

Learning from learners: a non-standard direct approach to the teaching of writing skills in EFL in a university context

Miguel Fuster-Márquez & Carmen Gregori-Signes

To cite this article: Miguel Fuster-Márquez & Carmen Gregori-Signes (2016): Learning from learners: a non-standard direct approach to the teaching of writing skills in EFL in a university context, Innovation in Language Learning and Teaching

To link to this article: <http://dx.doi.org/10.1080/17501229.2016.1142549>



Published online: 26 Apr 2016.



Submit your article to this journal [↗](#)





View related articles [↗](#)



View Crossmark data [↗](#)

Learning from learners: a non-standard direct approach to the teaching of writing skills in EFL in a university context

Miguel Fuster-Márquez  and Carmen Gregori-Signes 

Instituto Universitario de Lenguas Modernas Aplicadas (IULMA), Universitat de València, Valencia, Spain

ABSTRACT

Corpora have been used in English as a foreign language materials for decades, and native corpora have been present in the classroom by means of direct approaches such as Data-Driven Learning (Johns, T., and P. King 1991. 'Should you be Persuaded'- Two Samples of Data-Driven Learning Materials. In *Classroom Concordancing*, 1–16. Birmingham University. English Language Research Journal 4.). However, the suitability of using learners' output in classroom tasks remains controversial. This paper describes a pilot study in the application of a non-standard direct approach where Spanish university students are invited to reflect on their production. In the experiment, carried out in several sessions during the course, the students were exposed to a selection of erroneous sentences from their compositions. Prior to the classroom activity, the teacher contrasted the learners' sentences with correct versions produced by native English speakers. A relevant part of the methodology consisted in getting learners collectively involved in finding the errors and suggesting improvements. After that, solutions were discussed through the analysis of the alternative sentences provided by the native students. The results show that students are willing to accept this methodology as a supplement to textbooks' proposals. We claim that authentic and highly specific learner data obtained from a reliable ad hoc learner corpus and direct exposure to these data through controlled activities may cover certain learners' needs not found in textbooks.

ARTICLE HISTORY

Received 29 September 2015
Accepted 7 January 2016

KEYWORDS

Direct approach; learner corpus; EFL; data-driven learning; university education; learner autonomy

Introduction

Native and learner corpora have been widely used in corpus-informed English as a foreign language (EFL) materials for decades in what is known as the indirect approach (see Römer 2008, 113; also Braun 2005, 51, and more recently Granger 2015); however, some difficulties and challenges have been reported in relation to the introduction of corpora in the classroom by means of direct approaches such as Data Driven Learning (DDL) or similar methodologies (see Johns and King 1991; Bernardini 2004; Chambers 2010; Cheng 2010; Gilquin and Granger 2010; Philip 2010; Frankenberg-Garcia 2014), in EFL environments. Fuster-Márquez (2010) claims that in content courses such as English lexicology and morphology, offered in a university context where English is a foreign language, the selection of simple activities that involve the exploration of native corpora by means of an interface such as Mark Davis' *corpus.byu.edu* is a reasonable approach. This practice does not require complex technological knowledge; all it requires is that the instructor be familiar with the interface, the nature of the data in the corpus being explored and set reasonable teaching objectives.

On the other hand, the suitability of using learner data in the classroom instead of using examples produced by native speakers in direct approaches is more controversial. There are very few reports in

the literature of practitioners who have attempted to introduce learner data in the classroom via DDL. In all probability the reason is that presenting students with negative evidence is thought to be inappropriate. By contrast, EFL textbooks and other reference materials often include exercises based on errors produced by learners. A well-known example is Swan's *Practical English Usage* (2005), a reference work used by EFL teachers and students worldwide that has gone through several editions, although we are left in the dark about the way these pedagogic materials are employed by teachers or accessed by learners (cf. Granger 2015). The main objection to the implementation of learner data in the classroom in EFL literature is not related to the content, but rather to the way in which such data are presented to the learners and the teaching contexts where this approach might be desirable.

In this paper, we describe a pilot study that involves the direct exposure of students to learner data contained in an ad hoc learner corpus, partly following the model described by Seidlhofer (2002). The target group included University students from different nationalities in a General English language course taught during the fall semester, 2014 at the University of Valencia. These EFL learners, who ranged from an intermediate to an advanced level, were invited to reflect collectively on their own written production. The methodology is a combination of corpus and non-corpus based approaches. The process consisted of two main stages: the preparation stage and the implementation stage. The preparation stage involved the identification by the instructor of learner errors contained in the *ad hoc* corpus which contained the students' most recent essays. These errors were, in turn, checked against those found in the larger CASTLE database,¹ a learner corpus of half a million words which is currently being compiled at the University of Valencia containing the production of students in the degrees of English Studies and Translation Studies at said university between the years 2008 and 2015. This prior step helped to ensure that the selection of problem sentences in learners' essays that were to be shown and discussed in class, were sufficiently relevant for the whole group. Also, four native speakers of English (Erasmus students from British universities) were each asked to propose correct versions of the selected deviant sentences. A crucial part of the implementation in class consisted in getting learners collectively involved in identifying errors and suggesting improvements. After that, their solutions were discussed together with the analysis of the alternative sentences provided by the native students.

While the activity proved that students are willing to accept this methodology as a supplement to textbooks' proposals, the major contribution in this research comes from the fact that it focuses on the correction of actual deviant forms of learners, through the implementation of an eclectic direct approach. This kind of activity is neither found in textbooks, nor is it ordinarily provided when teachers give feedback.

In the next section we will review the relevant literature on direct corpus approaches concerning the use of native and learner corpora. This is followed by sections on the methodology and a discussion of the main findings in this pilot study. The conclusion summarises some of the most important points, outlines directions of future research and highlights the pedagogical implications of the study itself.

Native and learner corpora in direct approaches

Römer (2006) establishes two main pedagogical uses of corpora. On the one hand, 'the indirect approach is researcher-focused and centres on corpus evidence and the impact it can have on syllabus design or teaching materials' (2006, 125). The outcomes of the indirect approach are found in learners' dictionaries, pedagogical grammars or coursebooks (see discussion in Granger 2015). On the other, the direct approach allows 'teachers and learners get their hands on corpus data themselves, instead of having to rely on the researcher as mediator or provider of corpus-based materials' (Römer 2006, 124). The proposal discussed here may be classified as a non-orthodox application of the direct approach. We shall focus on the underpinnings of this methodology while highlighting similarities and differences with other direct approaches.

In direct approaches (see Römer 2008, 113; also Braun 2005, 51), learners explore corpus data either through activities designed by their teachers or independently by getting unmediated access to the data therein. The most widely discussed methodology in the literature is DDL, which was introduced by Johns, and defined (Johns and King 1991, iii) as 'the use in the classroom of computer-generated concordances to get students to explore regularities of patterning in the target language and to develop activities and exercises'. In Johns' view, DDL allows students to act as 'language detectives' since they are able to carry out corpus searches independently. Meunier claims that such an approach is suitable as an awareness-raising activity (2002, 130). For Odlin (1994, 319–320), DDL 'gives central importance to developing the learner's ability to "puzzle out" how the target language operates from examples of authentic usage'. In his view, DDL empowers learners by promoting their autonomy rather than turning them into passive receivers of information.

Although the implementation of DDL and other direct approaches have been considered to be motivating, a number of drawbacks have been reported. First, as argued by Philip (2010) and Gilquin and Granger (2010), DDL can be time-consuming and not always feasible. Likewise, despite the belief in the potential benefits of DDL, its effectiveness in language learning has not yet been confirmed (see Mukherjee 2009; Gilquin and Granger 2010). Boulton (2012) suggests adopting more realistic alternative methodologies, such as surfing the net, or using google-based concordancers such as *Webcorp* (<http://www.webcorp.org.uk/live>). Boulton argues that students are more acquainted with the internet and *google* and this may favour their acceptance over the use of corpora. Boulton (2012, 24) gives a list of advantages that the internet offers as a language learning resource:

While the web may not be a prototypical corpus in terms of linguist research, we can at least treat it as 'corpus surrogate' (Bernardini, Baroni, and Evert 2006: 10ff) which may be quite fit for purpose. Its advantages in language teaching include its size, recency, variety (whatever you want is probably there somewhere), availability (free), reliability (the web itself doesn't crash, or impose limits on the number of simultaneous users), speed, flexibility, and so on. Just as importantly, it is already familiar to learners, especially via Internet search engines such as Google.

These, among other reasons, may help us understand why corpus approaches are less popular among EFL teachers today. For Boulton (2012, 25) the Internet and web-based concordances could be adopted as convenient intermediate steps before moving to more hard core direct approaches. Nevertheless, both the direct approach and surfing the web are generally viewed by practitioners as complementary methods rather than alternative options to ordinary teaching methodologies. To date, there are practically no reports in the literature of corpus methodologies being adopted by teachers as the main sources of teaching. Also, most published research refers to controlled experiments and/or convenient additions to other normal classroom activities in university contexts, not in lower contexts (see Boulton 2009).

Unarguably, one of the attractive factors which has given corpus classroom methodologies a boost in foreign language teaching (FLT) contexts has been the possibility of allowing quick access to large amounts of authentic native data, that is, data which has not been manipulated by materials writers and otherwise unavailable. The issue of accessing authentic language data has been widely discussed in the literature, although the term 'authenticity' is open to different interpretations in second language acquisition (SLA) and FLT (see Widdowson 1978; Breen 1985; Gilmore 2007). For Buendgens-Kosten (2014, 457):

'Authenticity' is a frequently invoked and, at the same time, keenly debated notion in ELT which became an issue when the advent of Communicative Language Teaching in the 1970s brought with it a new focus on 'realism' in language learning materials and activities.

More often than not, the interpretation given to authenticity in SLA has been the use of genuine written or spoken language samples produced by native speakers. Strikingly, Minematsu (2014, 135) observes that '[e]ven though a communicative approach has emphasised the use of authentic texts, exercises in the course books or activities made by teachers in EFL classroom tend to be based on invented examples.' Such an interpretation of authenticity stems from Sinclair's (1996) linguistic

view when he referred to it as '[a]ll the material gathered from the genuine communications of people going about their normal business'. Earlier, Widdowson (1978, 89) used the term 'genuineness' to refer to this interpretation of authenticity. Therefore, it should be pointed out that the fact that a text is produced by a native speaker is not sufficient to grant it the status of authentic. In order to do so, the language sample should have been extracted from a real-life activity. By contrast, inauthenticity applies to language samples collected 'in experimental conditions or in artificial conditions of various kinds'. Instances of inauthentic data might be contrived texts and activities in textbooks specifically designed for learning purposes (see discussion in Clavel-Arroitia and Fuster-Márquez 2014).

Granger (2002, 8) reminds us that

the notion of authenticity is somewhat problematic in the case of learner language. Even the most authentic data from non-native speakers is rarely as authentic as native speaker data, especially in the case of EFL learners, who learn English in the classroom.

She contends that the context in which students learn adds a certain degree of artificiality due to the fact that learner production is tied to specific tasks and restrictions imposed by teachers. Nevertheless, she also takes the view that although assignments written by non-native speakers (NNS) in these EFL environments are produced in a very particular context, they cannot be classified as inauthentic (see also Mendikoetxea, Murcia Bielsa, and Rollinson 2010, 182–183).

In this respect, while the authenticity of native speech gathered from real-life contexts is a feature of most research reported in relation to DDL practices, the authenticity of our data is related to a learner corpus containing tasks gathered under specific conditions. We understand that our learner data, as acknowledged in the literature, is also authentic insofar as it represents a variety of non-native English, even if it has been collected in a classroom setting (see Granger 2003). We claim that classroom assignments produced by NNS such as free writing compositions are a manifestation of their language behaviour.

Authenticity needs to be related to sociolinguistic norms in specific language communities. Following Kachru's (1985) classification of English varieties around the world, we should acknowledge differences according to norms that exist in an inner circle (L1, so-called native speakers), an outer circle (L2, bilingual communities where English plays a relevant role) and an expanding circle (or L3, with speakers of English as a foreign or international language). The authenticity of learner production gathered from the least artificial EFL tasks such as free writing lies in the fact that it is the natural language produced by L3 speakers of the expanding circle, which represents not only their interlanguage but, more importantly, how they interact with other speakers of English in real-life situations. In the context of English usage worldwide, the identification of 'errors' arises from the comparison between the speech of L1 speakers, considered as norm-providers, and that of L3 speakers whose use of English is more occasional, produced in socially restricted linguistic interactions or in the learning conditions of a foreign environment, most typically the classroom. Although this issue will not be pursued any further in this paper, we cannot dismiss the political and cultural dimensions that lie behind the connection between nativeness and authenticity, having negative consequences on perceptions of NNS interlanguage, or in judgements of the adequacy of non-native teachers (see discussion in Alptekin 2002; Buendgens-Kosten 2014). Needless to say, the legitimacy of the English used by members of the inner circle as norm-providers has seldom been challenged in EFL (see, however, Seidlhofer 2005).

A different matter is to what extent it is pedagogically advisable that students be exposed to the data encountered in NNS corpora. For example, Nesselhauf (2004, 140) is in favour of this approach in a limited way 'for points which have already been covered in the classroom, possibly even repeatedly, but which the learners nevertheless still get wrong'. The moot point is how students should be exposed to such data, since non-native data constitutes 'negative evidence'. Nesselhauf (2004, 140) goes on to suggest that negative evidence be complemented by positive evidence and that follow-up exercises be proposed to stabilise correct usage. Mukherjee (2009, 213) proposes that

DDL practitioners consider the variant procedure of offering both negative and positive evidence from learner data, so that learners:

... do not get the impression of learner output being treated exclusively as a hotch-potch of mistakes and errors – it is neither desirable nor useful to establish a rigid dichotomy between good and correct usage in native data on the one hand and bad and incorrect usage in learner output on the other.

Whichever approach is adopted, be it by means of genuine native data or by means of learner data, Senior (2005, 65) recommends that teachers do not lose sight of their pedagogical purposes, that is, ‘what precisely we want our students to learn from these materials’. In SLA, the relevance of authentic texts and, indeed, of any kind of text or activity in the classroom should be primarily grounded in clear pedagogical objectives. As Breen (1985, 63) has aptly remarked, it is important to understand that ‘Perhaps *the* criteria to guide the teacher’s selection and use of texts (both written and spoken) reside initially, not in the texts themselves, but in the learners.’ With regard to exposure to learner errors, Mendikoetxea, Murcia Bielsa, and Rollinson (2010, 184) underscore that

[p]resenting the students with the errors they, or other co-learners, make and asking them to correct them is an activity appropriate for raising grammatical consciousness in an attempt to bridge the gap between what is known and the forms of the TL.

Methodologies which include learner production are typically applied to samples of languages with relevant learner errors. In this regard, we adhere to the views advocated by practitioners who favour that NNS should be exposed to tasks which also include native production, where learners may find the correct forms, of the negative evidence under scrutiny.

Methodological description

Similarities and differences with other approaches

Seidlhofer (2002, 216) claims that building a local learner corpus is pedagogically meaningful because it allows us ‘to be fine-tuned to specific learners and local conditions of relevance’. Gilquin and Granger (2010, 361) also underscore that such corpora ‘even go one step further, as they contain data produced by the very same students who will be using the corpus’. This kind of learner corpus, containing texts produced by Spanish students of EFL was the basis for this methodological approach. However, we implemented a more eclectic direct approach, which is partly based on DDL experiences reported in the literature with some significant departures from it. We claim that this methodology is at least as recommendable as some direct approaches or other more ordinary DDL methodologies, since we believe it contributes to improve the learning process by focusing not only on problem areas, but also on the solution of attested needs the students have by providing

Table 1. Differences and similarities between a typical DDL approach and an ad hoc learner corpus-based methodological approach.

	DDL approach	Ad hoc learner corpus-based approach
Learner direct exposure	Yes	Yes
Use of a large corpus	Yes	No
Data presented in KWIC (concordance) format	Yes	No
Quantitative analysis	Yes	No
Qualitative analysis	Yes	Yes
Inductiveness	Yes	Yes
Use of native data	Yes	Yes
Use of learner data	No	Yes
Authenticity	Yes	Yes/No
Focus on single language aspects	Yes	No
Focus on multiple language aspects	No	Yes
Teacher’s control	Yes/No	Yes
L2/L3 target learners	Yes	Yes

native input. The main similarities and differences between the two methodologies can be found in Table 1.

A relevant departure from standard DDL methodologies is that of exposing students to their data as learners. As illustrated in Table 1, a methodology that includes learner data together with non-genuine native data essentially differs from standard direct approaches.

As argued in the earlier section, exposing learners to genuine native language samples is a core aspect of DDL. An illustrative example of the way this direct approach would work in the classroom is shown in Figure 1. Thus, a teacher may wish to make learners aware of top noun collocates of synonymous adjectives such as *big* and *large*. This contrast can be carried out by showing them the results obtained from the *Corpus of Contemporary American English* (COCA), a corpus of 520 million words, collected between 1990 and 2015, through Mark Davies's interface [corpus.byu.edu/coca].

The frequency numbers next to each collocate point, in fact, towards the restriction of collocates, for example, *big trouble* is found 801 times (W1) in COCA, but there are no examples (W2) of **large trouble*; *large extent* has a frequency of 768 times (W1) in COCA, whereas **big extent* is found only once (W2).

However, in DDL such an overall observation of contrasting frequencies is not pedagogically sufficient and should be followed by a close analysis of concordances (i.e. contextualised examples culled from real native use) where each of the contrasted adjectives is seen together with specific nouns, as shown in Figure 2 featuring adjective *large*. In a Key Word in Context (KWIC) fashion, for this particular search, learners might scrutinise and learn from those genuine cases where natives prefer one or the other adjective. Finally, conclusions about the appropriate use of one or the other adjective in relation to lists of nouns are inductively drawn by learners.

In this regard, we do not claim that the data provided by the native speakers (NS) who participated in the task proposed in our study are authentic. The sentences produced by the informants were introduced as pedagogically useful additions, relevant to the task, with a view to shedding light on non-native features of learners and to provide in a novel way native alternatives to learner's production. However, while nativeness is represented in the sentences produced by our informants, these samples are correct forms that are entirely based on those that have been produced by learners; consequently, they cannot be regarded as genuine.

That said, the native rewrites in our approach bear some resemblance to Levenston (1978) and Cohen's (1983, 1989) proposal known as 'reformulation'. In several studies, Cohen discusses the benefits of using native speakers as 'reformulators' with the purpose of enhancing the second-language speaker's style, and/or as an addition to the insufficient traditional feedback on learner errors provided by teachers, who might happen to be nonnatives. As in the case of reformulators

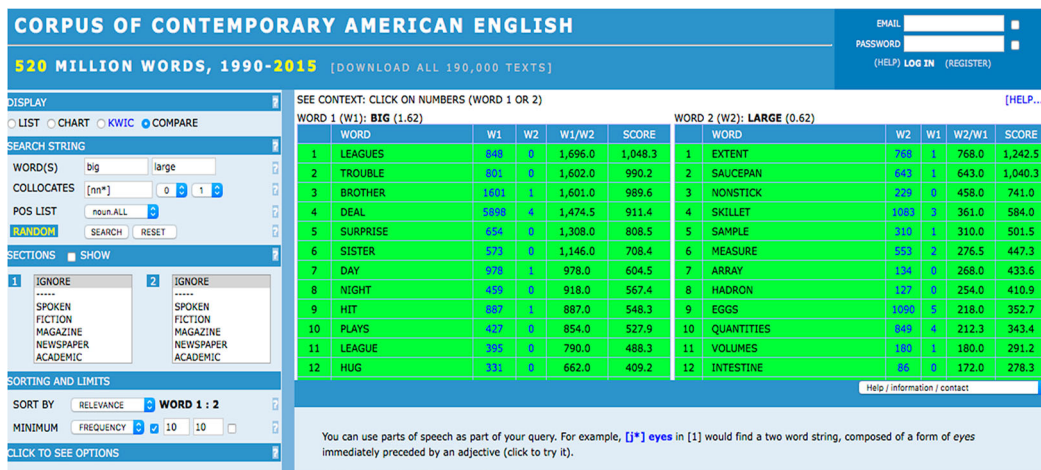


Figure 1. Contrast of noun collocates (one position right) with *big* and *large* in COCA [http://corpus.byu.edu/coca/].

Line	Year	Genre	Source	Text Snippet
58	2002	ACAD	GeographRev	... was the southernmost part of the district that first attracted large numbers of settlers , encouraged by railroad building in the
59	2003	MAG	MotherEarth	... large and small bulbs. As a general rule , the onions are kept for the table and the small ones are used
60	1994	NEWS	Chicago	... Solidarity union. # From the first, Gorecki composed for orchestral and choral forces as well as smaller ensembles. He
61	1997	SPOK	NPR_Science	... what it does is it disrupts the normal circulation over a part of the world , and it affects the weather from India
62	2009	MAG	USAToday	... empires, peacefully and without vast bloodshed: is still in part to lady Thatcher 's combination of strategic firmness and
63	1998	ACAD	EarNoseThroat	... There is evidence that the pillars are in contact with particels in the cell membrane, however there is no indication
64	1994	MAG	AmpSpect	... balance just barely in the black; but then two more payments push back into an overdraft of nearly \$10,000. One
65	2004	ACAD	SportBehavior	... then converted into percentages (see Table 1). A percentage of skaters strongly agreed or agreed with several
66	2004	MAG	OutdoorLife	... the bottoms of large rivers, lakes and sounds . A percentage of the diet of many divers consists of mollusks,
67	2001	FIC	FantasySciFi	... in that direction. I caught up with him beside a blazing fountain . A marble sea nymph was toting a dolphin who
68	2005	NEWS	Chicago	... is vital, he added. Fitchie Creek will have a pool exercise and fitness center, hot tub, library,
69	1999	SPOK	NPR_Science	... had a complicating situation with the fact that we had a population who were immun compromised with HIV-AIDS and, as
70	2012	MAG	Newsweek	... participated with a flowery parade. My father painted a portrait of the young queen and placed it at the front of
71	2003	NEWS	WashPost	... There are, however, a few funds that own a proportion of hedge stocks in their portfolios. One of the best
72	1999	MAG	Smithsonian	... from the southeast bring in cold, nutrient-rich water , a reason why the Galapagos has such lush marine life. Two divers
73	2008	FIC	Bk:Damien	... matter off with the flick of a long-fingered hand , the ruby of the ring on his middle finger winking one of its
74	1990	ACAD	SocialStudies	... poverty, crime, drug use, and unemployment or if segments of our population believe that they have no attachment
75	2005	MAG	CountryLiving	... Remove pan from the oven and transfer its contents to a servng bowl . 3 Finish the soup. Remove cup of the
76	2010	MAG	GoodHousekeeping	... tomatoes have softened. 4. Drain pasta; transfer to servng bowl . 3 Add vegetable mixture. With tongs, toss until
77	2007	SPOK	CBS_Sixty	... give it a try. God, that is a very share of Mr-MARTIN : Are you clear with everything I've said
78	1991	ACAD	NaturalHist	... plaster casts of the four teeth. Because of the unusually size of few of the Homo erectus specimens from Java,
79	2002	ACAD	OrthoNursing	... Some collar sizes are misleading. That is , a size refers to a long, slender neck and not a large
80	2007	ACAD	DrugIssues	... of loneliness was apparent, even among those reporting a social network and good social support. It is important to
81	1997	ACAD	Raritan	... down by their feet. Nonetheless, the belief in a southern land region led Ptolemy in the second century C.E. to
82	2002	ACAD	Archaeology	... an old woman tending a shrine at the foot of a stone statue . Angkor was built with great precision, taking

Figure 2. Concordance of *large* followed by a noun in COCA [<http://corpus.byu.edu/coca/>].

in Cohen's work, in our study native speakers silently correct learners output without altering the content and also without giving any additional reasons for their rewrites. Explaining learners' choice choices might require skills these native students do not possess and, strictly speaking, it is not pedagogically necessary for the purpose of the task.

However, our proposal is closer to what is known as 'reconstruction' (see Levenston 1978), where specific problem areas or errors are discussed at sentence level, although such reconstruction is carried out by native speakers, not the teachers. Another relevant difference with Cohen's 'reformulation' is that while that methodology is designed for individual students, in our case all learners participate actively in the methodology by having samples of every learner's output shown and analysed collectively in the classroom.

In spite of the highlighted departures from more standard practices, we agree in the main with Meunier (2002, 134) that by comparing learner and native data in direct approaches: (1) learners become researchers and discover the differences between their interlanguage and the target language (sense of discovery, of motivation), (2) have access to the errors or infelicities in their production but also to what is correct and valid and (3) reinforce negotiation, interactivity and interaction (among learners and between learners and teachers). Although we acknowledge that the methodology reported in this pilot study is not in keeping with standard DDL methodologies, the pedagogical purpose behind these contrasts between native and non-native data have striking parallels to DDL, while it avoids some of the main pedagogical drawbacks of DDL mentioned in the literature (see, e.g. Ädel 2010).

Data collection and task preparation

This research has been conceived as a small scale pilot study focused on a methodological approach where the target learners are a group of 35 students of *English Studies*, a four-year degree offered by the University of Valencia (Spain). In this degree students have to take a range of EFL courses, and also linguistics, culture and literary courses in one or more languages. The target group belongs to a group of *English Language 3*, a 15-week EFL course where the learners were expected to reach a B2 (intermediate) level of proficiency in English at the end of the fall semester of 2014. We collected one of their essays in an early session during the course.

Some key factors in the make-up of the group led us to the planning of this pilot study. The 35 students in the group came from different L1 backgrounds. Four different nationalities were present (Polish, Turkish, French and Spanish), although the majority were Spanish. Their ages ranged between 20 and 22 years old, as to their gender, 6 were males, and 29 female students.

This linguistic and cultural diversity is becoming a common feature in recent times at the University of Valencia in *English Studies* and it is illustrative of similar trends at other European universities as a result of the Erasmus exchange programme. Therefore, as regards their L1, the target group did not have a homogeneous profile. Moreover, the group had a mixed-ability class within the range of an intermediate-to-advanced continuum. It was also found that the set textbook was not fully satisfactory to cope with this rather heterogeneous learning context. These factors convinced us of the need to plan classroom activities that were sufficiently meaningful and motivating for all the students, where they could see their specific learning needs had been met.

The 35 students in the group were asked to write a short free argumentative essay (between 350 and 400 words) around the topic of parents' role in education. The task had to be completed as homework within a week and the students were allowed to use any kind of reference material (paper or online dictionaries, grammars, etc.).

A sample of 35 sentences was carefully selected from this ad hoc local corpus containing the 35 learners' compositions. Each of the sentences was taken from a different composition to ensure that errors found in each of the compositions were present. The choice of sentences was partly determined by overall frequencies of specific errors observed in the ad hoc corpus and larger databases such as CASTLE but, in the main, by the fact that they represented a variety of lexical, phraseological, syntactic or discourse problems this group of learners had. Thus, this selection process will, we assume, gather together common problem areas that are shared by most if not all students. These problems are firm candidates attested in a larger learner corpus, but most particularly in the ad hoc corpus that contains their own production.

On average, the selected sentences contained more than one error, which implied that an important number of problems could be discussed. All the examples shown were duly anonymised when presented in class. The psychological thrust for showing one sentence per learner was that every student could perceive that all their peers, even the most advanced, could still make mistakes, and therefore the procedure conveyed that errors were just a normal feature of interlanguage development. Pedagogically, the motivation was that we could get all students in class involved in our proposal.

Prior to the implementation of the activity in the classroom, and as part of the overall plan, after the selection of erroneous learner sentences, the native students who assisted us were asked to produce a correct version of each of these 35 sentences produced by the learners. This strategy was adopted instead of simply asking them to identify and mark essays. Marking learner writing is a complex assessment task which required training and experience the native British students did not possess. We gave NS two specific instructions. Firstly, that they could not consult with each other any aspect of their written version and secondly, that they should attempt to preserve as much as possible the style of the learner's sentence, only focusing on improvements related to specific problems or errors they detected. The first instruction was intended to obtain individual versions, and the purpose of the second was that they would not produce too distant versions from those produced by the learners so that they were difficult to handle in class. The result was that the instructor obtained four native versions of the sentences from the native speakers and, more importantly, most of these sentences focused on relevant problem areas. Consequently, in our presentation in class, we were able to follow the advice that in corpus approaches where learners are exposed to learner data, it is advisable to offer these together with correct native models. Further methodological features mentioned in Table 1 will be discussed in the following sections.

Classroom implementation

Four consecutive sessions, each one lasting 20 minutes out of a 2-hour class, were scheduled to deal with these 35 learner sentences, and an average of 2–3 problem areas in each sentence. This means that over 70 real errors could be discussed. The procedure consisted first in projecting on the screen a number of learner sentences and giving the whole class time to identify collectively in small groups

any possible errors. Here are three of those sentences which were analysed in class together with the learners:

- (1) By other hand I believe that each child needs a specific way of education.
- (2) However, there are advantages and inconvenients of that kind of education that combine learning and playing.
- (3) At last, a lot of studies shows that is better to let them choose their preferences before imposing your opinion and of course the funny learning, it is a fact that you learn more if you enjoy it.

Discussion would be followed between learners and teachers about the identification of problems and the suggestion of possible objective solutions. From a pedagogical viewpoint, this procedure could make learners feel empowered by actively participating in the process while perceiving that the teacher, though offering guidance, was not the sole authority to be heard in this respect, their peers could very well provide answers to problems. As in the case of DDL activities, the teacher adopted the role of mediator in the whole process. Immediately after the discussion with the learners in the classroom, the teacher showed the 'problem areas' and presented the students with the alternative rewritings produced by the native speakers. Relevant problem areas in each sentence and the specific solution(s) given by each native speaker were highlighted as shown below.

It should be noted that the KWIC view often adopted in DDL approaches did not make sense in this activity. KWIC is particularly useful to interpret a sufficient number of similarities, having in mind a single focus, in a corpus by means of a vertical interpretation, which was not the case. Instead, a sentence presentation was adopted where they could easily locate the problem areas and the correct forms provided by the NS as shown below for the three sentences we have just seen

(1) **By other hand** I believe that each child **needs a specific way of education**.

W. On the other hand, I believe that each child needs to be taught in an individual manner that suites the way in which they learn best.

X. On the other hand, I believe that each child needs to be educated differently.

Y. On the other hand, I believe that each child needs to be educated in their own specific way.

Z. On the other hand, I do believe that a child's core education is of the utmost importance.

(2) However, there are advantages and **inconvenients** of that kind of education **that combine** learning and playing.

W. However, there are advantages and disadvantages from the type of education that attempts to combine learning and playing.

X. However, whilst there are advantages, there are also inconveniences that come with this kind of education that combines learning and playing.

Y. However, there are both advantages and disadvantages of education that combines learning and playing.

Z. This is not to say the method of learning while playing games comes without its fault.

(3) **At last**, a lot of studies **shows that is better** to let them **choose their preferences before** imposing your opinion and of course **the funny learning**, it is a fact that you learn more if you enjoy it.

W. Finally, many studies have shown that it is more beneficial to let children follow their own preferences rather than imposing our own expectations on them. Also, it is true that you learn more if you enjoy the manner in which you are being taught.

X. Finally, a lot of studies show that it is better to let children choose their preferences before parents impose their own opinion. Of course, it is a fact that one learn more if one is enjoying the learning process.

Y. Finally, many studies show that it is better to let children make decisions rather than force them to follow your own ideas, as it is proven that children learn better doing things that they enjoy.

Z. Recently, many studies have shown that it is better to let the children decide what they wish to learn, before the parents try to impose their preferences. Moreover, the learning while playing method has proved to be effective, as children learn more if they enjoy it.

The selected sentences represented different kinds of complexity and for each one more than one problem area was identified. A wide range of infelicities were identified: lexical, grammatical, discursive, stylistic, etc. Learners noticed that these problem areas were all intimately related to the correct forms provided by the native speakers who were able to identify them, at least implicitly, as deviant language forms. It should also be noted that while in most cases, three native versions followed the

instructions given by the teacher, at times the last native student offered distant versions, so these versions were hardly useful in the discussions of specific error-solution plan. However, at least three of the native versions were enough for the purpose of error identification and the provision of suitable alternatives.

Discussion of results

The learners were exposed to their own production as learners and then it was correlated with correct native versions so as to round off their own reflections in relation to the identified errors and providing them with positive evidence. The fact that for each problem area 'slightly' different alternatives were offered by the native students was a key feature of this methodology which to a certain extent added subjectivity. Learning to cope with the subjectivity of the alternative native sentences forced learners to think inductively. During the whole process the teacher monitored the task and adopted a less authoritative voice. This kind of activity is thought to be beneficial at advanced learning stages, and particularly suitable with regard to the profile of our learners. University students have to cope with the richness of native choices in the target language and interpret them more critically and independently. However, despite the disparity of native alternatives for each sentence, the students were able to perceive a high degree of convergence in terms of (1) error detection or problem areas, and (2) very frequently agreement in the alternative solutions. Thus, subjective perceptions were quite often minimal and highlighted by the teacher as a final remark.

Not unlike a number of DDL experiments (see Martínez 2002; Berardo 2006), the sentences produced by learners and shown in class required more than a single focus, since they contained more than one error, and not infrequently even more than one. Some of the most obvious errors were related to discourse phrases. For example, the prepositional phrase 'by other hand' in sentence 1 was identified as an ill-formed sequential unit corresponding to native 'on the other hand', with a linking function; also, 'at last' in sentence 3 was misused when the learner might have used 'finally'. Corpus research shows that phraseological units are often wrongly constructed and acquired by the students and their discourse function not properly assimilated. Another interesting case was the deviant binomial sequence 'advantages and inconvenients' in sentence 2, where the learner had incorrectly transferred the expression 'ventajas e inconvenients' from Spanish. Two out of four NS produced the fairly common combination 'advantages and disadvantages' for that same purpose. Thus, the teacher could safely pinpoint that it was an excellent phraseological sequence to learn. In other cases, an erroneous construction was detected but no single ideal solution was given, such is the case of the combination 'the funny learning' in sentence 3, where the learner probably intended to mean that 'learning should be enjoyable or fun'. Nevertheless, the teacher could highlight that 'fun' and 'funny', though morphologically related, have quite different meanings. This common confusion among learners is often dealt with in intermediate courses and thus this activity helped to reinforce those found in their textbook since the correct form had not been assimilated yet.

Our methodological proposal is not as ambitious as Cohen's (1983, 1989), which envisages natives rewriting entire essays produced by second-language learners and perhaps further rewriting and improvements. Rather, our approach is closer to reconstruction. However, the teacher monitors the whole learning process but is not the agent of reconstructed sentences. Instead, the learners are granted the active role in the discovery of problems and the discussion of valid alternative sentences.

An important ingredient of any methodology is that it wins the acceptance of learners. In this respect, students appeared to be enthusiastic and willing to accept the proposal as a way of complementing exercises in textbooks. Partly the reason was that they were curious about specific errors they and their peers made in writing and eager to see the alternatives offered by the native Erasmus students, but also the reason was that it gave them a chance to escape from the routine of textbook activities.

Conclusion

We believe that the methodological approach outlined here is feasible for instructors in university education. A local learner corpus such as the one described in this study is not difficult to assemble, and the implementation of the activities proposed does not require technological expertise. Most direct approaches which make use of native data are designed as awareness-raising activities. The methodological approach described in this paper, however, may go beyond merely drawing students' attention to error since it proposes a collective revision process of learners' recent production which happens to be directly relevant to the enhancement of their writing skills. It does not deal with potential problems, but with the actual ones which are attested in their writing. However, we agree with Philip (2010, 10) that 'a balance has to be struck between the needs of the class as a whole and those of individual learners' in the selection of relevant tasks. Not only a careful prior analysis of the ad hoc corpus, but also of larger learner corpora together with observations from research on well-known difficulties of the recipient learners by the instructor can guarantee that the selection of data the target learner group will be exposed to is of interest and relevant to every student in class.

As in the empirical study reported by Mendikoetxea, Murcia Bielsa, and Rollinson (2010, 184), learners in our heterogeneous target group found the task highly motivating. However, it should be said that, to our knowledge, direct approaches reported in the literature are hardly ever described as substitutes of other regular teaching activities, but rather as activities where learners are given the opportunity to access more meaningful learning experiences. Here authenticity is to be interpreted as activities which are conducive to learning (see Breen 1985). Admittedly, General English textbooks may include exercises that take into consideration common errors associated with a specific level, and are often addressed to students who share the same L1s, but these textbooks do not show sufficient contextual background knowledge. In this respect, the compilation of an ad hoc learner corpus can be more than an adequate solution as it allows instructors to empirically observe in greater detail and in real time what specific areas of difficulty their students actually have. Learner corpora may be used in different ways. We acknowledge that the methodology described in this pilot study differs from standard practices, most particularly by the fact that the positive evidence provided by NS was not the product of a real-life activity. However, pedagogical reasons guided our choice, and our view is that the fact that learners were able to access different 'native' versions was attractive. Admittedly, the lack of availability of native speakers that the task requires can be a drawback. This difficulty was reported by Cohen (1989), who put forward several alternative solutions. However, our approach is not as demanding on natives as Cohen's approach. In our case, they are only required to rewrite sentences, but do not participate in discussions with learners. If native speakers are not directly available, teachers may (e-)mail sentences from learners to a few native students, together with instructions regarding the way the rewriting process should be carried out, either living in the area, or ask natives colleagues or students from one or more universities.

The methodology described in this paper does not cast doubt on the relevance of direct approaches or the implementation of other non-corpus approaches where native speakers and authentic material are used. On the contrary, it bears witness to the resourcefulness of corpus-driven approaches with native and non-native data. Yet, an unsolved problem in DDL experiments, and therefore also here, is its effectiveness. In this respect, a first step for future research might be to enquire about the acceptance of the methodology. This can be done by means of surveys, obtaining feedback from the target learners, etc. A second problem concerns the target EFL learners. Boulton (2010) remarks that most published DDL research draws on experimental work with university students and that it is difficult to ascertain whether similar experiments could be replicated at lower levels. Be that as it may, it may be a suitable methodology for the teaching of writing to EFL learners in a university context, where students are older, as a rule more proficient and, very importantly, expected to be more active, autonomous and to adopt more critical attitudes towards the information that reaches them.

Note

1. CASTLE: 'Learner Corpus of English as a Second and Third Language' (GV/2014/022).

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Miguel Fuster-Márquez is associate professors in the Department of English and German Philology and researchers of applied linguistics at IULMA (Instituto Universitario de Lenguas Modernas Aplicadas), University of Valencia (Spain). Fuster-Márquez's main research interests are corpus linguistics and its applications to persuasive discourse and the foreign language classroom in tertiary education. Recent publications include *The authenticity of real texts in advanced English language textbooks* (2014, co-authored by B. Clavel-Arroitia) and *Lexical bundles and phrase frames in the language of hotel websites*.

Carmen Gregori-Signes is associate professors in the Department of English and German Philology and researchers of applied linguistics at IULMA (Instituto Universitario de Lenguas Modernas Aplicadas), University of Valencia (Spain). Gregori-Signes's research interests lie in the area of discourse analysis, corpus linguistics and the integration of ICT in teaching and learning English as a foreign language. Recent publications include *Digital Storytelling and Multimodal Literacy in Education* (2014) and 'Share your Experience' *Digital Storytelling in English for Tourism* (2014, co-authored by M. Alcantud-Díaz & A. Ricart-Vayà).

ORCID

Miguel Fuster-Márquez  <http://orcid.org/0000-0003-3813-1770>

Carmen Gregori-Signes  <http://orcid.org/0000-0002-5735-5633>

References

- Ädel, A. 2010. "Using Corpora to Teach Academic Writing: Challenges for the Direct Approach." In *Approaches to English language Teaching*, edited by M. C. Campoy, B. Bellés-Fortuño, and M. L. Gea Valor, 39–45. New York: Continuum.
- Alptekin, C. 2002. "Towards Intercultural Communicative Competence in ELT." *ELT Journal* 56 (1): 57–64.
- Berardo, S. A. 2006. "The use of Authentic Materials in the Teaching of Reading." *The Reading Matrix* 6 (2): 60–69.
- Bernardini, S. 2004. "Corpora in the Classroom: An Overview and some Reflections on Future Developments." In *How to use Corpora in Language Teaching*, edited by J. Sinclair, 15–36. Amsterdam: John Benjamins.
- Bernardini, S., M. Baroni, and S. Evert. 2006. "A WaCky Introduction." In *Wacky! Working Papers on the Web as Corpus*, edited by M. Baroni and S. Bernardini, 9–40. Bologna: Gedit. Accessed June 1. <http://wackybook.sslmit.unibo.it>
- Boulton, A. 2009. "Testing the Limits of Data-Driven Learning: Language Proficiency and Training." *ReCALL*, 21 (1): 37–54.
- Boulton, A. 2010. "Learning Outcomes from Corpus Consultation." In *Exploring New Paths in Language Pedagogy: Lexis and Corpus-Based Language Teaching*, edited by M. Moreno Jaén, F. Serrano Valverde, and M. Calzada Pérez, 129–144. London: Equinox.
- Boulton, A. 2012. What Data for Data Driven? In *EuroCall (EUROCALL 2011 – The CALL Triangle: student, teacher and institution*. Accessed May 10. http://eurocall.webs.upv.es/documentos/newsletter/papers_20%281%29/07_boulton.pdf
- Braun, S. 2005. "From Pedagogically Relevant Corpora to Authentic Language Learning Contents." *ReCALL* 17 (1): 47–64.
- Breen, M. 1985. "Authenticity in the Language Classroom." *Applied Linguistics* 6 (1): 60–70.
- Buendgens-Kosten, J. 2014. "Authenticity." *ELT Journal* 68 (4): 457–459.
- Chambers, A. 2010. "What is Data-Driven Learning?" In *The Routledge Handbook of Corpus Linguistics*, edited by A. O'Keefe, and M. McCarthy, 345–358. Abingdon: Routledge.
- Cheng, W. 2010. "What can a Corpus Tell us about Language Teaching?" In *The Routledge Handbook of Corpus Linguistics*, edited by A. O'Keefe and M. McCarthy, 319–332. Abingdon: Routledge.
- Clavel-Arroitia, B., and M. Fuster-Márquez. 2014. "The Authenticity of Real Texts in Advanced English Language Textbooks." *ELT Journal* 68 (2): 124–134.
- Cohen, A. D. 1983. "Reformulating Second-Language Compositions: A Potential source of Input for the Learner." Revised version of a paper presented at the annual convention of TESOL, Toronto, Ontario, March 15–20: 1–25.
- Cohen, Andrew D. 1989. "Reformulation: A Technique for Providing Advanced Feedback in Writing." *Guidelines* 11 (2): 1–9.
- Frankenberg-García, A. 2014. "The Use of Corpus Examples for Language Comprehension and Production." *ReCALL* 26 (2): 128–146.

- Fuster-Márquez, M. 2010. "The Challenges of Introducing Corpora and their Software in the English Lexicology Classroom: Some Factors." In *Language Windowing Through Corpora*, edited by I. Moskowich-Spiegel, B. Crespo, I. Lareo, and P. Lojo, 269–288. Coruña: Universidade da Coruña.
- Gilmore, A. 2007. "Authentic Materials and Authenticity in Foreign Language Learning." *Language Teaching* 40 (2): 97–118.
- Gilquin, G., and S. Granger. 2010. "How can We Prepare Learners for Using Language Corpora?" In *The Routledge Handbook of Corpus Linguistics*, edited by A. O'Keeffe, and M. McCarthy, 359–370. Abingdon: Routledge.
- Granger, S. 2002. "A Bird's Eye of Corpus Learner Research." In *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*, edited by S. Granger, J. Hung, & S. Petch-Tyson, 3–36. Amsterdam: John Benjamins.
- Granger, S. 2003. "Error-Tagged Learner Corpora and CALL: A Promising Synergy." *CALICO Journal*, 20 (3): 465–480.
- Granger, S. 2015. "The Contribution of Learner Corpora to Reference and Instructional Materials Design." In *The Cambridge Handbook of Learner Corpus Research*, edited by S. Granger, G. Gilquin, & F. Meunier, 486–510. Cambridge: Cambridge University Press.
- Johns, T., and P. King 1991. "'Should you be Persuaded'- Two Samples of Data-Driven Learning Materials." *English Language Research Journal* 4: 1–16.
- Kachru, B. 1985. "Institutionalized Second-Language Varieties." In *The English Language Today*, edited by S. Greenbaum, 211–226. Oxford: Pergamon.
- Levenston, E. A. 1978. "Error Analysis of Free Compositions: The Theory and the Practice." *Indian Journal of Applied Linguistics* 4 (1): 1–11.
- Martinez, A. G. 2002. "Authentic Materials: An Overview on Jaren's Linguistic Issues." Accessed September 10. <http://www3.telus.net/linguisticsissues/authenticmaterials.html>.
- Mendikoetxea, A., S. Murcia Bielsa, and P. Rollinson. 2010. "Focus on Errors: Learner Corpora as Pedagogical Tools." In *Corpus-Based Approaches to English Language Teaching*, edited by M. C. Campoy, B. Bellés-Fortuno, and M. L. L. Gea-Valor, 180–194. London: Continuum.
- Meunier, F. 2002. "The Pedagogical Value of Native and Learner Corpora in EFL Grammar Teaching." In *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*, edited by S. Granger, J. Hung, and S. Petch-Tyson, 119–142. Amsterdam: John Benjamins.
- Minematsu, K. 2014. "Using a Corpus in an EFL Classroom to Develop Depth of Vocabulary Knowledge." *The Tsuru University Review* 79: 133–146. Accessed June 6. <http://trail.tsuru.ac.jp/dspace/bitstream/trair/656/1/Y-079133.pdf>.
- Mukherjee, J. 2009. "The Grammar of Conversation in Advanced Spoken Learner English: Learner Corpus Data and Language-Pedagogical Implications." In *Corpora and Language Teaching*, edited by K. Aijmer, 203–230. Amsterdam: John Benjamins.
- Nesselhauf, N. 2004. "Learner Corpora and Their Potential for Language Teaching." In *How to use Corpora in Language Teaching*, edited by J. M. Sinclair, 125–156. Amsterdam: John Benjamins Publishing.
- Odlin, T. 1994. *Perspectives on Pedagogical Grammar*. Cambridge: Cambridge University Press.
- Philip, G. 2010. "Classroom Concordancing in the 21st Century: The New Generation." *Language Forum* 36 (2): 1–29.
- Römer, U. 2006. "Pedagogical Applications of Corpora: Some Reflections on the Current Scope and a Wish List for Future Development." *ZAA* 54 (2): 121–134.
- Römer, U. 2008. "Corpora and Language Teaching." In *Corpus Linguistics. An International Handbook (volume 1), [HSK series]*, edited by A. Lüdeling and M. Kytö, 112–130. Berlin: Mouton de Gruyter.
- Seidlhofer, B. 2002. "Pedagogy and Local Learner Corpora: Working with Learning-Driven Data." In *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*, edited by S. Granger, J. Hung, and S. Petch-Tyson, 213–234. Amsterdam: John Benjamins.
- Seidlhofer, B. 2005. "English as a Lingua Franca." *ELT Journal* 59 (4): 339–341.
- Senior, R. 2005. "Authentic Responses to Authentic Materials." *English Teaching Professional* 38: 71.
- Sinclair, J. 1996. *EAGLES. Preliminary Recommendations on Corpus Typology*. Accessed September 14. <http://www.ilc.cnr.it/EAGLES96/corpusyp/corpusyp.html>.
- Swan, M. 2005. *Practical English Usage*. Oxford: Oxford University Press.
- Widdowson, H. 1978. *Teaching Language as Communication*. Oxford: Oxford University Press.