilitates data processing stages: research, analysis, collection and new ways in form of wording. Corpus-informed didactic materials help learners to simulate the translator/interpreter practice under the variety of situations of a professional life.

There are some activities presented below that are based on the following available free online corpora:

- monolingual corpora: *Business Letter Corpus* (BLC), *Corpus of Spanish* (*Corpus del Español*), *Corpus of Portuguese* (*Corpus do Português*); *Corpus of Modern Spanish of the Royal Academy* (*CREA – Corpus de referencia del español actual*), *Russian National Corpus* (*Национальный Корпус Русского Языка*), *Reference Corpus of Contemporary Portuguese* (*CRPC – Corpus de Referência do Português Contemporâneo*); *Corpus of Contemporary American English* (*COCA*);
- bidirectional parallel corpora: *COMPARA* (parallel corpus of English and Portuguese), *Lingee*¹;
- multidirectional parallel corpora: parallel Corpus of the University of Vigo (*Corpus Paralelo CLUVI*);
- multimedia corpora: *English Language Interview Corpus as a Second-Language Application* (*ELISA*), *Scottish Corpus of Texts and Speech* (*SCOTS*); *Russian National Corpus* (*Национальный Корпус Русского Языка*).

2. Procedures: production stage

The effectiveness of the corpus-informed didactic materials depends on a variety of factors. Online corpora are databases that do not explain the regularities of the usage of linguistic items and structures. That is why teachers or other specialists, who are interested in designing of corpus-based activities for pedagogical purposes, have to determine the following work segments.

Analyzing teaching context.

¹ *Lingee* is the electronic translation tool, combination of an editorial dictionary and a multilingual search engine. *Lingee* is similar to bidirectional corpora.

Didactic models that include corpus-informed activities can combine a series of approaches to be adequate to a certain curriculum content and meet the necessary assessment criteria. Effective outcomes depend on how far the framework definitions and requirements definitions are taken into consideration.

Framing progress indicators.

At the intermediate stages, students have to demonstrate understanding of the main principles of the corpus-based research and be able to participate in a classroom and supervised autonomous corpus-informed translation studies. The advanced stage determines autonomous corpus-based research in a wide range of electronic corpora (monolingual, multilingual, parallel, multimedia, etc.), and the application of some elements of the corpus-based analysis for professional purposes.

Specifying the degree of difficulty.

The effectiveness of a task/activity/exercise depends on its degree of difficulty that can be based on four categories: “text-specific difficulties, translator-dependent difficulties, pragmatic and technical difficulties” (Nord, 2005: 168). The stages of the translation process such as “comprehension, deverbilization, reformulation, and justificative analysis” (Beeby Lonsdale, 1996: 46) may be taken into consideration within an acting performance.

Emphasizing the main objectives.

Introduced corpus-informed activities have to be applied with the emphasis on the main objectives of the program. For instance, reinforcement of vocabulary, syntactic structures, a certain style or discourse type, etc.

Enhancing motivation.

Finally, it is important to assume that motivation is a key factor that mediates learning. “From a neuroscientific perspective, when we are doing something, these mind-body states represent our motivation. We are likely to identify them by the emotion or mood most obvious to us at the moment, such as ‘I’m getting bored with reading this textbook’. Although our mind-body state may seem stable as we proceed with a task, in reality it is in a state of flux, diminishing, strengthening, or changing into another state. On the single page of a book or in the span of five minutes in a course, we may go from feeling inspired, to feeling frustrated, to feeling creative, and then inspired again” (Wlodkowski, 2008: 18). Therefore, planned activities have to be compatible with a variety of factors which determine learning process and influence learners’ effort.

3. Didactic models

3.1. *Exploring terminology and formal/informal discourse*

New developments bring new concepts as well as new ways of their representation. It is fundamental that within a translation course, learners gain self-confidence in transmitting terms and expressions related to a variety of fields, translating/interpreting and producing formal and informal contexts in a native language and a non-native language. The following examples demonstrate the series of the corpus-informed activities that are suitable for such purposes.

Activity 1 (context: Spanish-English translator trainings, intermediate level).

1. 1. In small groups read and discuss in Spanish the proposed contexts according to the table below (Table 1).

- A. “*Como cualquier otra inversión, la que se hace en bolsa tiene sus riesgos, por lo cual los inversores individuales suelen buscar el asesoramiento de los expertos, ya sea de agencias o de sociedades de valores, o bien de especialistas asesores concretos. Aparte, están las inversiones colectivas, que tienden a diluir el riesgo. De entre ellas, nos referiremos brevemente a las más notables: fondos de inversión, sociedades de cartera, y clubs de inversiones. Por obligación legal, los fondos de inversión mobiliaria trabajan a través de una sociedad gestora, y de un banco depositario, que, por lo general, es su propio promotor. Se constituyen sobre la base de la emisión de certificados de participación, cotizables en bolsa*” (retrieved from the CREA, key word: ‘depositario’).

B. “*Producto natural sería la penicilina, pero al emplear este análogo de nutriente se obtiene un compuesto modificado. Manipulación de rutas metabólicas específicas. A, B, C, son intermediarios metabólicos de la ruta que conduce a la molécula de interés D. E1, E2, E3, son las enzimas que catalizan las etapas de la ruta. (a): la línea discontinua refleja la retroinhibición que ejerce D sobre E1. (b): E4 es una enzima capaz de degradar D. (c): X es un producto que modifica la actividad de la enzima E5, cuyo sustrato y producto naturales son, respectivamente, G y H. (d): A*, B*, C*, D*, son los análogos correspondientes de A, B, C y D, respectivamente... Todas las posibilidades que se acaban de mencionar para la fermentación con cepas normales pueden aplicarse en la utilización de cepas mutantes. Pero, además, estas últimas pueden aportar ventajas que aquéllas no ofrecen. Es por ello por lo que su empleo presenta algunas características propias*” (retrieved from the CREA, key word: ‘aportar ventajas’).

Table 1. Text analysis

Género:
Emisor:
Receptor:
Función:
Tema:
Discurso:

Términos clave en español y sus equivalentes en inglés:
(The table is adopted from Gallego Hernández, 2012: 201)

Using the information from the table, orally present the contents of the two contexts in English.

1. 2. Go to the *Corpus del Español*, type *cheque* in a search string, indicating CHART in display, NEWSPAPER and 1900s in sections, and sort by RELEVANCE; click on a bar NEWS for context. In a new Word document, save the sentences that contain the combinations ‘verb+ *cheque*’ (for example, ...depositar el cheque en el banco..., ...reducir el impuesto al cheque..., ...girar un cheque en blanco..., etc.), and translate them into English.

1. 3. Go to the *Corpus of Contemporary American English (COCA)*, type *putt* in a search string, indicating CHART in display, NEWSPAPER in the both sections, click on a bar NEWS for context. In a new Word document, save the five phrases that contain *putt* in the roles of different parts of speech (for example, “*McIlroy carelessly knocked his next putt eight feet past the hole*”, “*Scott didn't have a birdie putt longer than 12 feet*”, “*But when I saw him miss that putt, I bailed on that idea*”, etc. Translate these phrases into Spanish. Discuss the translations with the group members.

Activity 2 (context: Portuguese-English translator trainings, intermediate level).

2. 1. Imagine that you are working at an international tourist information office. You have to send an apology letter to excuse your company for providing some kind of service not in time. You suppose to use some formal expressions related to the subject, but you don't know exactly how to formulate the phrases. In this case you can seek for the information in the *Business Letter Corpus*. First of all, as your apology letter requires apologizing for the inconvenience, type *inconvenience* in a search line (search type: start with; sort type: unsorted). Choose a necessary phrase to start your apology letter. Depending on a context, introduce the key words in the corpus for retrieving as more concordances as possible that are suitable for your letter (for example, ‘compensate’, ‘offer’, ‘discount’, ‘understanding’, etc.). In little groups compose your letters and exchange them with your colleagues. Analyze the results. Translate into Portuguese the letter of your colleague, send it to your colleague again, and analyze the results together.

2. 2. Go to the *Corpus do Português*, find the frequency and distribution of the synonyms of [=reluzente].[j*] (in NEWS and FICTION), [=fraco].[j*] (in ACADEMIC and ORAL), and [=alegre].[j*] *] (in FICTION and NEWS). Display: LIST, POST LIST: NOUN for all words. In small groups compare the frequencies of the usage of the synonyms and discuss the possible reasons for this. Retrieve three complete sentences for each synonym and translate them into English. Analyze the translations in little groups.

2. 3. Go to the *Corpus de Referência do Português Contemporâneo (CRPC)*, open *Spoken Corpus Mozambique 1986-87*, retrieve the concordances for “machimbombo”. After analyzing the contexts, give a detailed explanation of the significance of this word in English.

Activity 3 (context: English-Portuguese translator trainings, advanced level).

3. 1. Translate the following contexts from English into Portuguese:

- A. *The water overflowing from the great southern river has flooded large areas of land, giving rise to the planting of orchards and allotments, but above all creating one of the planet's most productive habitats: the salt-marsh.*
- B. *The contrasts in the landscape are in evidence in the humanised and fertile landscape with plantations, fruit trees and allotments and, also, the existence of a wilder landscape with heathers and brooms which is intersected by water lines* (retrieved from the *TURIGAL-CLUVI*, key word: *allotment*).
- C. *Situated in the Algarve's eastern region, this complex includes a pavilion, three multi-purpose sports halls, two gymnasiums, four tennis courts, a football stadium with a grass pitch, two grass football pitches for training purposes, an athletics track with a synthetic surface, a training area for throwing events and a cross country course.*
- D. *For centuries these two sources were a meeting place for the women of the village who would fill their water pitchers and do their washing.*
- E. *Football is one of the sports for which the Algarve is equipped with some of the best facilities, namely full-sized grass pitches with the appropriate official measurements, training pitches, gymnasiums, multi-purpose complexes and keep-fit circuits* (retrieved from the *TURIGAL-CLUVI*, key word: *pitch*).

Go to the *CLUVI*, open *TURIGAL (corpus de turismo português-inglês)*, type *allotment*, in a search line ‘palavra ou expresión en inglés’, and compare your translations with the given ones based on the questions: what are the differences? what translations have higher quality? why?

Repeat the same procedure with a key word *pitch*.

3. 2. Go to the *Lingee: English-Portuguese*, insert an English key word *pool* and find its Portuguese translation equivalents on the top of the table in *Dicionário inglês-português* (for example, *piscina, sinuca, tanque, lago, agrupamento, jogo de bilhar*, etc.). In the *Lingee* and *COMPARA*, find two examples for each equivalent among the parallel concordances in Portuguese. Analyze their translations in full contexts.

Do the same procedures with a key word *snap*. Analyze the translations in full contexts.

3. 2. Transmitting figurative language

Cultural markers, idioms, proverbs, connotative words and expressions, etc., belong to complex figurative language that has to be regularly investigated even by professional translators/interpreters. Well-formed translation equivalents of culturally determined metaphors, comparisons or euphemisms underline the quality of producing work and language-learning experience of a specialist. During practical seminars for a translation course, teachers can introduce a large number of corpus-based tasks for figurative language analyses (see Table 2). Such analytical procedures can reveal a variety of language patterns that are not familiar to not native speakers. It helps learners to reflect on possible ways of their translation.

Table 2. An abstract from the task *Finding idioms* (Anderson & Corbett, 2009: 54)

The BNC [*British National Corpus*: <<http://www.natcorp.ox.ac.uk/>>] contain the following six instances of *gander* forming part of the expression *What is sauce for the goose is sauce for the gander*.

...what is sauce for the goose must be sauce for the **gander**. If this is the justification we seek for abandoning...

Sauce for the goose but not, apparently the **gander**! There is very little likelihood of confusion arising...

What's sauce for the goose is sauce for the **gander**. As for covering up the work...

...sauce for the goose was likely to be sauce for the **gander**. Because they saw the world as composed of groups...

... what's cash for the goose is cash for the **gander**...

What's sauce for the goose is sauce for the **gander**. And er if it's appropriate...

Which of these do you think are standard uses of the expression? Which are creative uses?
What are the implications of your findings for automatically retrieving examples of idioms and semi-fixed expressions from a corpus?

Now look for other idioms in the BNC. For example, *when in Rome...don't count your chickens, dead as ashort of a ...* (as in *a few sandwiches short of a picnic*). Do these behave in the manner you expect? What problems do you encounter in finding examples?

3.3. On teaching of interpretation

Interpreting practices require high quality training opportunities that determine investment in technological equipment (Díaz-Cintas, 2008). The trainings based on professional software or on tailor made activities are beyond compare. However, the advantage of multimedia corpora is that the professionals who carry out interpreting studies can familiarize learners with speech and cultural components and provide a lot of benefits from corpus-informed perception tasks and simulation activities for interpretation. There are some examples.

Activity 1 (context: Russian-English interpreting trainings, intermediate level).

Go to the Russian National Corpus (*Национальный Корпус Русского Языка*), open the Multimedia corpus (*Мультимедийный корпус*), and type a key word *приветствовать* in the field *лексико-грамматический поиск*. Scroll down the interface, in the option *жесты*, click on *эмоции*, and mark all of the variants. They are *нет / слезы / смех / улыбка*. Scroll down the interface and click on *искать*. On the left side in a new window you can find four video files, and, on the right side, their transcriptions. Watch all the fragments that demonstrate the meeting-greeting situations 1) in London, 2) in a park, 3) at the entrance to an apartment, and 4) in a circus. In small groups interpret of the discourses and their speech components into English (for subtitling and voice-over formats). Present and compare your interpretations with other groups.

Activity 2 (context: English-Spanish interpreting trainings, advanced level).

2. 1. Go to the *English Language Interview Corpus as a Second-Language Application (ELISA)*. For watching the video files, it is necessary to download free *RealPlayer*. Scroll down the page and find the following interviews: 1) “Working on a computing services helpdesk” (the interview with Linda, who manages the computer help desk at Oxford University), 2) “Music is my life - a guitar teacher from Edinburgh” (the interview with Phil, a guitar teacher from Edinburgh), 3) “The Santa Fe Chamber of Commerce” (the interview with the (former) president of the Chamber of Commerce in Santa Fe, New Mexico, US). Watch the interviews. In small groups characterize the pronunciations of the English, Scottish and American speakers: how do they articulate the vowel sounds? who of them do use more weak forms? what are the differences in their intonations when they make statements or exclamations? Train to interpret of the discourses from English into Spanish in a voice-over format. During your interpretation, ask your group members to analyze your work and give a feedback (they can use full transcriptions of the interviews).

2. 2. Go to the *Scottish Corpus of Texts and Speech (SCOTS)*, scroll down the page and find the “Conversation 33: Two female students on Scotland, and University education”. Watch the video and transcribe the dialog. Compare your transcription with the audio transcription from the corpus. Prepare Spanish subtitles for each transcribed line.

4. Final remarks

Translation and interpreting studies predetermine insights into the nature of authentic contexts. It directs attention to the search for the sources of information and methodology that support this compulsory condition. Corpus-based activities carry out the learner-centered principles and are based on free tech resources. By developing skills to work with electronic corpora, the specialists can design appropriate didactic materials that suit to a variety of learning needs. However, as online corpora are still limited in their functions, there are some challenges to overcome. The examples of activities that are given above demonstrate particular advantages of the corpus-based approach compared to other approaches – time-consuming process of preparation and consultation, free access to large databases, engaging in thought-provoking learning procedures. In fact, coherent explanations, adapting course materials, motivating instructions, and enhancing meaning in learning practices are key pillars for professional involvement of the corpus-informed materials in the context of TIS.

References

- Anderson, W., Corbett, J. (2009). *Exploring English with Online Corpora*. Basingstoke: Palgrave MacMillan.
- Beeby Lonsdale, A. (1996). *Teaching Translation from Spanish to English: Worlds beyond Words*. Canada: Ottawa Press.
- Beeby, A., Rodríguez-Inés, P., Sánchez-Gijón, P. (Eds.) (2009). *Corpus Use and Translating: Corpus Use for Learning to Translate and Learning*. Amsterdam-Filadelfia: John Benjamins.
- Corpas Pastor, G. (2008). *Investigar con corpus en traducción: los retos de un nuevo paradigma*. Frankfurt am Main: Peter Lang.
- Cronin, M. (2012). *Translation in the Digital Age*. London: Routledge.
- Díaz-Cintas, J. (Ed.) (2008). *The Didactics of Audiovisual Translation*. Amsterdam: John Benjamins.
- Gallego Hernández, D. (2012) *Traducción económica y corpus: del concepto a la concordancia. Aplicación al francés y al español*. Alicante: Servicio de Publicaciones de la Universidad de Alicante.
- Nord, C. (2005) *Text Analysis in Translation: Theory, Methodology, and Didactic Application of a Model for Translation-Oriented Text Analysis*. Amsterdam, NY: Rodopi.
- Robinson, D. (2012). *Becoming a Translator: An Introduction to the Theory and Practice of Translation*. NY: Routledge.
- Wlodkowski, R. (2008). *Enhancing Adult Motivation to Learn: A Comprehensive Guide for Teaching all Adults*. San Francisco: Jossey-Bass A Wiley Imprint.

Online corpora

Business Letter Corpus (BLC): <<http://www.someya-net.com/concordancer/>>.

COMPARA: <<http://www.linguateca.pt/COMPARA/psimples.php>>.

Corpus del Español: <<http://www.corpusdelespanol.org/>>.

Corpus do Português: <<http://www.corpusdoportugues.org/x.asp>>.

Corpus Paralelo CLUVI: <<http://sli.uvigo.es/CLUVI/index.html>>.

Corpus de Referência do Português Contemporâneo (CRPC): <<http://www.clul.ul.pt/pt/recursos/183-reference-corpus-of-contemporary-portuguese-crpc>>.

English Language Interview Corpus as a Second-Language Application (ELISA): <http://www.uni-tuebingen.de/elisa/html/elisa_index.html>.
Linguee: <<http://www.linguee.com>>.
Real Academia Española: Banco de datos (CREA). Corpus de referencia del español actual: <<http://www.rae.es>>.
Russian National Corpus: <<http://www.ruscorpora.ru/en/index.html>>.
Scottish Corpus of Texts and Speech (SCOTS): <<http://www.scottishcorpus.ac.uk>>.