The *Toledo Teacher Trainees* corpus (TTT): Bridging the gap between students' narratives and corpus linguistics

Fátima Faya-Cerqueiro[a] – Gema Alcaraz-Mármol[b]
University of Santiago de Compostela[a] / Spain
University of Castilla La Mancha[b] / Spain

Abstract – In recent decades a few research methods have resorted to L2 learners in order to analyse several aspects aiming at methodological improvements. One of them is corpus linguistics, which has largely contributed to the study of language production from a quantitative perspective. A very different one has been the compilation of perceptions of the L2 learning process using 'narrative inquiry' and qualitative methods of analysis. However, scholars have not addressed the combination of both methods. In this proposal we examine their main individual features and offer an interwoven line of research, applying the quantitative approach of corpus linguistics to the genre of language learning narratives. Thus, we present a new corpus of L2 learners' perceptions and provide detailed information on its structure, compilation and categorisation. The interdisciplinary status of this proposal will enable the exploration of new research possibilities that can ultimately benefit the teaching-learning process.

Keywords – narrative inquiry; L2 learning narratives; corpus linguistics; teacher trainees; corpus-based perceptions

1. INTRODUCTION[1]

The twenty-first century has witnessed a growing interest in how second language (L2) students conceive their own learning process. The interest and effort to improve this process on the part of the teaching community and linguists is not new. Nonetheless, this interest has traditionally left aside the reflections and perceptions of the learners themselves, focusing on more objective aspects such as the product of the learning process. It is precisely in the light of these facts that learner corpora were born at the beginning of the 1990s, centred on the analysis of the written and oral productions of L2

students. However, learner corpora generally focus on the learners' L2 output, not on the learning process itself. In this sense, Pavlenko (2007: 163) affirmed:

> In the past decade, language memoirs, linguistic autobiographies, and learners' journals and diaries have become a popular means of data collection in applied linguistics. It is not always clear however how one should go about analysing these data.

These memoirs, autobiographies and journals allow a new approximation to the learning process from the learners' standpoint itself. These elements offer valuable data for researchers, teachers and teacher trainers, which lead to the potential improvement of teaching methodologies.

However, despite offering some directions for systematic analysis, the focus of these studies is eminently qualitative. Pavlenko's statement reveals there is still room to pose different research questions and adopt new perspectives when tackling this genre. In this sense, corpus linguistics may constitute a methodological framework where data, such as the content of these narratives, can be approached from a perspective more in line with a quantitative standpoint. What is suggested here is the use of corpus within a broader scope and to go beyond L2 output analysis, paying attention to aspects which have more to do with the exercise of reflection and self-report about the learning process itself.

In what follows, we present the *Toledo Teacher Trainees* corpus (TTT), based on L2 learners' narratives. We aim at building a bridge between two approaches, namely corpus linguistics and learners' perceptions, in an attempt to address the data taken from the latter through the method promoted by the former. We will first describe the theoretical framework behind our proposal (cf. Section 2). This framework explores the contribution of L2 learning narratives to the perception of the teaching learning process, as well as the potential role of corpus linguistics in Second Language Acquisition in general and in the analysis of these narratives in particular. We pay particular attention to learner corpus research, a relatively recent area of study that shares an interest in learners' language learning process. In fact, a comparison between learner corpora, narratives and our own methodology is established. Section 3 provides a detailed description of our corpus, including compilation methods, participants' profile, main steps taken for analysis and categorisation. Our paper finishes with perspectives and research directions regarding potential benefits and advantages that our corpus offers for future studies (cf. Section 4).

2. PREVIOUS RESEARCH

Over the last 35 years, corpus linguistics research has dealt with the description of different types of registers, among which we can find informal conversation, journals, reports, female and male language or even dialects. As regards language learning and teaching, more and more researchers highlight how corpus linguistics can help this discipline (Aijmer 2002; Reppen 2010; Frankenberg-Garcia *et al.* 2011). In Römer's (2011: 205) words, "in the field of applied linguistics, more and more researchers and practitioners treasure what corpus linguistics has to offer to language pedagogy." This particular field within corpus linguistics, dealing with the collection of linguistic evidence from non-native speakers, is known as learner corpus studies.

The origins of learner corpora are found in the late 1980s, but they did not really take root as a discipline until the beginning of the 1990s. A learner corpus brings together corpus linguistics and language learning, aiming at providing descriptions of learner language, and offering new perspectives on second language learning, and sometimes changing the ones already established. It helps us gain better understanding of how languages other than the first language (L1) are learned (Ortega 2009). Paquot and Plonsky (2017: 2) define it as an "interdisciplinary enterprise which sits its crossroads between a variety of disciplines."

The focus of a learner corpus is to be found in two types of studies. The first type pursues the analysis of L2 output in the written or spoken form, exploring several linguistic patterns of different nature within the learners' interlanguage. The linguistic patterns of interest for research include syntactic complexity (Vyatkina 2013), frequency and correctness of vocabulary (Laufer and Waldman 2011) and even pragmatics (Chen 2010; Polat 2011). The second type delves around what is called 'Error Analysis', which helps us understand the development of the learning process (Granger 2002). Among this second category, we find Divsar and Heydari (2017) and Botley *et al.* (2007). The former analyse several categories of errors in Iranian English as a Foreign Language learners' essays, whereas the latter pay attention to spelling.

The range of learner corpora that exist nowadays varies in number and category. The *Cambridge Learner Corpus* (Nicholls 2003) and the *International Corpus of Learner English* (Granger *et al.* 2009) are among the most ambitious, with data banks of learners from different L1 backgrounds and millions of words. These two corpora use the written language as a medium. However, others such as the *College Learner Spoken English*

Corpus (Yang and Wei 2005), the *Corpus of the National Institute of Information and Communications Technology for Japanese Learners of English* (Izumi *et al.* 2004), or the *Corpus of Young Learner Interlanguage* (Housen 2002) rely on spoken language. They are usually smaller than the written ones but equally relevant.

Learner corpus linguistics has adopted this standpoint based on the analysis of L2 learners' output, paying attention to the learners' interlanguage and errors in order to contribute to the improvement of language teaching methodology. We share this same aim but, in our case, the object of analysis is not the L2 production, but the learners' perceptions on their own L2 learning process.

These reflections are the object of research of what it is known as 'narrative inquiry.' Narrative inquiry, broadly defined by Barkhuizen *et al.* (2014: 3) as "an established umbrella term for research involving stories," became a prominent research method in social science at the turn of the twenty-first century (Clandinin and Connelly 2000). The recognition of narratives as valid data and the need of qualitative approaches to cover personal experiences were major assets for the development of this methodology (Huber *et al.* 2013: 217). Some authors have claimed the relevance of narratives together with a need for systematicity in their study (Atkinson and Delamont 2006), and several proposals regarding design, collection and analysis have actually provided a consistent foundation for research (Pavlenko 2007; Riessman 2008; Wells 2011). Nowadays, this "interdisciplinary method that views lives holistically" (Marshall and Rossman 2014: 157) is a well-established and productive approach encompassing work from different areas, as attested by the variety of disciplines included in first volumes devoted to narrative inquiry (Clandinin 2007) or by the journal that bears the same name.

Within social science, the field of education, probably due to its broad scope, has adopted this methodology for an array of studies, where narrative inquiry can provide a better understanding of teaching and teacher education, classroom practice, classroom management or curriculum design, among others (Johnson and Golombek 2002; Craig 2011; Kitchen *et al.* 2011; Huber *et al.* 2013). As regards L2 learning the most common areas of research using learners' narratives include motivational issues (cf. Thompson and Vásquez 2015; Farahani *et al.* 2019), identity (Early and Norton 2012; Benson *et al.* 2013) and also more specific perceptions on given topics, like the native speaker myth, related to the lack of confidence perceived by non-native language teachers (Reis 2011), test-takers' perspectives (Rajendram *et al.* 2019; Sinclair *et al.* 2019) or assessment

(Franco 2020). Although most narratives deal with general issues regarding L2 previous experiences, research can also include more specific or tailored narratives, such as study-abroad experiences (Benson *et al.* 2013) or learners' perceptions of teachers (Oxford 2001). The analysis of teachers' or learners' narratives, in particular, has also demonstrated its relevance for practitioners, future students and learners to reflect upon their experiences. Similarly, research on educational stakeholders' perspectives can be used to implement or improve educational practices, being thus a cyclic process when teaching and learning perceptions can eventually change teaching and learning attitudes.

In the L2 context, narrative inquiries have mostly been approached through face-to-face interviewing processes, which makes data gathering a costly and time-consuming method of compilation requiring interaction between the learner and the researcher. Moreover, the number of participants in L2 narrative inquiry research is generally limited due to this procedure, since when using interviews the time devoted to compilation and transcription can equal a low number of participants in those studies. Quantitative approaches to L2 self-perceptions are virtually non-existent. Yet, a study by Baker and MacIntyre (2000), which examines a French immersion programme in Canada, could be in the line of a quantitative perspective. As a way to complement their survey, based largely on quantitative research methods and supported by statistical tests, they asked students to describe a positive or negative experience about speaking French. Although they refer to this part of the study as qualitative, it does include percentages of different lexical items regarding the experiences in the immersion and non-immersion groups.

Table 1, below, shows the differences and similarities found in traditional learner corpora, corpus of learners' perceptions and narrative inquiry, as applied to L2 learning experiences. The comparison is established upon eight features. The three fields of study share some features such as the participants' profile, the format and the possible application of results. By contrast, important differences are observed, namely the scope, the number of participants and the types of results obtained. Traditional learner corpora are mainly concerned with L2 production, while the other two analyse self-reflection on the language learning process, which can be expressed in the learners' L1 or in L2, depending mainly on language proficiency.

	Traditional learner corpus	Corpus of learners' perceptions	Narrative inquiry
Participants' profile	L2 learners	L2 learners	L2 learners
Number of participants	Representative	Representative	Case studies
Format	Oral/written	Written	Oral/written/multimodal
What do they study?	Linguistic data	Metalinguistic data	Metalinguistic data
Language	L2	L1	L1/L2
Aim	Interlanguage / errors	Learning experience	Learning experience
Application	Methodology	Methodology	Methodology
Type of results	Quantitative	Quantitative	Qualitative

Table 1: Differences and similarities between traditional learner corpus, corpus of learners' perceptions and narrative inquiry

The review of the literature reveals a gap, as the combination of both elements – corpus linguistics and narrative inquiry – has not yet been tackled. Thus, our proposal differs from the abovementioned works in two main aspects. First, research on learners' perceptions has been developed with a small number of participants, and generally focused on a particular element within the learning process. Second, those studies adopt a qualitative perspective, maybe partly due to the small number of participants. We consider, therefore, that corpus linguistics can open new lines of research as regards narrative inquiry and students' perceptions, adopting a more quantitative perspective and with a significantly higher number of participants.

3. *TOLEDO TEACHER TRAINEES* (TTT): A CORPUS OF L2 PERCEPTIONS

3.1. Compilation, structure and main principles

The *Toledo Teacher Trainees* corpus (TTT) aims to be at the crossroads between learner corpora and L2 narratives, as it provides a collection of learners' perceptions which has been designed to make it a representative source of information for different research purposes. The participants were 354 future teachers of Primary and Infant Education from the Faculty of Education in Toledo, Spain. The four academic years in these degrees are represented in our sample with 162 students in their first year, 123 in their second, 42 in their third and 27 in their fourth. After completion of their second year, students achieve a B1 level, while the last two years belong to a special language training group that grants a B2 level. Most of the participants are taking a Degree in Primary Education (269) and the rest study a Degree in Infant Education (85). Their ages range from 18 to 45 with a

mean of 20.0 years. Regarding gender they fall into the following categories: 245 female, 109 male. All of them declare Spanish as their native language except for four, three of whom spoke Romanian and one Valencian as L1. Only 14% of them have not studied a third language (50). Among the 354 students, 147 have taken part in a bilingual CLIL programme (English/Spanish) at some stage before university.

The compilation of the corpus was done through an *ad-hoc* online questionnaire. Learners had to fill in the questionnaire individually. Different sessions were organised for each group to enable simultaneous participation without time limit and they were supervised by their language teacher and/or researcher to answer possible questions and to provide information on the purposes of the compilation. The questionnaire was not shown to the participants previously although they were briefly informed about its main features. All the contributions were compiled during the same academic year.

The questionnaire consisted of two main parts. The first part contained questions regarding the profile of the participants such as age, gender, degree, L1, additional languages and educational background. As for the second part it included specific items concerning their experience in the different pre-university stages. Participants had to answer open questions with no word limit and a closed question where they were required to write three keywords related to their experience, which could summarise their perceptions of each stage. Open questions comprehended different aspects of their academic life such as teachers, materials and resources, methodology, contents and assessment process.

The corpus comprises around 170,000 words and includes participants' answers to the open questions regarding each of the mentioned educational stages as well as the keywords associated to the different school periods.

All the questions were written in Spanish and this language was also required for their responses. The reason why their L1 was adopted was to guarantee they could express themselves without linguistic limitations. This is common practice in studies that focus on metacognitive or metalinguistic features and also in L2 narratives (Pavlenko 2007: 172):

> in studies of subject and life reality where the speakers' L2 proficiency is low and the L1 is shared with the researcher, the choice of L1 as the language of data collection is justified.

Accordingly, Farahani *et al.*'s (2019) research B1 level participants use their L1, Persian, when writing language learning histories.

3.2. Our approach

The present study is part of a wider research project that intends to explore perceptions about the L2 teaching-learning process of university students in earlier stages. The corpus analysis is developed in two stages. First, we focus on the keyword list of each stage, a phase which includes systematisation and analysis of these words, which are classified into broad semantic categories. This first step enables the comparison of different patterns identified in the keywords. A second stage will resort to this classification for the analysis of the answers to the open questions.

AntConc (Anthony 2019) was used to alphabetically organise and count all the keyword instances, which follow the clearance of non-valid answers, such as whole clauses, prepositions and other non-content words since our main aim was to see students' major topics of interests. The selection focused on nouns, verbs and adjectives. For practical reasons all the tokens were unified into types, so these types included masculine and feminine adjectives (*cómodo/a*, 'comfortable') and word categories with the same lemma (*diferencia/diferente* 'difference/different'; *aprendizaje* 'learning'/verb forms of *aprender* 'to learn').

Each type was classified according to broad semantic fields in different stages. For this purpose there was a progression from narrower sets or number of categories to broader ones. Three evaluators were simultaneously immersed in this process, which consisted of assigning a label to each token or more than one in an initial stage. In the first explorations categories amounted to near 30 and were finally reduced to 11.

Umbrella terms were applied to categories. The idea was to present them as neutral concepts as possible. In this way, they would become comprehensible enough to admit all the different possibilities within a continuum. Some of the final categories are 'emotion/feelings', 'complexity', 'innovation' or 'usefulness.' They show a great variability regarding number of members, token/type ratio, range of word categories and also the type of concepts that they define (abstract vs. concrete ideas).

However, semantic categories are not always absolute terms and blurred lines between categories do occur. Simon-Vandenbergen and Aijmer (2008: 12) know the

difficulty when defining "the semantic field a priori and there is no obvious and well-defined number of dimensions structuring the field." This idea contrasts with the mosaic conception of lexical fields, where components fit nicely. In fact, Geeraerts *et al.* (1994) warn that there might be overlapping between different fields given their fuzzy boundaries. Accordingly, non-standard characteristics are established, which are far from the traditional established ones. In those cases, agreement on the category which better fits the word is supported by dictionaries and thesaurus.

Far from theorising on semantics itself, semantic concepts are employed to create useful categories for the purposes of our research, always within coherence. The literature offers a plethora of definitions that remark the relationship of words or lemmas that share semantic features. According to Mackey (1965: 76), a semantic field is:

> made up of basic key-words, which command an army of others. The semantic area may be regarded as a network of hundreds of associations, each word of which is capable of being the centre of a web of associations radiating in all directions.

Words with common semantic associations can be categorised as members of such a network. As noted by Kittay and Lehrer (1992: 3), "words applicable to a common conceptual domain are organised within a semantic field by relations of affinity and contrast (e.g., synonymy, hyponymy, incompatibility, antonymy, etc.)." Following this idea, the notion of semantic field is understood in a broad sense here, comprising different hierarchical associations. Core members of each semantic field or category are easily identified, while more peripheral members require the support of dictionary definitions. In some cases, definitions of a given entry are insufficient and further searches of words within the definition are needed.

The category *cambio* 'change', which is among the shortest categories, can serve as an example. It would include not only 'change' alone, but also 'lack of change' and all the different options within the cline 'change-lack of change.' Thus, the following members have been identified: *ágil* ('agile'), *cambio* ('change'), *diferencia* ('difference'), *dinámico* ('dynamic'), diverso ('diverse'), *evolución* ('evolution'), *monótono* ('monotonous') and *repetitivo* ('repetitive').

4. PERSPECTIVES AND RESEARCH DIRECTIONS

As already noted, narrative inquiry can bring about immediate benefits for education students. Reflecting on previous experiences as learners may also lead future practitioners to become aware of the kind of teacher they want to be, the methodological approaches to use and other aspects concerning their teaching practice. At the same time, this narration of their L2 learning experiences can provide researchers with valuable data to detect weaknesses and strengths regarding L2 teaching and aspects that might influence L2 learning in various ways. These data can turn into suggestions for action at different levels.

In our opinion, corpus linguistics can enrich the potential of these narratives, providing a new approach to their analysis. The compilation of a structured corpus with different variables and the introduction of quantitative methods to the genre of L2 narratives will enable sharper research questions and more solid generalisations on how the teaching-learning process of different school periods is perceived. The approach to categorisation presented above is just an example of the possibilities offered by the corpus. Some of these possibilities for further research have already been suggested. Similarly, the specific and precise categorisation of lexical items will facilitate corpus searches in the narratives, not only between elements of the corpus itself, but beyond. Although in the present article only one category has been sketched, there are more semantic categories which have been established on the basis of the keyword analysis. In turn, these categories can be completed with new members identified in the corpus beyond the keywords. These semantic categories can also serve as a starting point for carrying out searches for specific terms of interest related to the teaching and learning process.

The wide range of criteria taken into account for data collection will allow us to make systematic comparisons including variables such as gender, linguistic background, educational stage and even aspects on which dense information has been obtained such as teachers, assessment or resources, which stand as object of reflection for our participants. Moreover, comparisons can be made on the basis of the different academic years to which the participants belong. Adopting an apparent-time perspective and having students from all years in the same degree, we have the opportunity to observe the evolution of perceptions from the first to the last year of their university stage.

In this first contribution to corpus-based perceptions we have introduced our compilation approach. Nevertheless, our corpus offers a wide range of possibilities yet to be explored, which can be of interest to linguists, teachers and teacher trainers. However, a whole corpus remains to be exploited beyond the keywords, which are the easiest part to identify within the corpus, but the most constrained as well.

The specific nature of our corpus encourages us to contrast our results with other specific or more general corpora. Some of these comparisons should be taken with caution, since our corpus is framed within a highly specific context, where most students speak one and the same L1. Besides, most of them belong to the same geographical area in the province of Toledo. Therefore, some evidence from our corpus cannot be extrapolated. However, in spite of the possible limitations derived from context, the corpus structure, its longitudinal and transversal nature will allow to establish comparisons with similar corpora in different contexts of a multilingual nature.

Not only do the possibilities for comparison remain intracorpus, but they can also be found intercorpus. Indeed, one could study aspects of interest to the field of learner corpus but from a perception perspective. That is, as mentioned above, errors are one of the most explored aspects in learner corpora, and these are analysed from a strict perspective of linguistic correction. In our corpus, errors may constitute an element for reflection among the participants, being thus treated from a metalinguistic point of view. Similarly, we could deal with aspects related to oral and written production, within a framework of reflection on the learning process itself.

On the other hand, comparisons could also be carried out with aspects tackled in narrative inquiry research, such as aspects of identity, self-concept and attitude towards the L2. The corpus perspective would provide a quantitative basis in this sense. Furthermore, the potential of this corpus goes beyond the quantitative, since it offers material that can be qualitatively analysed through subsequent interviews with a small group of participants, taking objective data as a starting point.

Therefore, and in spite of the specificity that characterises the corpus presented here, we consider that the study points to the interdisciplinary value of second language research, particularly in the field of perceptions, where the systematicity offered by corpus methodology paves new paths of research.

5. REFERENCES

Anthony, Laurence. 2019. *AntConc*. Tokyo: Waseda University.

Aijmer, Karin. 2002. Modality in advanced Swedish learners' written interlanguage. In Sylviane Granger, Joseph Hung and Stephanie Petch-Tyson eds. *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam: John Benjamins, 55–76.

Atkinson, Paul and Sara Delamont. 2006. Rescuing narrative from qualitative research. *Narrative Inquiry* 16/1: 164–172.

Baker, Susan C. and Peter D. MacIntyre. 2000. The role of gender and immersion in communication and second language orientations. *Language Learning* 50/2: 311–341.

Barkhuizen, Gary, Phil Benson and Alice Chik. 2014. *Narrative Inquiry in Language Teaching and Learning Research*. New York: Routledge.

Benson, Phil, Gary Barkhuizen, Peter Bodycott and Jill Brown. 2013. *Second Language Identity in Narratives of Study Abroad*. Basingstoke: Springer.

Botley, Simon, Faizal Hakim and Doreen Dillah. 2007. Investigating spelling errors in a Malaysian learner corpus. *Malaysian Journal of ELT Research* 3/1: 74–93.

Chen, Hsin-I. 2010. Contrastive learner corpus analysis of epistemic modality and interlanguage pragmatic competence in L2 writing. *Arizona Working Papers in SLA and Teaching* 17/1: 27–51.

Clandinin, D. Jean ed. 2007. *Handbook of Narrative Inquiry: Mapping a Methodology*. Los Angeles: Sage Publications.

Clandinin, D. Jean and F. Michael Connelly. 2000. *Narrative Inquiry: Experience and Story in Qualitative Research*. San Francisco: Jossey-Bass.

Craig, Cheryl. 2011. Narrative inquiry in teaching and teacher education. *Narrative Inquiries into Curriculum Making in Teacher Education* 13: 19–42.

Divsar, Hoda and Robab Heydari. 2017. A corpus-based study of EFL learners' errors in IELTS essay writing. *International Journal of Applied Linguistics and English Literature* 6/3: 143–149.

Early, Margaret and Bonny Norton. 2012. Language learner stories and imagined identities. *Narrative Inquiry* 22/1: 194–201.

Farahani, Ali Akbar Khomeijani, Abbas Ali Rezaee and Robabeh Moshtaghi Zonouz. 2019. Motivational trajectories in language learning: Evidence from highly-motivated English as a foreign language learners. *Electronic Journal of Foreign Language Teaching* 16/2: 281–299.

Franco, Ashleigh. 2020. Not all Finns think alike: Varying views of assessment in Finland. *International Education Studies*. 13/1: 1–10.

Frankenberg-Garcia, Ana, Lynne Flowerdew and Guy Aston. 2011. *New Trends in Corpora and Language Learning*. London: Continuum.

Geeraerts, Dirk, Stefan Grondelaers and Peter Bakema. 1994. *The Structure of Lexical Variation: Meaning, Naming and Context*. Berlin: Mouton de Gruyter.

Granger, Sylviane. 2002. A bird's-eye view of learner corpus research. In Sylviane Granger, Joseph Hung and Stephanie Petch-Tyson eds. *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam: John Benjamins, 3–33.

Granger, Sylviane, Estelle Dagneaux, Fanny Meunier and Magali Paquot. 2009. *International Corpus of Learner English*. Louvain-la-Neuve: Presses Universitaires de Louvain.

Housen, Alex. 2002. A corpus-based study of the L2 acquisition of the English verb system. In Sylviane Granger, Joseph Hung and Stephanie Petch-Tyson eds. *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam: John Benjamins, 77–116.

Huber, Janice, Vera Caine, Marilyn Huber and Pam Steeves. 2013. Narrative inquiry as pedagogy in education: The extraordinary potential of living, telling, retelling and reliving stories of experience. *Review of Research in Education* 37/1: 212–242.

Izumi, Emi, Kiyotaka Uchimoto and Hitoshi Isahara. 2004. *Nihonjin 1,200-nin no Eigo Speaking Corpus (A speaking corpus of 1,200 Japanese learners of English)*. Tokyo: ALC Press.

Johnson, Karen E. and Paula R. Golombek. 2002. *Teachers' Narrative Inquiry as Professional Development*. Cambridge: Cambridge University Press.

Kitchen, Julian, Darlene Ciuffetelli Parker and Debbie Pushor eds. 2011. *Narrative Inquiries into Curriculum Making in Teacher Education*. Bingley: Emerald.

Kittay, Eva Feder and Adrienne Lehrer. 2012. Introduction. In Adrienne Lehrer, Eva Feder Kittay and Richard Lehrer eds. *Frames, Fields and Contrasts: New Essays in Semantic and Lexical Organization*. New York and London: Routledge.

Laufer, Batia and Tina Waldman. 2011. Verb-noun collocations in Second Language writing: A corpus analysis of learners' English. *Language Learning* 6/2: 647–672.

Mackey, William F. 1965. *Language Teaching Analysis*. London: Longman and Indiana University Press.

Marshall, Catherine and Gretchen B. Rossman. 2014. *Designing Qualitative Research*. Los Angeles: Sage Publications.

Nicholls, Diane. 2003. The Cambridge Learner Corpus– error coding and analysis for lexicography and ELT. In Dawn Archer, Paul Rayson, Andrew Wilson and Tony McEnery eds. *Proceedings of the Corpus Linguistics 2003 Conference*. Lancaster University: University Centre for Computer Corpus Research on Language, 572–581.

Ortega, Lourdes. 2009. *Understanding Second Language Acquisition*. London: Hodder Education.

Oxford, Rebecca L. 2001. The bleached bones of a story: Learners' constructions of language teachers. In Michael P. Breen ed. *Learner Contributions to Language Learning: New Direction in Research*. London: Longman, 86–111.

Paquot, Magali and Luke Plonsky. 2017. Quantitative research methods and study quality in learner corpus research. *International Journal of Learner Corpus Research* 3: 61–94.

Pavlenko, Aneta. 2007. Autobiographic narratives as data in applied linguistics. *Applied Linguistics* 28/2: 163–188.

Polat, Brittany. 2011. Investigating acquisition of discourse markers through a developmental learner corpus. *Journal of Pragmatics* 43/15: 3745–3756.

Rajendram, Shakina, Jeanne Sinclair and Elizabeth Larson. 2019. International graduate students' perspectives on high-stakes English tests and the language demands of Higher Education. *Language and Literacy* 21/4: 68–92.

Reis, Davi Schirmer. 2011. "I'm not alone:" Empowering non-native English-speaking teachers to challenge the native speaker myth. In Karen E. Johnson and Paula R. Golombek eds. *Research on Second Language Teacher Education*. New York: Routledge, 45–63.

Reppen, Randi. 2010. *Using Corpora in the Language Classroom*. Cambridge: Cambridge University Press.

Riessman, Catherine Kohler. 2008. *Narrative Methods for the Human Sciences*. Los Angeles: Sage Publications.

Römer, Ute. 2011. Corpus research applications in Second Language teaching. *Annual Review of Applied Linguistics* 31/1: 205–225.

Simon-Vandenbergen, Anne-Marie and Karin Aijmer. 2008. *The Semantic Field of Modal Certainty: A Corpus-Based Study of English Adverbs*. Berlin: Walter de Gruyter.

Sinclair, Jeanne, Elizabeth Jean Larson and Shakina Rajendram. 2019. "Be a machine:" International graduate students' narratives around high-stakes English tests. *Language Assessment Quarterly* 16/2: 236–252.

Thompson, Amy S. and Camilla Vásquez. 2015. Exploring motivational profiles through language learning narratives. *The Modern Language Journal* 99/1: 158–174.

Vyatkina, Nina. 2013. Specific syntactic complexity: Developmental profiling of individuals based on an annotated learner corpus. *The Modern Language Journal* 97/1: 11–30.

Wells, Kathleen. 2011. *Narrative Inquiry*. Oxford: Oxford University Press.

Yang, Huizhong and Naixing Wei. 2005. *The Construction of and Research on the COLSEC*. Shanghai: Shanghai Foreign Language Education Press.

Corresponding author
Fatima Faya-Cerqueiro
University of Santiago de Compostela
Department of Applied Didactics
E-15782 Santiago de Compostela
Spain
e-mail: fatima.faya@usc.es