

Learner corpora in the classroom: a useful and sustainable didactic resource

Fanny Meunier – CECL, UCL, Belgium

fanny.meunier@uclouvain.be

Introduction

Learner corpora have been collected and analyzed for approximately 20 years now. They can be defined as electronic collections of (near-)natural spoken or written texts produced by foreign or second language learners in a variety of language settings and assembled according to explicit design criteria.¹ They have mainly been used to better describe (near-)authentic learner language use through cross-sectional and longitudinal approaches, foster the understanding of second language acquisition processes (Myles 2008, Granger 2009, Granger and Meunier 2010) and enhance pedagogical materials such as learners' dictionaries, vocabulary and grammar books (Grabrielatos 2005, Meunier and Gouverneur 2009, Römer forthcoming 2011). They have also been used as a data source for the creation of data-driven learning activities (Boulton 2008, 2009a and 2009b, Gilquin and Granger 2010), whether carried out in class, made available on the web, or on language-learning platforms.

The present paper focuses on one of the many uses that learner corpora have been put to, i.e. their use (or under-use) as a pedagogical resource in the foreign language classroom. Whilst many language teachers have often 'heard about'² (learner) corpora, their actual market penetration rate (to use a business-oriented metaphor) in classrooms is almost negligible (as clearly stated by Römer 2006, Breyer 2009, Granger 2009 and Meunier 2010). In the present paper I first comment on the reasons behind this lack of uptake of learner corpora in classrooms. In section two, I explain why, despite the obstacles listed in section 1, learner corpora are useful, multi-purpose and sustainable didactic resources for teachers. Concrete suggestions for a fuller integration of such resources in the classrooms are presented in the last section of the paper.

1. Reasons for the lack of uptake of learner corpora in classrooms

In a recent article (Meunier 2011) I list several reasons for the lack of uptake of corpus-oriented tools and methods in the classroom, some of which are summarized below:

- many teachers are not aware of the possibilities offered by (learner) corpora and of the changes that corpus methods have brought to materials that they are using. It is still very often the case that no (or hardly any) time is devoted to corpus issues in pre- or in-service teacher training. Another reason for this lack of awareness can be found in the sometimes vague statements found in the introductions to teaching materials where the corpus-informed nature of the materials is mentioned cursorily but not explained in detail;
- some teachers believe that linguists who advocate the use of corpus methods have no idea of what teaching is about, despite the fact that some of those linguists are also teachers. This feeling of distance is often reinforced by the fact that the types of examples provided in the literature are mainly related to EAP/ESP courses (e.g. Feak and Swales 2010, Jones and Schmitt 2010, Römer and Wulff 2010), which does not always facilitate a transposition to learners with less advanced proficiency levels; the term 'corpus' is – rightly so – associated with computers, which might be problematic for some schools where either equipment or expertise is not easily available;

¹ The present definition is a combination of two definitions of learner corpora given by Granger in 2002 and 2009.

² To use an oft-quoted comment by teachers in the framework of in-service teacher training sessions (personal experience)

- the term ‘corpus’ is – often wrongly so – associated exclusively with quantification, figures and statistics, which might constitute a strong disincentive for some potential users who do not perceive the added value of such an approach in their teaching practice;
- there is a lack of empirical studies exploring the actual impact of corpus methods on the learning outcomes. In addition, the results of the available studies (Vannestal and Lindquist 2007, Belz and Vyatkina 2008, Boulton 2009, and Breyer 2009) present a contrasted picture and show that using corpora with students: i) may require substantial support in some cases, ii) may be time-consuming, iii) does not appeal to all the students, and iv) may prove beneficial for some skills and tasks but not for others.

Whilst the five reasons presented here above are valid for any type of corpus (native or learner), two additional reasons for the specific lack of use of learner corpora both in classrooms and in L2 syllabuses and materials³ are:

- the lack of availability of learner corpora;
- the fact that “the topics covered in most existing learner corpora are often unsuitable for the everyday needs of the vast majority of L2 school teachers, who target the L2 for general purposes, often for a teenage audience” (Meunier, 2010: 211).

In sum, all these obstacles have naturally led to a feeling of lack of ‘pedagogical relevance’ (Braun 2005) and ‘authentication’ (Belz and Viyatkina 2008). Despite such limitations, concrete projects have nonetheless been undertaken to promote a use of learner corpora which caters for learners’ needs and creates a sense of pedagogical relevance and authentication. Some of these projects will be presented in section 2, together with a plea for widening the understanding of the learner corpus concept.

2. Learner corpora as useful and sustainable didactic resources

Despite the obstacles presented in section 1, learner corpora do constitute an innovative, multi-purpose and sustainable teaching resource. As is the case for the introduction of any new sustainable resource in our lives, the use of learner corpora requires some initial investment in time and a change of habits, but, to use another business metaphor, the return on investment may prove substantial.

2.1. Fostering pedagogical relevance and authentication

Learner corpora will find their way to classrooms if they are used as tools promoting the creation of a sense of community. As clearly explained by Belz and Viyatkina (2008) corpora will only be used by language learners if they can interpret, analyze and understand them in a personally meaningful way. This direct involvement of learners in corpus collection and use corresponds to what Granger (2009: 25) calls “corpora for immediate pedagogical use (IPU)”, i.e. data “collected by teachers as part of their normal classroom activities [...] [and where] the learners are at the same time producers and users of the corpus data”. This IPU can be found in telecollaborative projects in which L2 learners interact, using the oral or written mode, with native speakers or other non-native speakers of the target language. The oral and written interactions, once archived, can be used in the framework of pedagogical interventions. One rather early example is Kasper and Rose (2002) who used the students’ interactions as a resource for the creation of exercises focusing on specific linguistic forms produced by the learners during their meaningful interactive communicative tasks. More recently, Belz and Viyatkina (2008: 35) describe their own study as part of a

language learning configuration in which distally located learners use Internet communication tools for social interaction, dialogue, and debate with NS age peers [...] English-speaking learners of German at a large public university in the United States engaged in genuine interactions with German-speaking

³ This does not apply to language for academic/specific purposes - domains where teachers seem more inclined to use native and advanced learner corpora (see, for instance, Flowerdew 2003, Gilquin et al. 2007, or Paquot 2008).

learners of English at a German teachers' college for eight weeks [...] All correspondence was archived in a web-based teleconferencing program as they were produced.

The corpus of language produced by the learners and the native speakers was then used to create data-driven learning material for the teaching of German particles to the very learners who had been involved in the tele-collaborative project, hence creating a sense of clear authentication. The pedagogical relevance was ensured by what Braun (2006) calls the 'homogeneity' and 'topical relevance' of such types of corpora, which, she adds, are more important than the issues of corpus size or representativity.

In the two examples above, the focus was on grammatical aspects, but learner corpora can also be used to analyze the pragmatic features of interactions. Rizzardi, Pedrazzini and Nava (2008), through their analysis of a Role Play learner corpus revealed the learners' difficulties in using appropriate relational/pragmatic features of the language. The corpus collected included 57 spoken interactions (role-plays) between Italian learners of English. A few years ago, I carried out another study on speech acts using a learner corpus of synchronous written interactions. One of my French-speaking groups of second-year students majoring in English at the Université Catholique de Louvain (UCL, Belgium) collaborated with another group of Irish learners of French at Trinity College Dublin (TCD). This joint project took place within the framework of a writing course. A task-based methodology was adopted and learners had to discuss various topics. One of the topics proposed as a prompt for discussion was an information technology (IT) topic which read as follows: What have been the three most important developments in information technology and computers over the past 40 years, and why? The pedagogical relevance of the task included, among others:

- synchronous communication with native speakers in the target language
- chat-like online discussion
- reinforcement of tandem/peer-to-peer learning⁴
- management of messages and assignments
- reinforcement of learner autonomy
- focus on a written communication medium that had to be very precise (no body language, no sound, compulsory expression of non-verbal elements if necessary in the discussion, etc.)
- use of new technologies

All these aspects are in line with what Ruiz-Madrid (2007: 57) advocates, i.e. the use of web resources "not to teach the same thing in a different way but rather to help our students to enter a new realm of collaborative inquiry and construction of knowledge".

In the UCL-TCD project, the users' log session transcripts were transferred to the project coordinators. The initial logged session looked as follows (see Figure 1 below):

⁴ The session lasted approximately one hour. During the first 30 minutes students communicated in English and in the second half hour they communicated in French. Thus the students were, in turn, both native speakers and learners.

Sample of log session
(MaryC = NS of English / You = NNS of English)

MaryC says “oops”
MaryC says“sorry, i accidentally disconnected for a moment”
You say “ it's nothing”
MaryC says “daphne, i hope you don't mind me correcting you... you can say 'it's nothing' in english but it's not very common; not like saying 'rien' in french, maybe better to say "it's ok" or “no problem””
You say “I'm sorry”
MaryC says “don't be sorry! just a suggestion”
You say “You're kind”

Figure 1. Sample of log session

Once collated, all the sessions constituted a new corpus⁵ which was truly bilingual (interplay of native and learner turns), and which, despite the use of the written mode, was really representative of spontaneous interactions (synchronous, immediate response from a ‘real’ person, i.e. not just the teacher or the classmates). To enable further automatic analysis, the corpus was annotated using basic search and replace strategies. Sentence boundaries were included and tags were assigned to each native and learner turn. As shown in Figure 1, the log sessions display only the name of one of the participants, whilst the other is labeled as ‘You’. The purpose was to anonymize the data and identify each participant with a unique code depending on his/her role in the interaction (e.g. John Smith is ENS1 – English native speaker 1 – when he communicates in his English mother tongue and FNNS1 – French non native speaker nr 1 – when he communicates in French; similarly Jeanne Durant is FNS1 – French native speaker 1 – when she communicates in her French mother tongue and ENNS1 – English non native speaker 1 – when she communicates in English). Figure 2 shows the same extract as in Figure 1 but with the new annotation system.

<ENS1>“oops”</ENS1>
<ENS1>“sorry, i accidentally disconnected for a moment ”</ENS1>
<ENNS1>“ it's nothing ”</ENNS1>
<ENS1>“daphne, i hope you don't mind me correcting you... you can say 'it's nothing' in english but it's not very common; not like saying 'rien' in french, maybe better to say "it's ok" or “no problem””</ENS1>
<ENNS1>“I'm sorry”</ENNS1>
<ENS1>“don't be sorry! just a suggestion ”</ENNS1>
<ENNS1>”you're kind”

Figure 2. Annotated version of the corpus.

Output from the corpus was then used in class with the learners to work on speech acts. Figure 3 presents a screenshot of concordances of the ‘don't know’ chunk. The concordances were done on the English NNS subcorpus with the help of *WordSmith*; They were later used in class in the framework of language awareness activities on ways of ‘asking for help’.

⁵The design criteria of the corpus were: writers of a similar age (c. 20 y.), similar educational background (2nd y. at university), with controlled L1s and L2s, performing the same task, within the same time limit.

N	Concordance	Set	Tag	Word No.	File	%
1	ht there was something wrong with the chatroom" <S> "I don't know" AngelaB> im just warning you your French			7,359	or-1.doc	5
2	is it? " Benedicte> The town isn't so big" Benedicte> I don't know exactly but you need 25 minutes to go round i			56,705	or-1.doc	33
3	S> "ok I'll use my imagination!!" Nadja> let's go" <S> "I don't know how many words we have I'd say we enough!			30,362	or-1.doc	18
4	eeek?????" <S> "It's pretty early for you, you lucky boy!(I don't know how to say it?)" <S> "Last week I did the glo			42,433	or-1.doc	25
5	orry but i don't understand what you mean" <S> "Sorry I don't know how to put the accents on the words and my			1,215	or-1.doc	1
6	<S> "It's called Universite Catholique de Louvain" <S> "I don't know if I have understand well the question" Barry			43,957	or-1.doc	26
7	" Nadja> Yes, enough!!! We have here a congregation (don't know if it is the right word) of all the german speak			493	or-1.doc	0
8	ur Dad important to you? Nadja> Do you speak gaillic? Don't know if it is written like that!" Fritz> it? how so?"			75,765	or-1.doc	44
9	? " Nadja> What kind of initiations"" <S> "I don't know if it's the right word? In French it is called Ba			13,965	or-1.doc	8
10	is instead: I Nadja> I have to pass an English test an I don't know if my English is good enough." <S> "your en			398	or-1.doc	0
11	s one french expression"bouffe" which means food. But I don't know if that is the same." <S> "see I can't even ge			73,078	or-1.doc	42
12	here are many Kot à projets (project kot)" Benedicte> I don't know if you know what a kot is. Do you know?" <S			56,394	or-1.doc	33
13	ings as well. Is your town near the Dutch border" <S> "I don't know if you know that there are two languages in o			44,160	or-1.doc	26
14	musical called 'Guys and Dolls'. I hope it's good." <S> "I don't know it? Is it a new one?" Emmak> Je penserai			13,322	or-1.doc	8
15	bouf- & don't even know if that's right??" Melanie> see I don't know many expressions either. lol is about the onl			73,039	or-1.doc	42
16	d ie" <S> "ok" RoryG> you're very selfconfident" <S> "I don't know this word" <S> "I can speak easily?" RoryG			18,142	or-1.doc	11
17	h is very good" <S> "No, because you believed my that I don't know what Coke is." <S> "Sorry believed me" Dick			20,609	or-1.doc	12
18	cab" Benedicte> ok" Benedicte> reseau" Benedicte> I don't know who to say about that , I'm rn not very good at			31,482	or-1.doc	19

Figure 3. Concordances of 'don't know' - UCL/TCD corpus – English NNS section

The concordances include example sentences not directly related to the 'asking for help' language function (such as 'I don't know exactly but you need 25 minutes to...') but many of them are directly related (e.g. 'I don't know how to say it?', 'I don't know if it's the right word?', 'I don't know what Coke is.').

The versatility of the corpora presented allows teachers to work on specific types of learner populations, to assess longitudinal progress, to compare learner and native populations, but also to address aspects of text types and genres (written vs spoken, formal vs informal). When compared to multi-million word corpora, the corpora presented above may look rather small, but these highly pedagogically relevant learner corpora prompt teachers and learners alike to adopt a more qualitative, manageable and sound pedagogical approach which is defined by Seidlhofer (2002) as 'learner-centred', 'context-dependent' and 'culture-bound'. The fact that learners analyze their own productions favours the individualization of learning and teaching, and helps learners monitor their own productions, together with the effects that those productions have on others.

2.2. Learner corpora as multi-purpose and sustainable resources

Learner corpora such as those presented in section 2.1. are not only multi-purpose in nature (they can be used for numerous types of focus on form activities: lexis, grammar, pragmatic aspects, etc.) but are also sustainable resources in that they can be built on progressively over time, be safely stored, and reused. Week after week and year after year, teachers spend long hours correcting, annotating or commenting on learners' tasks and performances and, very often, much of this feedback enterprise gets lost along the way. Teachers hand the sheets back to students, students keep those sheets for a year (at best) and, at the end of the school year, some well-organized students eventually put the sheets in an archive box which is never opened again. Taking the 'digital turn' (Wible 2008) makes it possible to store and re-use the time-consuming and highly valuable feedback work done by the teachers. Correcting essays and giving feedback (be it positive or negative feedback) constitutes a substantial part of teachers' everyday work. Collecting essays in electronic format, recording some of the oral interactions and keeping track of the feedback provided are the very first steps of learner corpus use in the classroom.

Beyond the sustainability issue, working with corpora also facilitates individualization of teaching and learning. Teachers can easily go back to an individual learner's productions (not through an exclusive quantitative access to progress as validated through grades given, but a real qualitative

access to the work itself) and learners can easily, and regularly, go back to or use their own productions as yardsticks against which to assess their own progress (or lack thereof).

2.3. Widening understanding of the learner corpus concept

The collection of learner corpora is usually coordinated by university teams, as shown by the list of learner corpora around the world which is available on the website of the Centre for English Corpus Linguistics in Louvain (see <http://www.uclouvain.be/en-cecl-lcWorld.html>).⁶ This might give the impression that learner corpus collection is a university-led enterprise. At the same time, an increasing number of teachers – also at primary or secondary levels of education – use learning-management systems (LMS), course management systems (CMS), or virtual learning environments (VLE) which include the key ingredients and tools for successful collection, annotation and exploitation of learner corpora. This sometimes leads to paradoxical situations where teachers, when asked whether they use learner corpora in their classes, reply negatively, whilst they regularly use an LMS⁷ to collect learner data which actually qualify as learner corpus data (see definition provided in the introduction to the article). One of the reasons behind this mutual misunderstanding is purely lexical: Users of LMSs will use more general terms such as ‘aggregated log sessions’, ‘student’s productions’, ‘essays’, ‘tasks’ but will never (or hardly ever) use the specific term ‘corpus’. The ‘corpus’ is only one of the ingredients of LMSs which include both educational and administrative tools. The Innovative Learning website (see http://www.innovativelearning.com/learning_management/index.html) lists the following features for LMSs: Management of users, roles, courses, instructors, and facilities, generation of reports, learner messaging and notifications, assessment/testing, handling of student pre/post testing, display of scores and transcripts, grading of coursework, chat and forum facilities, wikis, web-based or blended course delivery.

An initial interim conclusion that may be drawn is that there might potentially be many more learner corpora out there than first expected. A further interim conclusion is that teachers who collect these learner-corpora-to-be do not advertise them, which is a pity. A possible corollary of this is that teachers have learner data/corpora at hand (or could easily collect learner corpora) but do not always know how to exploit them to the full (see Section 3 for some suggestions).

3. Creating, using (and reusing) your own learner corpus

The present section will inevitably be limited in scope as its aim is not to provide an exhaustive list of all the options available but rather to encourage teachers to start collecting their own corpus by:

- presenting some freely available tools that can help teachers collect a learner corpus;
- illustrating some of the pedagogical tasks that can be carried out on the basis of learner corpora.

The suggestions provided in this section start from the premise that the target user can be one single teacher and that access to a computer room is not compulsory. The only prerequisite is that the teacher and the pupils should have access to one computer and internet connection, whether at home or elsewhere.

As Krajka explains (2007: 36):

The prevalence of computers, increased opportunities of Internet access, availability of large amounts of target language data, all of these call for language teachers’ greater interest in the active use of corpora both for the classroom (in materials development) and in the classroom (for learner discovery tasks). Contrary to the pre-Internet era, when corpus consultation procedures were largely restricted to linguists and lexicographers due to technological, financial and logistical considerations, the

⁶ The list also includes information on the target language, the first language of the learners, language medium, text or task type, proficiency level, number of words and availability of the learner corpora.

⁷ Whilst LMSs, CMSs and VLEs each have their specificities, the term LMS will be used here as an umbrella term

language teachers of the Web 2.0 age will find it much easier to access, compile and consult corpora for language teaching.

In his article, the author (*ibidem*) presents an overview of available corpus resources and explains how teachers can compile customized native corpora (via the web, for instance) to best suit the needs of particular teaching demands. However, the article focuses exclusively on native corpora. In the present section two examples will be presented for the collection of teacher-made⁸ learner corpora.

3.1. Using freely available LMSs as a corpus tool box

One of the key advantages of using an LMS (see section 2.3. for a more complete list) is that it enables the collection, storage and management of electronic documents. These documents can be uploaded to the system by students (and/or teachers) at any time. The documents can later be shared with other users, annotated, marked, revised by learners and/or teachers, and be used as raw material for the creation and development of further teaching and learning practices. Several open-source LMSs are available (such as Moodle⁹, Claroline¹⁰ or Blackboard¹¹) but Moodle has been selected to illustrate this section. The community of Moodle users is extremely large and information and tutorials on Moodle are easily accessible. As stated on the Moodle.org website (<http://moodle.org/about/>), it was created to give educators the best tools to manage and promote learning. It can be used for very large student populations but is also meant to be used in small-scale projects. Users have access to numerous activity modules (forums, databases, wikis) to build richly collaborative communities of learning but can simply use Moodle as a way to deliver content to students and assess learning using assignments or quizzes. A demonstration site is also accessible to new users.

The most interesting feature for learner corpus collection is that learners, once logged onto the system, can upload their files. The teacher can then download all the submissions and create his/her own customized corpus. Submissions can also be annotated by teachers, grades can be assigned and comments or feedback provided. The system is very flexible and teachers can grant access to annotated files to individual students or to groups of students.

The teacher can use the raw files (text only) or the annotated files (e.g. with error codes inserted in front of erroneous words or expressions) to create data-driven learning materials for subsequent use in the classroom. Teachers can also download the interactions between the students and use them as a basis to create a corpus (see examples provided in section 2). Using such an LMS helps develop the digital literacy of students and teachers alike and offers an easy way of collecting ad hoc learner corpora.

The problem with Moodle (and many other LMSs) is that other tools have to be used to analyse or annotate the corpus created. One such example is Byrne (2007) who reports on the use of a software program called Markin to insert feedback into the students' essays. He exported the students' forum contributions from the LMS he used into the Markin programme (which has extensive feedback possibilities in the form of annotation buttons) and once the annotation phase was done, he reimported the students' annotated work in the LMS. Similarly, if the teacher wants to retrieve concordances, he/she will have to use a concordancing programme (see articles in Sinclair 2004, O'Keeffe et al. 2007, and Aijmer 2009 for more concrete examples on how to use corpora in the classroom; see also Krajka 2007 for a presentation of some online concordancing tools).

As the use of multiple tools or software may, however, prevent teachers from actively using corpora, a more integrated tool, specifically designed for teachers, will be presented in the following section.

⁸ For an overview of already existing learner corpora and their availability, see <http://www.uclouvain.be/en-cecl-lcWorld.html>

⁹ <http://moodle.org>

¹⁰ <http://www.claroline.net/index.php?lang=en>

¹¹ <http://www.blackboard.com/Platforms/Learn/Overview.aspx>

3.2. GOLD: a web-based learner corpus collection system for teachers

The Centre for Advanced Language Proficiency Education and Research (CALPER) at Pennsylvania State University has developed the Graphic Online Language Diagnostic (GOLD), which is described as a user-friendly web-based assessment instrument. Michael McCarthy, the initial corpus project director, created GOLD¹² as a means of assessing the development of advanced proficiency through learner corpora. GOLD “enables language teachers to create their own corpora from their students’ spoken and written performance and to analyze and benchmark student performance. It also allows teachers to trace the development of individual students or groups of students over time” (<http://calper.la.psu.edu/corpus.php>).

Using GOLD requires registration, but teachers can get free access. They simply have to fill in a short form explaining what use they intend to make of the system. Once registered, teachers can easily upload files from their students into the system. In addition to its extreme user-friendliness, it contains built-in tools both for basic statistical analysis and for concordancing facilities. Very clear tutorials are also provided (see http://calper.la.psu.edu/corpus_portal/tutorial_overview.php) and include the following topics:

- learn what a corpus is and find out more about how corpora have been developed;
- discover how corpora have been used in a variety of different areas and for a variety of different purposes;
- download corpus tools and get started doing corpus analysis;
- discover how corpora can help both language learning and teaching;
- take a tour of some corpus-based materials and see how they are different from more traditional ones;
- consider the broader potential of corpora and reflect on how you can incorporate what you have learnt in this tutorial in your own teaching;
- discover how teachers can obtain corpora they need.

GOLD also provides facilities which can be very useful when uploading files, such as the ‘paste special’ and ‘unformatted text’ which helps avoid copying image files which could be included in a text. Another example is the ‘compare corpus’ facility which enables quantitative and qualitative comparisons between two learner texts or corpora, but also between some native corpora that the teacher would have uploaded to the system.

4. Concluding remarks

The important role of corpus consultation literacy has been put forward by O’Sullivan (2007) and can reasonably be said to be one of the multiple literacies of our time. Gilquin et al. (2007) have gone as far as labelling learner corpora the missing link of pedagogy. Corpus literacy is pedagogically relevant as it involves learners in language awareness/discovery procedures and can be used to foster the co-construction of knowledge when collaborative projects are carried out.

The purpose of this paper was basically threefold: First, to explain the reasons behind the current lack of uptake of learner corpora in classrooms; second, to present the advantages of using learner corpora, which are truly multi-faceted and sustainable resources; and finally, to describe some existing and freely-accessible tools that teachers can access to collect and analyze their own tailor-made learner corpus.

Learner corpora should no longer be ignored by teachers and should become part and parcel of their teaching toolbox. The examples presented in section 2 have shown how learner corpora can be used in the classroom to promote authentication and pedagogical relevance. They have demonstrated how the learners’ own productions can be better exploited to promote learning and individualisation.

However, the limits of, or difficulties in, using learner corpora should not be ignored. Teachers have to invest some of their time to discover the new tools which are available and to learn how to use

¹² GOLD was developed by CALPER faculty and staff and funded by a grant from the U.S. Department of Education and through support from the Centre for Language Acquisition at Penn State University.

them, but it should be added that the availability of web-resources such as GOLD will probably ease their task. It is to be hoped that teachers will be able to co-operate in teams both within their own schools and across several schools, as is the case for much larger projects such as the one described by Wible et al. (2001). The options offered by new technologies (accessibility, data sharing, storage) will probably provide enough incentives to boost learner corpus literacy in classrooms.

References

- Belz, J. & N. Vyatkina (2008) "The Pedagogical Mediation of a Developmental Learner Corpus for Classroom-Based Language Instruction", *Language Learning & Technology* 12 (3), pp. 33-52.
- Boulton, A. (2008) "DDL: reaching the parts other teaching can't reach?", in A. Frankenburg-Garcia (ed.) *Proceedings of the 8th Teaching and Language Corpora Conference*, Associação de Estudos e de Investigação Científica do ISLA-Lisboa, Lisbon, pp. 38-44.
- Boulton, A. (2009a) "Testing the limits of data-driven learning: language proficiency and training", *ReCALL* 21 (1), pp. 37-54.
- Boulton, A. (2009b) "Data-driven learning: reasonable fears and rational reassurance", *Indian Journal of Applied Linguistics* 35 (1), pp. 81-106.
- Braun, S. (2005) "From pedagogically relevant corpora to authentic language learning contents", *ReCALL* 17 (1), pp. 47-64.
- Braun, S. (2006) "ELISA - a Pedagogically Enriched Corpus for Language Learning Purposes", in S. Braun, K. Kohn & J. Mukherjee (eds) *Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods*, Peter Lang, Frankfurt, pp. 25-47.
- Breyer, Y. (2009) "Learning and Teaching with Corpora: Reflections by Student Teachers", *Computer Assisted Language Learning* 22 (2), pp. 153-172.
- Byrne, T. (2007) "Marrying Two Existing Software Packages into an Efficient Online Tutoring Tool", *Computer Assisted Language Learning* 20 (5), pp. 459-469.
- Feak, C.B. & J.M. Swales (2010) "Writing for publication: Corpus-informed materials for post-doctoral fellows in perinatology", in N. Harwood (ed.) *English Language Teaching Materials. Theory and Practice*. Cambridge University Press, Cambridge, pp. 279-300.
- Gabrielatos, C. (2005) "Corpora and Language Teaching: Just a fling or wedding bells?", *TESL-EJ Journal* 8 (4), pp. 1-35.
- Gilquin G. & S. Granger (2010) "How can data-driven learning be used in language teaching?", in A. O'Keeffe & M. McCarthy (eds) *The Routledge Handbook of Corpus Linguistics*, Routledge, London, pp. 359-370.
- Gilquin, G., S. Granger & M. Paquot (2007) "Learner corpora: the missing link in EAP pedagogy". *Special issue of Journal of English for Academic Purposes* 6 (4), pp. 319-335.
- Granger, S. & F. Meunier (2010) "SLA research and learner corpus research: Friend or foe?", paper presented at the 20th Annual Conference of the European Second Language Association. Reggio Emilia, Italy, 1-4 September 2010.
- Granger, S. (2009) "The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation", in K. Aijmer (ed.) *Corpora and Language Teaching. Studies in Corpus Linguistics* 33, Benjamins, Amsterdam, pp. 13-33.
- Granger, S., J. Hung & S. Petch-Tyson (eds) (2002) *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Benjamins, Amsterdam.
- Jones, M. & N. Schmitt (2010) "Developing materials for discipline-specific vocabulary and phrases in academic seminars", in N. Harwood (ed.) *English Language Teaching Materials. Theory and Practice*, Cambridge University Press, Cambridge, pp. 225-248.
- Kaltenböch, G. & B. Mehlmauer-Larcher (2005) "Computer corpora and the language classroom: on the potential and limitations of computer corpora in language teaching", *ReCALL : the Journal of EUROCALL* 17 (1), pp. 65- 84.
- Kasper, G. & K. Rose (2002) *Pragmatic development in a second language*. Blackwell, Oxford.
- Krajka, J. (2007) "Corpora and Language Teachers: From Ready-Made to Teacher-Made Collections", *CORELL: Computer Resources for Language Learning* 1, pp. 36-55.

- McCarthy, M. (2007) "Assessing Development of Advanced Proficiency Through Learner Corpora", position paper. Available at http://calper.la.psu.edu/downloads/pdfs/CALPER_ALP_Corpus.pdf.
- Meunier F. & C. Gouverneur (2009) "New types of corpora for new educational challenges: collecting, annotating and exploiting a corpus of textbook material", in K. Aijmer (ed.) *Corpora and Language Teaching*. Benjamins, Amsterdam, pp. 179-201.
- Meunier, F. (2010) "Learner Corpora and English Language Teaching: Checkup Time", *Anglistik: International Journal of English Studies* 21 (1), pp. 209-220.
- Meunier, F. (2011) "Corpus linguistics and second/foreign language learning: exploring multiple paths", . RBLA International special issue, Corpus studies: future directions. S. Gries (guest editor), *RBLA* 11 (2), pp; 459-477.
- Myles, F. (2008) "Investigating learner language development with electronic longitudinal corpora: Theoretical and methodological issues", in L. Ortega & H. Byrnes (eds) *The Longitudinal Study of Advanced L2 Capacities*, Routledge, New York, pp. 58-72.
- Nesselhauf N. (2004) "Learner corpora and their potential for language teaching", in J. Sinclair (ed.) *How to Use Corpora in Language Teaching*, Benjamins, Amsterdam, pp. 125-152.
- O'Sullivan, I. (2007) "Enhancing a Process-Oriented Approach to Literacy and Language Learning: The Role of Corpus Consultation Literacy", *ReCALL* 19 (3), pp. 269-286.
- O'Keeffe, A., M. McCarthy, & R. Carter (2007) *From Corpus to Classroom. Language use and language teaching*. Cambridge University Press, Cambridge.
- Rizzardi M.C., L. Pedrazzini & A. Nava (2008) "The Role-Play Learner Corpus: a resource for investigating learner language", in A. Martelli & V. Pulcini (eds) *Investigating English with Corpora. Studies in Honour of Maria Teresa Prat*. Polimetrica Publisher, Monza, pp. 379-398.
- Römer, U. & S. Wulff (2010) "Applying corpus methods to writing research: Explorations of MICUSP", *Journal of Writing Research* 2 (2), pp. 99-127.
- Römer, U. (2006) "Pedagogical Applications of Corpora: Some Reflections on the Current Scope and a Wish List for Future Developments", *Zeitschrift für Anglistik und Amerikanistik. Special Issue: The Scope and Limits of Corpus Linguistics – Empiricism in the Description and Analysis of English*, V. Gast (ed.) 54 (2), pp. 121-134.
- Römer, U. (2011 – forthcoming) "Corpus research applications in second language teaching". *Annual Review of Applied Linguistics*.
- Ruiz-Madrid, M. N. (2007) "Pedagogical Boundaries for Web Resources. An Integrative Proposal for the FL Classroom", *CORELL: Computer Resources for Language Learning* 1, pp. 56-67.
- Seidlhofer, B. (2002) "Pedagogy and local learner corpora: Working with learning-driven data", in S. Granger, J. Hung & S. Tyson (eds) *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Benjamins, Amsterdam, pp. 213-234.
- Vannestal, M.E. & H. Lindquist (2007) "Learning English Grammar with a Corpus: Experimenting with Concordancing in a University Grammar Course", *ReCALL* 19 (3), pp. 329-350.
- Wible, D. (2008) "Multiword expressions and the digital turn", in F. Meunier & S. Granger (eds) *Phraseology in Foreign Language Learning and Teaching*, Benjamins, Amsterdam, pp. 163-181.
- Wible, D., C.W. Kuo, F. Chien, A. Liu & N.L. Tsao (2001) "A web-based EFL writing environment: integrating information for learners, teachers, and researchers", *Computers and Education* 37, pp. 297-315.