

## DISCOVERY LEARNING IN THE LANGUAGE-FOR-TRANSLATION CLASSROOM: CORPORA AS LEARNING AIDS

Silvia Bernardini\*  
University of Bologna

**Abstract:** This contribution reviews the idea of discovery learning with corpora, proposed in the 1990s, evaluating its potential and its implications with reference to the education of translators today. The rationale behind this approach to data-driven learning, combining project-based and form-focused instruction within a socio-constructivist inspired environment, is discussed. Examples are also provided of authentic, open-ended learning experiences, thanks to which students of translation share responsibility over the development of corpora and their consultation, and teachers can abandon the challenging role of omniscient knowledge providers and wear the more honest hat of “learning experts”. Adding to the more straightforward uses of corpora in courses that aim to develop thematic, technological and information mining competences – i.e., in which training is offered in the use of corpora as professional aids –, attention is focused on foreign language teaching for translators and on corpora as learning aids, highlighting their potential for the development of the three other European Master’s in Translation (EMT) competences (translation service provision, language and intercultural ones).

**Keywords:** Data-driven learning. Discovery learning. Translator education. EMT competences.

---

\* Silvia Bernardini: PhD in Translation studies from the University of Middlesex, UK (2007). MPhil in English and Applied Linguistics from the University of Cambridge, UK (1998). Laurea in Translation, University of Bologna, Italy (1997). Professor of English Language and Translation, Department of Interpreting and Translation, University of Bologna, Italy. E-mail: [silvia.bernardini@unibo.it](mailto:silvia.bernardini@unibo.it)



## **APRENDIZAJE POR DESCUBRIMIENTO EN LA CLASE DE LENGUA PARA TRADUCTORES: LOS CORPUS COMO HERRAMIENTAS DE APRENDIZAJE**

**Resumen:** Este artículo revisa el concepto de aprendizaje por descubrimiento con corpus, propuesto en los 90, y evalúa su potencial e implicaciones en la formación de traductores actual. Se discuten los fundamentos que subyacen a esta aproximación al aprendizaje basado en datos en la que se combinan la formación basada en proyectos y la enseñanza centrada en la forma en un entorno de inspiración socioconstructivista. Se aportan ejemplos de experiencias de aprendizaje auténticas y de final abierto, gracias a las cuales los estudiantes de traducción comparten responsabilidad sobre el desarrollo y consulta de corpus, y los profesores pueden abandonar el exigente papel de omniscientes proveedores de conocimiento para pasar a ser “expertos en aprendizaje”. Aparte de los usos más evidentes de los corpus en cursos en los que se intenta desarrollar competencias temáticas, tecnológicas y de minería de la información (es decir, cursos en los que se ofrece formación en el uso de corpus como herramientas profesionales), se pone la atención en la enseñanza de lenguas para traductores y en los corpus como herramientas de aprendizaje, a la vez que se subraya su potencial para el desarrollo de otras tres competencias de la red de Másteres Europeos en Traducción (EMT), que son la competencia lingüística, intercultural y en prestación de servicios de traducción.

**Palabras clave:** Aprendizaje basado en datos. Aprendizaje por descubrimiento. Formación de traductores. Competencias EMT.

### **1. Introduction**

The aim of this contribution is to go back to the idea of discovery learning from corpora, re-evaluating its potential for the education of translators in the light of the changes in attitudes and priorities that have occurred in the last fifteen years within translation schools. I will refer mainly to the context I am more familiar with, the Italian one, but the general points should apply more widely, not least because of the convergence of educational systems in the EU resulting from the establishment of the European Higher Education Area. Discovery learning activities will be compared

with the typical corpus-based activities that still appear to be more mainstream in the translation curriculum. The potential (as well as limits) of the former will be compared with those of the latter, and in the light of the common framework of competences adopted by the European Master's in Translation (EMT) Network.<sup>1</sup> In so doing I hope to stimulate reflection and discussion on a (still) relatively neglected use of corpora for translator education as well as, more generally, on the need for closer attention to ways of developing those transversal competences that are more intangible, therefore more easily overlooked, and yet essential to becoming a translator.

## **2. What is Discovery Learning (and Why Might it be Relevant to Translator Education)?**

Pedagogic concerns have constantly been central to corpus linguistics – stretching as far back as Palmer's *Second Interim Report on English Collocations* (1933) and driving forward the milestone *COBUILD* project (Sinclair, 1987). Yet the focus has largely been on the insights that learners can obtain from corpora, either through the mediation of experts (teachers, textbook authors, lexicographers) or by directly consulting them, as one would a dictionary. And rightly so: the corpus revolution (Hanks, 2012) could not but affect approaches to foreign language learning as well. The “natural alliance between linguistic description and pedagogic prescription” (Seidlhofer, 2003,78) has also been criticized, however, on the grounds that the norms of a target user community, as observed in a corpus, may not be an appropriate model for language learners (Widdowson, 1991), and that descriptive insights should merely inform, but not determine what is taught in a foreign language classroom.

---

<sup>1</sup> The European Master's in Translation Network, originating from an initiative of the European Commission's Directorate-General for Translation (DGT), provides a forum for European postgraduate programmes in translation. It currently has 64 members.

Discovery (or exploratory) learning with corpora attempts to bridge these positions by making the most of corpora as learning aids rather than as sources of descriptive evidence only. This approach is closely associated with the name of Tim Johns and with his work on data-driven learning (Johns, 1991). Working as a lecturer in EFL at the University of Birmingham in the '80s and '90s, he witnessed the coming of age of the corpus paradigm, but was also ideally positioned to approach it from the perspective and with the priorities of an applied linguist. His memorable statements (*Research is too serious to be left to the researcher; Each student a Sherlock Holmes*) describe, in a nutshell those pedagogic practices that rely on “the hands-on use of authentic corpus data [...] for inductive, self-directed language learning [...] empowering learners to explore language corpora and come to their own conclusions (Boulton 2011,571). Through the metaphor of *the learner as traveller*, Bernardini (2004a) further emphasized the importance of focusing on the learning experience, the voyage, rather than its destination: whether inductive learning from corpora is more effective than explicit instruction for the acquisition of specific language structures is, in a sense, besides the point, as is the descriptive accuracy of the observations made by learners. Instead, discovery learning as described by Bernardini (2004b) values and attempts to optimize exposure to authentic language data, in the course of open-ended information-gap activities that provide the bases for collaborative assignments such as academic articles or mock-conference presentations. This approach appears to be potentially well-suited for translator education, for reasons that can only briefly be summarized here.<sup>2</sup>

First, being open-ended and self-directed, discovery learning activities are inherently learner-centred, favouring autonomy and having substantial transfer potential, as students practice and develop routines for analyzing corpus data that will no doubt also

---

<sup>2</sup> A recent insightful discussion of the value of data-driven learning for second language learning in general can be found in Flowerdew (2015).

come in handy in their profession. While autonomy and learner-centredness are widely recognized as valuable in second language learning – at least since Nunan (1988), they are especially important for learners at advanced levels, on their way to becoming language professionals. Many years of language instruction may have instilled in these learners the (generally deceitful) presumption that they do not require further language input; at the same time, they make it difficult for teachers to identify common areas that all students in a class need to work on. Instead, project-based discovery learning tasks using corpora are “fully adaptable to the learner’s individual needs and preferences [,] where the learner has an ability to select from an unrestrictive range of responses, or even to come up with responses not envisaged by the teacher” (Leech, 1997, 11-12).

Second, discovery learning is an inductive process that is meant to favour the *noticing* (Schmidt, 2010) of patterns, i.e. correlations between form and meaning in language performance, or *preferred ways of putting things* (Kennedy, 1992). No translator can afford to ignore the pervasiveness of collocational patternings, for instance, and the role they play in disambiguating word senses (Sinclair, 1996), making texts fluent and natural (McCarthy and O’Dell, 2006), and providing a springboard for irony and creativity (Louw, 1993; Kenny, 2001). Nor can they afford to ignore the ways in which the context of situation and context of culture affect such patternings (Halliday and Mathiessen, 2014). Register- and genre-specific corpora are ideal resources to factor in the contextual dimension.

Third, discovery learning is a form of communicative, or more specifically “situated”, learning (in the Vygotskian sense), whereby knowledge is “co-constructed through collaborative dialogue and negotiation with guidance and support mediated by the teacher or student” (Flowerdew, 2015, 19). By working in teams for the solution of open-ended linguistic problems (including the collection of data, the drafting of report and the delivery of presentations), a social context is created in which learning can take place and at the same time collaborative working methods can develop. This view is very much in line with Kiraly’s (2000, 184-

185) recommendations concerning the education of translators in a social constructivist, “community-of-practice”-inspired approach:

Implementing truly communicative foreign language instruction in translator education programmes would be an excellent first step towards turning translator education into a practice-oriented enterprise. It would encourage students to begin taking responsibility for their own learning [...] [i]t would help create a spirit of community within the institution and break the mould of the transmissionist model of teaching, with students no longer being treated like empty vessels that need to be filled with knowledge.

Before moving on to corpus-based discovery learning activities that could contribute to making translator education more practice-oriented, let us briefly appraise the current situation.

### **3. Corpora in Translator Education: What Role(s) so far?**

Based on the undisputed potential of corpora as sources of linguistic facts, several researchers have argued that they could (or should) become part of a translator’s toolkit, alongside dictionaries, translation memory software and the like. The general point made is that “language corpora, if selected and used appropriately, are able to provide more abundant and reliable information to the translator than traditional reference tools” (Aston 2009: ix). Much could be said in support of this view, but expounding on it would take us too far afield (the interested reader can find insightful reflections and teaching suggestions in the various CULT proceedings volumes, e.g. Beeby et al, 2009).

If we agree that translators need to use corpora, then translation students need to *learn* to use corpora. Such is Sánchez-Gijón’s argument concerning specialized translation: “[g]iven the importance of documentation in the translation process, it should

not be neglected in translator training programmes” (2009, 111), and “[i]n order to make full use of the documentary resources available to translators today, trainers should be considering the methodologies offered by corpus linguistics (ibid, 114). Rodríguez-Inés (2009, 129) makes the point explicitly when she proposes “one specific instrumental sub-competence, namely the ability to use electronic corpora adequately in order to solve translation problems in an appropriate manner”.

This approach appears to be rather mainstream - and very promising. In terms of the EMT competence framework (EMT expert group 2009), corpus work of this kind could indeed contribute to the development of three of the six recommended competences, i.e. *thematic* (by learning to search for information and develop one’s knowledge of specialist fields), *technological* (by learning to use translation-related software, in this case mainly for documentation purposes) and *information mining* competence (by developing strategies for extraction of terminology and phraseology, evaluation of sources etc.).

I will provide a short example of a teaching scenario that I believe is rather representative of this approach. In the Master’s in Specialized Translation of the University of Bologna, corpora are used mainly in the *Translation methods and technology* course (worth 15 ECTS credits and including a module on *information mining and terminology* alongside CAT/localization and MT/post-editing) and the *Specialized translation courses* (40 ECTS credits altogether, including two modules on *technical and specialized translation (B > A and A > B)* alongside translation for the publishing industry and multimedia translation/AVT). The assumption is that students learn the technical and methodological skills needed in the *Translation methods and technology* course, and then apply them to actual translation tasks in the *Specialized translation courses*. Both courses are taught in a computer lab, and have a practical/professional orientation.

A typical set of activities carried out across the two courses could involve first of all building a bilingual corpus for the simulation

of a terminology or translation task (requiring, among others, decisions on criteria for text selection, evaluation of sources, file format and metadata), and extracting typical terms and phrases (requiring technical and methodological know-how in using a corpus query tool and the ability to search for and interpret data – devising queries, sorting results, detecting patterns, discarding noise and so forth). These activities would be coherent with the instrumental orientation of a learning activity focusing on methods and technology for translators. Within a specialized translation course one could then shift the focus back to understanding and producing texts, putting to good use the knowledge and skills previously honed.

Introducing corpora to future translators within learning activities with a clear professional orientation may be the most obvious choice for lecturers won to the cause, but it is not necessarily the best choice for learners who still need to be convinced of the return on investment. As suggested by Aston (2009, ix-x), “corpus use is anti-economic in the short term, and [therefore] has not yet become widely established among professional translators”. The very same students who are perfectly happy to toil over translation memory software – particularly if industry-leading – are seldom engaged by corpora and concordancers, which “[u]nlike translation memory or machine translation systems, [...] do not instantly present a preferred candidate for the user to accept, modify or reject” (ibid.).

Thus, as we attempt to convince learners of the potential of corpora as professional aids, we have to make sure that they are able to appreciate such potential and that they possess the skills for critical thinking and cognitive processing necessary to “overcome the difficulty of simultaneously handling both the linguistic complexity of the material and the conceptual complexity of the tool” (Molés-Cases and Oster, 2015, 207). These priorities cannot be pursued in instrumentally-focused and/or professionally-oriented courses that attempt to simulate working conditions. They require an academic orientation.



#### 4. Moving Forward: Project-based Discovery Learning with Corpora

Most second cycle degree courses in translation that I am familiar with cover, along with the more professionally-oriented subjects, such as specialized translation and translation technology, other subjects that are meant to fine-tune two other competences recognized by the EMT expert group, namely *language* and *intercultural* competences (EMT expert group 2009). In Italy, these typically come under the headings of *Culture and literature*, and especially *Language and linguistics* of the A/B/C languages – the Master's offered by the University of Bologna currently devotes as many as 24 ECTS credits to these contents).

Even though the teaching of language(s) and linguistics to translation students is inherently different from general language teaching, even at advanced levels (Bernardini, 2004b, Kiraly, 2000), traditionally it has not been focused upon in the literature on translator education. The situation does not seem to have improved noticeably since 2000, when Kiraly (2000, 181) pointed out that “[t]he wealth of articles, monographs and conferences on translation studies [...] is marked by a virtual absence of contributions dealing with the role of second language learning and teaching in translator education”. The question then is, how can activities of corpus building and use be employed in language and linguistics courses to favour the acquisition of language competence of the sort needed by translators, and the development of sociolinguistic and textual competences founded on intercultural awareness? In the remainder of this section I will describe my own attempt at answering this question.

In academic year 2014/15 I taught a course in English language and linguistics to first year students enrolled in the University of Bologna MA in Specialized translation (6 ECTS credits, two semesters, one 2-hour class per week, in a lecture room with pc and projector). Coming from many years of experience teaching translation and translation technology, this was quite a challenge, since the students' attitudes and expectations were rather negative,

motivation was low, and metalinguistic competence about the English language rudimentary.

The course had two parts, the first more theoretical and teacher-centered, in which I introduced the building blocks of systemic-functional linguistics (SFL) applied to the English language,<sup>3</sup> and the second more practical and project-centred, in which students in teams were asked to carry out an original piece of corpus-based research. They were free to choose one linguistic feature of their choice and a text type of their choice, as long as the study compared instances of “constrained communication” (Lanstyák and Heltai, 2012) involving English and Italian. More specifically, they were to focus on *native* vs. *non-native* (“*English lingua franca*”, or “*ELF*”) texts produced by native speakers of Italian, or texts *translated* from Italian into English vs. *non-translated* (“original”) English texts. In corpus terms, the perspective to be adopted was monolingual comparable.

The project-based part took up eight meetings/weeks, and followed the structure set out in table 1. Evaluation for this part of the course was limited to the final group assignment (the academic paper), and was worth 50% of the final mark.

---

<sup>3</sup> The topics covered were Units and constituency, Mood, Modality, Grammatical metaphor, Transitivity, Clause complex, Theme, Cohesion, Genre and Register. The textbook adopted was Suzanne Eggins’s very accessible *Introduction to systemic functional grammar* (London: Continuum, 2004, 2<sup>nd</sup> ed.). This part of the course was evaluated through an oral end-of-course examination (worth 50% of the final mark), aiming to ascertain that the students understood the basic linguistic notions and were able to describe them using appropriate metalanguage.

**Table 1:** Structure of the project-based part of the course

Week	Activity (general)	Activity (detailed)
1	Corpus building	Identification and demarcation of text type and topic Operationalization criteria for text inclusion/exclusion Selection of representative texts Conversion to simple text format and storage
2	Hypothesis development	Identification, operationalization and retrieval (with Antconc)* of relevant (sets of) linguistic feature(s) for comparison
3	Hypothesis testing	Comparison of quantitative data across differently sized corpora (normalization, $\chi^2$ )
4-6	Presentation (conference format)	20' talk with slides (description of genre/register, literature review, method, results, discussion, limits, further work), followed by 10' of discussion
7-8	Paper write up	About 3,000 words, following lecturer guidelines and general conventions for academic writing in linguistics

\* <http://www.laurenceanthony.net/software/antconc/>

There were 13 groups in two classes, ranging in size from 3 to 5 members. Each group chose a name and a logo for themselves. The topics and text types each group ended up working on for their projects are listed in Table 2.

**Table 2:** Groups and project information

Group	Linguistic feature	Text type	Type of constrained communication**
Chocolate	Involvement markers	EU speeches	ELF vs. Native
Coffee breakers	Complexity markers and passive voice	Museum websites ( <i>history</i> sections)	Translated vs. non-translated
Golden angels	Cohesive conjunctions	MA dissertation introductions	ELF vs. Native
Hope	Hypotaxis	Company mission statements	Translated vs. non-translated
Octopus	Transitivity/process types	Product launch press releases	ELF vs. Native
Platypus	Parataxis, hypotaxis and embedding	Academic course descriptions	ELF vs. Native
Rainbow	Cohesion and language complexity	Patents	ELF vs. Native
Snoopy	Transitivity/process types	Wine descriptions	Translated vs. non-translated
Spectacles	Personal and impersonal approach	PhD thesis abstracts	ELF vs. Native
Tiramisu	Connector usage (cohesion and clause complexing)	PhD thesis conclusions	ELF vs. Native
Tupperware	Ideational Grammatical Metaphor/nominalization	PhD thesis abstracts	ELF vs. Native
Wallaby	Nominalization	Food company websites ( <i>history</i> and <i>about us</i> sections)	Translated vs. non-translated
Walmart	Mood adjuncts and interpersonal grammatical metaphor	Letters to shareholders	ELF vs. Native

\*\* Determining whether a text is written by an ELF author directly in English or whether it was translated into English is far from an easy task. While it was not always possible to find evidence in one sense or the other, the students were sensitized to the issue and asked to make a hypothesis, try to (dis)confirm it, and argue

for it. Notwithstanding the uncertainty of the outcome, it seems important for students of translation to be aware of (and reflect on) the more and more blurred dividing line between the two types of communication.

The uncorrected abstract in Table 3 below, by the *Wallaby* group, is quite representative of the kind of corpus work the students carried out for their projects and the insights they gained.

**Table 3:** An example abstract

<p>Nominalization and web texts: A comparative analysis of native and non-native food company websites.</p> <p>Nowadays, English is considered the lingua franca of the web and many Italian food companies have their web page translated into English in order to communicate to the widest possible public. Nominalization is very common in Italian and may be transferred in English translations of such texts. However, previous work (Hervey et al., 2000) has shown that the way nominalization is used varies substantially between English and Italian and has highlighted the different usage of this linguistic feature by English natives and non-natives. How nominalizing affects the features of a specific genre and register has not been thoroughly investigated yet. This paper compares the use of nominalization in pages either called "History" or "About us" of food companies websites and aims to prove that nominalization is more common in texts translated from Italian than in native texts. The study was based on two corpora: 25 native texts and 25 translated texts. All occurrences were counted in both corpora. Nominalization was found to be exploited to a greater extent in translated texts. This is probably due to how translators calque nominalization when translating from Italian into English. On the basis of these results, we can conclude that nominalization affects the structure and readability of a genre such as the web text, which differs from conventional writing in length, perception and purpose. This research intends to be a useful resource for translators from Italian into English. It might also be a starting point for further studies of ELF website texts, particularly the ones that have been translated from Italian.</p>
--

If we were to evaluate this piece of academic research writing in terms of scientific soundness and argumentative rigour and elegance, we could point out several flaws. Yet the objective of the course is not that of producing accurate linguistic descriptions based on corpora; rather, it is to create a learning setting in which students are motivated to explore corpora on their own terms, with scaffolding as required. As can be gleaned from the abstract, these students had first of all to come up with a linguistic issue that could be problematic for translators from Italian into English (combing the literature and reflecting on their own experience); they then had to agree among themselves on corpus construction parameters and evaluate single specimens for inclusion; once the corpus was in place, they had to search for ways of identifying nominalizations in corpora; and once the data had been obtained, they had to normalize them and evaluate the statistical significance of the difference observed. Each step required “not only technical skill [...], but above all critical thought” (Aston 2009, ix). By downplaying the importance of finding solutions quickly and easily, and instead opening up a space for deeper (and more leisurely) exploration of corpus data, project-based discovery learning activities of this kind can provide a suitable growing environment for those skills and competences that students and future translators need (if they are to use corpora to their advantage), but that have so far proved difficult to develop within translation practice or technology courses.

While I firmly believe that this type of corpus work is indispensable for future translators, there is no denying the many challenges it raises. In the context of the course described in this contribution, student feedback consistently suggested that the perceived workload of the workshop part was excessive (while there were virtually no complaints about the study workload for the background/theory part). Furthermore, they complained that not enough guidance was offered by the teacher, and suggested that more corpus analysis (as well as essay writing) work could be carried out in class, so as to have easier and constant access to expert

advice - in short a safety net. Although this is in a sense a natural reaction to an unfamiliar and more demanding learning approach, one should not underestimate the risk of overload, resulting in resistance to involvement and uptake. Gauging the right amount of feedback and support is arguably one of the greatest challenges for educators adopting a pedagogy of discovery.

With these caveats in mind, in the final section below I would like to go back to the wider picture, and reflect on the promises of discovery learning activities not just for learning to use corpora, but for the education of future translators at large.

### **5. Summing up: Why Translator Educators should Embrace a Pedagogy of Discovery**

Project-based activities like those I described in the previous section are undoubtedly challenging for lecturers and students alike. The former have to provide constant monitoring and substantial feedback, and be prepared to cope with the unexpected, since “the multiplicity of possible results and interpretations makes the outcome of tasks difficult to predict” (Molés-Cases and Oster 2015, 205). The latter, for their part, are faced with a steep learning curve, as they struggle to learn the technicalities of corpus exploration techniques and to apply them to the empirical analysis of a foreign language (Frankenberg-Garcia, 2010). Furthermore, learner differences in terms of autonomy and motivation, critical thinking, language competence and corpus-analytic skills are likely to have a greater impact on project-centred pedagogy than they would in traditional lecture-style contexts (Braun, 2007; Mitchell-Schuitevoerder, 2014). Therefore the suitability of the approach I am advocating should be carefully evaluated in the light of the specificities of each learning setting. However, there are very good reasons why I believe that an attempt should be made.

As I suggested in 3 above, researchers in corpus-based translation studies have convincingly argued that corpora have

a role to play in the development of *thematic*, *technological* and *information mining* competences. However, there is also a risk, in a professionally oriented field like translator education/training, of excessive preoccupation with technical expertise. While the job market does demand that graduates possess advanced technological skills the moment they enter it, there is also a competing, if less often voiced, need for us to prepare them for a life-long career. Thus, the current challenge seems to be to strike a balance between academic and vocational priorities: “[t]echnology can easily and successfully be taught and learnt from the manufacturers’ manuals, but such a method does not encourage a critical and analytical mind”, yet it is “the enhancement of critical thinking during their studies [that] will prepare students to make well-founded decisions and choices in their [...] careers. (Mitchell-Schuitevoerder, 2014, 241; see also Mossop, 2003). Corpus-based discovery learning offers an environment where the two priorities - the vocational and the academic - can be reconciled, and where our students can develop those research competences that are required, if translation is to be regarded as “a profession whose members are competent and recognized academically” (Vandepitte, 2013, 145).

Back to the EMT competences, projects like the ones listed in Table 2 - see Hatim (2013, chapter 14) for many other project ideas - could make an important contribution to the development of (meta-) *language* competence and *intercultural* competence - through observation of the effects of the context of situation (register) and the context of culture (genre) (Halliday and Matthiessen, 2004) on sets of native and non-native/translated texts. Furthermore, the observation and evaluation of the output of multiple translators afforded by the corpus approach could raise students’ awareness of professional practices, thus also contributing to the development of the *translation service provision* competence, arguably the one competence that would seem to require the active exercise of translation skills. As claimed by Vandepitte (2013, 142),



[e]ven carrying out small research projects may be of benefit: [...] although it may remove [students] from the usual activities with texts in translation exercises, [...] [t]heir knowledge of the effect certain translations have on the readers will be expanded: students studying explicitation phenomena in translations should have improved their insight into the inferences that readers may (not) need to process more or less explicit viable target texts.

Finally, one of the limits of corpora often pointed out in the literature, namely that they “rarely provide immediate answers to a translator’s problems” (Aston, 2009, ix), can be turned into a didactic asset if we focus on the quest rather than its result. As insightfully argued by Kiraly (2000, 182),

[t]hrough our very teaching methods, we language teachers demonstrate to our students our own understanding of how language works. If we teach language as a set of artefacts, and translation skills as objectifiable, transmittable strategies, we can expect our students to develop a translator’s self-concept that sees their role as that of insignificant bilingual scribes, mechanically transcoding from one language into another.

Kiraly makes a strong connection between language competence and translation skills, suggesting that a translator’s perception of how language works influences their own self-concept as translators. The area of overlap between language and translation teaching is currently being investigated by several researchers, particularly as concerns the role of translation in the foreign language classroom (Laviosa, 2014, Pym et al., 2013). Equally important, however, is the role of language in the translation classroom. I concur with Carreres (2014, 129) that “[w]ell-designed translation-based language learning activities can help both the general language learner and the future translator enhance both

their language skills and their sensitivity to some of the challenges encountered in real-world translation” and that “those involved in translator training [should] engage more actively with research in SLA and with practices in language teaching, [and] recognise the natural connection that exists between language and translation education” (ibid., 131). Now that translation departments have firmly established themselves as having well-defined and distinct learning objectives from modern language ones, a re-appraisal of the common ground between their two subjects seems worthwhile.

Summing up, my suggestion is that corpus-based discovery projects focusing on translation studies topics could contribute to developing *all* the competences described by the EMT expert group (2009), at the same time reconciling training and education, practice and academe (González Davies 2004). Several open questions remain, particularly concerning the different conditions for use of this didactic approach – i.e., “for what types of learners, what minimum resources, what language points, how it can be integrated with other techniques, and so on” (Boulton 2009: 51) – and empirical confirmation of its merits – i.e., “whether the use of corpora, parallel or comparable, helps students develop a critical mind and whether it enhances their actual translation skills” (Mitchell-Schuitevoerder 2014: 30). These issues will have to be addressed through educational research in the *translation-with-language* classroom(s).

## References

Aston, G. "Foreword." *Corpus Use and Translating: Corpus use for Learning to Translate and Learning Corpus Use to Translate*. Ed. Beeby et al. Amsterdam: Benjamins, 2009. IX-X.

Beeby, A., Rodríguez-Inés, P., Sánchez-Gijón, P. *Corpus Use and Translating: Corpus use for Learning to Translate and Learning Corpus Use to Translate*. Amsterdam: Benjamins, 2009.

Bernardini, S. "Corpora in the Classroom." *How to Use Corpora in Language Teaching*. Ed. John McHardy Sinclair. Amsterdam: Benjamins, 2004. 15-36.

\_\_\_\_\_. "Corpus-aided Language Pedagogy for Translator Education." *Translation in Undergraduate Degree Programmes*. Ed. Malmkjær. Amsterdam: Benjamins. 97-111.

Boulton, A. "Testing the Limits of Data-driven Learning: Language Proficiency and Training." *ReCALL* 21.1 (2009): 37-54.

\_\_\_\_\_. "Data-driven Learning: The Perpetual Enigma." *Explorations across Languages and Corpora*. Ed. Stanislaw Goźdz-Roszkowski. Frankfurt: Peter Lang, 2011. 563-580.

Braun, S. "Integrating Corpus Work into Secondary Education: From Data-driven Learning to Needs-driven Corpora." *ReCALL* 19.3 (2007): 307-328.

Carreres, A. "Translation as a Means and as an End: Reassessing the Divide." *The Interpreter and Translator Trainer* 8.1 (2014): 123-135.

EMT expert group. "Competences for Professional Translators, Experts in Multilingual and Multimedia Communication." 2009. [[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)]

Frankenberg-Garcia, A. "Raising Teachers' Awareness of Corpora." *Language Teaching* 45.4 (2010): 475-489.

- Flowerdew, L. "Data-driven Learning and Language Learning Theories." *Multiple Affordances of Language Corpora for Data-driven Learning*. Ed. Leńko-Szymańska and Boulton. Amsterdam: Benjamins, 2015. 15-36.
- González Davies, M. "Undergraduate and Postgraduate Translation Degrees: Aims and Expectations." *Translation in Undergraduate Degree Programmes*. Ed. Malmkjær, Kirsten Amsterdam: Benjamins, 2004: 67-81.
- Halliday, M.; Matthiessen, C.. *An Introduction to Functional Grammar* (4th edition). London: Routledge, 2014.
- Johns, T. "From Printout to Handout: Grammar and Vocabulary Teaching in the Context of Data-driven Learning." *ELR Journal* 4 (1991): 27-45.
- Hanks, P. "The Corpus Revolution in Lexicography." *International Journal of Lexicography* 25.4 (2012): 398-436.
- Hatim, B. *Teaching and Researching Translation* (2nd ed.). London and New York: Routledge, 2013.
- Kennedy, G. "Preferred Ways of Putting Things (with Implications for Language Teaching)." *Directions in Corpus Linguistics*. Ed. Jan Svartvik. The Hague: Mouton de Gruyter, 1992. 353-373.
- Kenny, D. *Lexis and Creativity in Translation: A Corpus-based Study*. London and New York: Routledge, 2001.
- Lanstyák, I., Heltai, P. "Universals in Language Contact and Translation." *Across Languages and Cultures*, 13.1 (2012): 99-121.
- Laviosa, S. *Translation and Language Education*. London: Routledge, 2014.
- Leech, G. "Teaching and Language Corpora: A Convergence." *Teaching and Language Corpora*. Ed. Ann Wichmann, Steven Fligelstone, Anthony McEnery and Gerry Knowles. London: Longman, 1995. 1-23.

Leńko-Szymańska, A., Boulton, A. *Multiple Affordances of Language Corpora for Data-driven Learning*. Amsterdam: Benjamins, 2015.

Louw, B. "Irony in the Text or Insincerity in the Writer? The Diagnostic Potential of Semantic Prosodies." *Text and Technology: In Honour of John Sinclair*. Ed. Mona Baker, Gill Francis and Elena Tognini-Bonelli. Amsterdam: Benjamins, 1993. 157-176.

Malmkjær, K. *Translation in Undergraduate Degree Programmes*. Amsterdam: Benjamins, 2004.

McCarthy, M., O'Dell, F. *English Collocations in Use. How Words Work Together for Fluent and Natural English*. Cambridge: Cambridge University Press, 2006.

Mitchell-Schuitevoerder, R. *A Project-based Syllabus Design. Innovative Pedagogy in Translation Studies*. PhD thesis, Durham University, 2014. [Online: <http://etheses.dur.ac.uk/10830/> ]

Molés-Cases, T., Ulrike, O. "Webquests in Translator Training: Introducing Corpus-based Tasks." *Multiple Affordances of Language Corpora for Data-driven Learning*. Ed. Leńko-Szymańska and Boulton. Amsterdam: Benjamins, 2015. 199-224.

Mossop, B. "What should be Taught at Translation School?" *Innovation and E-learning in Translator Training*. Ed. Anthony Pym, Carmina Fallada, José Ramón Biau and Jill Orenstein. Tarragona: Intercultural Studies Group, Universitat Rovira i Virgili, 2003. 20-22.

Nunan, D. *The Learner-centred Curriculum*. Cambridge: Cambridge University Press, 1988.

Palmer, H. E. *Second Interim Report on English Collocations*. Tokyo: Kaitakusha, 1933.

Pym, A., Malmkjær, K., Gutierrez-Colon Plana, M. M. *Translation and Language Learning: The Role of Translation in the Teaching of Languages in the European Union*. Luxembourg: Publications Office of the European Union, 2013.

Rodríguez-Inés, P. "Evaluating the Process and not just the Product when using Corpora in Translator Education." *Corpus Use and Translating: Corpus use for Learning to Translate and Learning Corpus Use to Translate*. Ed. Beeby et al. Amsterdam: Benjamins, 2009. 129-149.

Sánchez-Gijón, P. "Developing Documentation Skills to Build Do-it-yourself Corpora in the Specialised Translation Course." *Corpus Use and Translating: Corpus use for Learning to Translate and Learning Corpus Use to Translate*. Ed. Beeby et al. Amsterdam: Benjamins, 2009. 109-127.

Schmidt, R. W. "The Role of Consciousness in Second Language Learning." *Applied Linguistics* 11 (1990): 129-158.

Seidlhofer, B. *Controversies in Applied Linguistics*. Oxford: Oxford University Press, 2003.

Sinclair, J. M. *Looking up: An Account of the COBUILD Project in Lexical computing and the Development of the Collins COBUILD English Language Dictionary*. London: Harper Collins, 1987.

\_\_\_\_\_. "The Search for Units of Meaning." *Textus* 9.1 (1996): 75-106.

Vandepitte, S. "Research Competences in Translation Studies." *Babel* 59.2 (2013): 125-148.

Widdowson, H. "The Description and Prescription of Language." *Georgetown University Roundtable on Language and Linguistics*. Ed. James E. Alatis. Washington D.C.: Georgetown University Press. 11-24.

Recebido em 04 janeiro de 2016  
Aceito em 22 de fevereiro de 2016  
Publicado em abril de 2016