Data-driven learning: taking the computer out of the equation.

Alex Boulton

Data-driven learning: Taking the computer out of the equation

Alex Boulton. 2010. *Language Learning*, 60/3.

alex.boulton@univ-nancy2.fr

CRAPEL – ATILF/CNRS, Nancy-Université

3 place Godefroi de Bouillon
BP 3397
54015 Nancy – cedex
France

Abstract

Despite considerable research interest, data-driven learning (DDL) has not become part of mainstream teaching practice. It may be that technical aspects are too daunting for teachers and students, but there seems no reason why DDL in its early stages should not eliminate the computer from the equation by using prepared materials on paper – considerably easier for the novice learner to deal with. This paper reports on a simple experiment to see how lower-level learners cope with such paper-based corpus materials and a DDL approach compared to more traditional teaching materials and practices. Pre- and post-tests show both are effective compared to control items, with the DDL items showing the biggest improvement, and questionnaire responses are more favourable to the DDL activities.

Keywords

data-driven learning; worksheets; concordance print-outs; paper-based materials; lower levels

The potential applications of electronic corpora in language teaching and learning have received considerable attention in recent years. They are now ubiquitous in the construction of reference materials – especially dictionaries but also grammars and usage manuals – and increasing numbers of textbooks and other teaching aids make at least nominal use of them in the selection and sequencing of language items to cover. It is also possible for teachers and learners to access corpora themselves, as evidenced by the growing numbers of scholarly papers in data-driven learning (DDL), to use Johns' (1991a) term. Briefly, learners are not taught overt rules, but explore corpora to detect patterns among multiple examples, a far more "natural" approach (Gaskell & Cobb, 2004, p. 304; Scott & Tribble, 2006, p. 6; Boulton, 2009). However, it is notable that DDL has not crossed over into mainstream practice or been taken up by major publishers: very few DDL materials are available ready to use (Boulton, 2008b).

The potential advantages of DDL have been detailed in numerous articles (e.g., Aston, 1998), as have a number of potential barriers (for detailed discussion, see Chambers, 2007a; Farr, 2008), three of which involve the nature of DDL research to date and which are of particular interest here. Firstly, it may simply be that the work so far has failed to convince a wider audience that the investment – in terms of time, effort, money and resources – brings sufficient pay-offs. Chambers (2007a) found that most empirical research in DDL involves small-scale, qualitative studies, while Boulton (2008a) notes that many are concerned with annex questions such as learners' behaviour, their reactions, or use of corpora as a reference tool. Furthermore, those few that do attempt a quantitative analysis of learning outcomes tend to produce results that are not statistically significant, or are relatively small. This in fact reflects much empirical research in language learning, where broad, long-term achievements are difficult to analyse, and the focus inevitably falls upon specific, immediate learning outcomes, with correspondingly limited results. Such studies are not without their uses, however, as it is a reasonable hope that a

number of focused studies might allow a larger picture to appear. The experiment described in the present paper has similarly specific aims, and can thus be seen as an additional small contribution to our overall understanding of the possibilities and the limitations of DDL.

Secondly, the bulk of the literature in DDL describes courses with advanced learners approaching complex language points using sophisticated equipment and techniques (Boulton, 2008a). This is partly inevitable, as such academic articles are usually produced by researchers working with advanced learners in university environments and comparatively abundant resources – not least the researchers' own time and skills (Chambers, 2007a, p. 12-13). There is also the permanent drive towards publishable output, hence to be original, to push boundaries further, to explore new areas of technological development. Nonetheless, the upshot may be a tendency to assume that DDL is *only* useful for advanced learners in a computer laboratory, and with experts (i.e., the researchers) devoting considerable time to training learners in small groups and to developing sophisticated corpora, software and techniques. It is perhaps not surprising that other teachers may not then see the relevance to their own local circumstances (Gabrielatos, 2005); but "to make a serious contribution to language teaching, corpora must be adopted by ordinary teachers and learners in ordinary classrooms" (Mauranen, 2004, p. 208; also Mukherjee, 2006, p. 14).

The third point follows from this: hands-on manipulation is the dominant paradigm in DDL research – Boulton's (2008a) survey of 50 empirical DDL studies found that all but eight involved learners using the computer themselves. However, it is apparent in many studies that effective hands-on DDL requires considerable training (Turnbull & Burston, 1998), and technological considerations are among the major sources of problems (Yoon & Hirvela, 2004). It is also clear that many teachers view class sessions in the computer room with trepidation for a variety of reasons (Farr, 2008), while others simply do not have regular access to sufficiently

equipped computer laboratories and technical back-up. The question which interests us here is what happens when the computer is taken out of the final equation, with learners using only printed materials prepared in advance.

Paper-Based Materials in DDL

With prepared materials and tasks, many of the decisions are taken out of the hands of the learner (Thompson, 2006), and one might fear claims that such exercises are against the spirit of DDL (Todd, 2001; Cobb, 1997, chap. 2) – if indeed they constitute DDL at all. As Bernardini (2001) puts it, "it would be inherently contradictory to prescribe a methodology when the aim of the approach is to give learners the instruments to develop their own methodologies and make their own discoveries" (p. 228). More pragmatic is Mukherjee's (2006) observation that "DDL activities can be plotted on a cline of learner autonomy, ranging from teacher-led and relatively closed concordance-based activities to entirely learner-centred corpus-browsing projects" (p. 12). What is clear is that printed data were part of part of Tim Johns' original vision (also Higgins & Johns, 1984), and one he never abandoned (cf. Johns, 2002): "experience in using concordance data reactively has indicated that it could be used proactively also in a more traditional teacher-centred setting, and has suggested also a range of concordance-based exercise types" (Johns, 1991b, p. 31). Although there have been only occasional attempts by others to promote prepared concordance print-outs from publicly available resources (e.g., Chambers, 2007b; Chambers & Kelly, 2004), the consensus seems to be that paper-based materials are not in themselves incompatible with DDL (e.g., Breyer, 2006; Frankenberg-Garcia, 2005b).

This is not to deny that direct consultation is essential if learners are to benefit fully and autonomously from the possibilities corpora have to offer, but at the same time there seems no reason to rule out more teacher-oriented applications in earlier stages:

> In order to cater for students' differing abilities and familiarity with inductive learning strategies, a gradual introduction to concordance work and extensive guidance in using concordancing strategies is recommended. This would allow individual learners to progress at their own pace towards conducting independent and productive concordance investigations. (Turnbull & Burston, 1998, p. 12)

One obvious way of ensuring this, as Lamy and Karskov Mortensen (2007) point out on the *ICT4LT* Web site, is "to provide plenty of practice with paper-based exercises first, so that students get used to inductive reasoning before they are asked to cope with the additional burden of manipulating a piece of software, however simple it may seem" (section 4.1). In other words, learners may find it easier to graduate from "soft" to "hard" DDL (Gabrielatos, 2005), or from what Cresswell (2007) calls "deductive DDL" (i.e., starting with teacher-led exercises) to fully "inductive DDL" (i.e., starting with the data on their own). While many courses introducing DDL do start with teacher-initiated queries, and even teacher-prepared data, published research tends to gloss over this and concentrate on the later, hands-on stages of corpus consultation.

Prepared materials such as concordance print-outs are certainly limited in many ways compared to hands-on investigation, but that is part of their advantage (cf. Aston, 1997; Chambers, 2007b). In particular, they reduce some of the cognitive burden in initial stages by allowing learners to focus on a single new element. It is unsurprising that learners find it difficult to get to grips with new material (the corpora), new technology (the software) and a new approach (DDL) all at once (Gavioli, 2005, p. 40) – especially at lower levels of language ability. The methodology itself is "revolutionary" enough to warrant keeping other things simple, and one way to do this is to take the computers out of the equation at the start: much research has found the technological aspects to be a substantial source of frustration (e.g., Farr, 2008), and students may even be "technophobic" (Bernardini, 2002). The use of paper-based materials has

other benefits too: no need for a computer laboratory with computers that break down and Web sites that crash, the inevitable unexpected findings, and so on. In other words, printed materials are likely to be more accessible for immediate use by regular teachers and learners in ordinary classroom situations. If it can further be shown that such materials bring benefits in themselves – and not solely as a half-way house to full, hands-on corpus consultation – then this might help to promote the concept of DDL to a wider community.

In his survey of 50 empirical DDL studies, Boulton (2008a) found only eight using printed materials, three of which were not concerned with language learning outcomes per se: Stevens (1991) looked at the use of printed concordances for testing purposes, while Johns (2002) and Hadley (2002) both gathered data on learners' reactions to and perceptions of DDL by means of questionnaires. Such research questions are decidedly important, but do not address the crucial issue of whether DDL "works". Ciezielska-Ciupek (2001) did attempt this, using printed corpus materials as a supplement to a coursebook with a group of secondary-school students in Poland. Tests on language items covered produced positive results, which persisted four weeks later with no further revision, although parts of the procedure were rather informal and the test measures not subjected to any statistical analysis.

More in-depth was a study by Allan (2006) of eighteen students of English enrolled for an advanced exam preparation course in Ireland. The experimental group worked with printed concordances from a small newspaper corpus over twelve weeks, tasks being assigned out of class. Detailed analysis of results from the Vocabulary Knowledge Scale for self-assessment showed the experimental group making greater gains than the control of five students. However, conditions between experimental and control groups were not identical, and there was considerable variation in the number of concordance tasks completed and the time spent on them.

Koosha and Jafarpour (2006) controlled their study rigorously, although the focus was considerably narrower, looking at learners' ability to derive prepositional collocations from concordances. Two hundred university students of English in Iran took part in this large-scale test over 15 classes, half as a control with traditional instruction, half in an experimental group using concordance print-outs. Cloze tests allowed comparisons before and after the experimental condition as well as between the two groups. The experimental group scored significantly higher in the use of the target language.

Two papers by Boulton (2008c, 2009) are experiment-based, examining the ability of French learners of English to cope with particular aspects of DDL with no explicit training even at relatively low levels. The first of these found significant improvement from pre- to post-tests, suggesting that lower-level learners are able to extract relevant patterns from concordance print-outs of phrasal verbs and apply the findings to new contexts. Although the focus is extremely narrow, this stage is clearly crucial for DDL to function, and has tended to be taken for granted (though see Todd, 2001). The findings were supported in the second study with similar learners, who were able to use concordances more effectively as a reference than traditional information sources (dictionaries and grammar-usage manuals) for connectors. A final test ten days later showed significant improvement compared to the pre-test, suggesting that learning did occur, although no significant difference was found between the DDL and traditional groups at this stage.

The present study builds on these past attempts at empirical analysis of paper-based materials. It is hoped that a larger body of such evidence will help DDL break out of its current research confines and into more mainstream teaching practices (Mukherjee, 2006, p. 21), especially if it can be shown that DDL has benefits for lower-level learners without expensive resources or extensive training. Simple materials of this sort, if they are seen to be effective and

practical, might counter a number of frequent objections to DDL, and contribute to greater awareness of its potential as they require little training, are easily shared, and can be incorporated into published materials (Chambers & Kelly, 2004).

Method

The experiment was designed to compare traditional teaching with paper-based DDL materials in helping lower-level learners to cope with typical problem areas, data being collected by means of pre- and post-tests as well as questionnaires. As there is no regular or reliable access to a computer laboratory with Internet, paper-based materials are essential if the students are to have access to DDL. Time is also at a premium, as all learners have to cover the same imposed syllabus for the end-of-year exam, leaving little time for training or dealing with inevitable technical problems.

The rest of this section describes the population sample, and outlines in more detail the language items, materials and procedures involved.

Participants

The learners who participated in the study were second-year students enrolled at an architectural school in the north-east of France. Although 89 students were officially enrolled, some were absent at various stages; in total 71 questionnaires were collected following the experimental session, along with 62 complete sets of test data. Of these students, 61% were women; the median age at the time was 19½; the vast majority had French as a mother tongue, although there were also three Arabic and three Chinese speakers.

The students have compulsory 90-minute English classes once a week in groups of 15-20, but none had any prior experience of DDL. The students' language learning history is typical for

France: most have seven years of English at school prior to entering higher education, though levels are not high. This is reflected in a start-of-year levels test based on a full-length *Test of English for International Communication* (TOEIC)[i]: the mean score for these students was 52.9% overall, around 450 on the official TOEIC scale (i.e., towards the lower end of the "intermediate" band of 405-600), equivalent to A2 or B1 on the *Common European Framework of Reference for Languages* (Council of Europe, 2001). This is typical of much higher education in France, closely matching the situation in earlier DDL studies of our own with engineering students (Boulton, 2007, 2008c, 2009).

Language Items

Fifteen problematic language items were selected from the learners' own productions during a regular written assignment. This method was chosen to ensure that students could see the relevance of the language involved in the experiment (cf. Seidlhofer, 2000) rather than having it imposed from general lists of problem areas such as *Common Mistakes at First Certificate and How to Avoid Them* (Tayfoor, 2004), or even those compiled from general learner corpora such as the *International Corpus of Learner English* (e.g., Granger, 2003).

On this occasion, the students were allowed 30-40 minutes to produce an argumentative essay on a topic of their choice loosely related to architecture and building, without using reference tools (see topics listed in Appendix A). Problem areas were identified, and a number of language items selected on the basis of frequency of occurrence and generalisability. The focus was on grammar/usage items as these tend to lend themselves especially well to a DDL approach rather than, say, purely spelling or semantic problems (Chambers & O'Sullivan, 2004, p. 166), and learners seem aware of this (Yoon & Hirvela, 2004). It may be that such errors do not "matter" in the sense that they rarely impede communication, as Osborne (2004, p. 259)

diligently remarks in his own study on using corpora in error-correction; his point, however, is that the reality of the institutional context means that such errors are expected to be eliminated, for better or for worse.

A preliminary list was rated by English teachers in another university, so that the final list consisted of fifteen items (Appendix B) which were generally considered typical problem areas for French learners – the type of language points which are frequently taught at lower levels but rarely well assimilated. This suggests that traditional teaching of these items is relatively unsuccessful, leaving open the possibility for alternative techniques such as DDL.

A third of the 15 items featured in the pre- and post-tests as a control and received no treatment, while the other ten were included in the experimental session. For students in half of the groups, items 1-5 were taught using traditional methods, items 6-10 using the DDL materials; the situation was reversed for the students in the other groups. In this way, no group or language item receives special treatment, and can serve as a control for the others, a technique used in other studies of corpora in language learning (e.g., Stevens, 1991; Bowker, 1999; Estling Vannestål & Lindquist, 2007).

Teaching Materials

The teaching materials consisted of two short booklets covering the same language items for all students, but reversing the traditional and DDL treatment for each. The first two pages consisted of a short introduction in French to corpora and their potential applications, along with examples. This document then presented five of the items using corpus data and DDL techniques, five using dictionary entries and traditional teaching methods. Every attempt was made to produce equivalent materials for both DDL and traditional treatment: two introductory questions for each item followed by a single page of either corpus data or dictionary information,

interspersed with specific questions to focus attention. Two examples each of the corpus-based and dictionary-based materials are provided in Appendix D.

The corpus-based materials were compiled from the BYU-BNC (Davies, n.d.) interface to the British National Corpus, 100 million words of spoken and written British English collected in the early 1990s. The interface has a number of limitations (in particular with regard to sorting or copying the concordances), but it allows queries by register, accepts part-of-speech queries and wild cards, provides information on frequency and collocates, enables direct comparison of search terms, and so on. The main advantages of this interface are that it is fast, simple to use, and available free on line – in other words, it is the type of corpus resource which is well within the reach of any English teacher with access to the Internet. The data presented consisted mainly of concordance lines, accompanied where appropriate by graphic presentations of frequency and distribution by register – variety is essential to avoid concordance burn-out (Thurston & Candlin, 1998, p. 278). The concordances themselves were in the KWIC format, with the keywords centred in bold and an average of 11 words left and right; the concordances were truncated to fit onto the page, but were not otherwise edited in any way. They were selected and sorted, and presented in manageable groups of usually 5 to 30 lines, each set of concordances introduced with a specific question to guide the interpretation. This is in line with Johns' (1991a) rationale: "all the citations shown in the handout are authentic, [although] there is in this handout a degree of 'rule-hiding' in the selection of citations, the categories adopted, and the sequencing of citations within each category" (p. 4). It is indeed one major advantage of prepared paper-based materials that such selection can be made in advance to focus attention rapidly on the desired elements, reducing the risk of overwhelming the learners with huge quantities of data and background noise and limiting the range of possible answers (Thompson, 2006).

For the traditional resources, dictionaries provide an obvious point of comparison (Yoon & Hirvela, 2004; Frankenberg-Garcia, 2005a), one of their main advantages being not the definitions, but the examples they propose (Cobb, 2003). The information here was also collected from on-line sources – again, free and easily accessible to teachers. The *Reverso* Web site (http://dictionary.reverso.net/) allows access to both bilingual and monolingual dictionaries from the same publisher (*Collins English French Electronic Dictionary*, 2005; *Collins COBUILD English Dictionary for Advanced Learners*, 2003), affording relatively coherent presentation. Monolingual or bilingual entries were selected as appropriate, and presented in the same layout on the page as the originals.

Test Instruments

The main test instruments were the same for all participants, consisting of pre- and post-tests in identical format. Each test consisted of 30 questions, two on each of the 15 language items described above. To reduce the possibilities for inadvertent error, instructions and layout were closely modelled on the familiar multiple-choice gap-fill format of the TOEIC part V for "incomplete sentences"; examples are provided in Appendix C. The questions themselves were largely based on example sentences found in pedagogical materials, though adapted as necessary to correspond to the particular problem points covered. The main source was the *Oxford Advanced Learner's Dictionary* (2007) available free on line, supplemented by other resources as necessary. One advantage of pedagogical examples is that the contexts are simple and have few distractions, enable clear and unambiguous answers, are typical of traditional testing materials and so are familiar to the students, and if anything favour the traditional teaching paradigm over the experimental condition.

In addition to the pre- and post-tests, a short questionnaire was compiled in French to be completed individually by participants present at the experimental session (Appendix E). It combined closed questions on a 5-point Likert scale (ranging from *strongly agree* to *strongly disagree*) with open questions to be completed in the students' own words.

Procedure

The experiment took place during normal class time with the students' regular teachers. One of the teachers had some second-hand knowledge of DDL, but none had any prior experience of using it themselves. A one-hour training session was organised on DDL, and especially on the materials and their use in class in order to maximise consistency across all groups. While all teachers exhibited interest in the experiment, some remained unconvinced the DDL approach would work with their students.

The pre-test was conducted in week 1, the experimental condition in week 2, followed by the post-test after an interval of three weeks so that the post-test would be of medium-term recall. The pre- and post-tests were conducted towards the end of class and collected in for scoring by the researcher, but no feedback on test scores was provided during the experiment. They were presented as TOEIC practice and adhered precisely to that format; the students were also told that they included only items which had been the source of errors in their own earlier work.

The experimental session itself lasted one hour of a 90-minute class in week 2, introduced by the teacher informing the class that they were going to look at some of the language points covered in the previous week's test. As these are known to be difficult, the teacher explained s/he was going to introduce a new technique for some of them. The booklets were distributed, and the teacher spent about five minutes going through the introductory section and the examples. Each language point was then introduced by means of two whole-class discussion questions, after

which the students were directed to study the information and the guiding questions in the booklets.

For the traditional treatment, teachers did not need to change their usual preferred style of teaching, although they were asked to adopt a comparatively traditional "knowledge transmission" paradigm prevalent in France (Brown, 2007), which can be caricatured as follows: "Your teacher is the guide and mentor, who will show you what to learn and how to learn it. Listen to your teacher and do as you are told" (Willis, 2003, p. 167). For the DDL treatment, students were encouraged to work in pairs or small groups so that they could come up with their own answers, with minimal teacher intervention. Once each item had been completed, whole-group feedback allowed the teacher the opportunity to clarify or correct the students' findings, but only if necessary. While even such limited teacher input may seem to go against the spirit of hands-on DDL, it is nonetheless a frequently reported practice (e.g., Estling Vannestål & Lindquist, 2007), and is expected in the case of traditional paper-based materials; it would also have been ethically questionable to have deliberately let students go away with erroneous ideas in the name of research.

Teachers were told to cover the language points in their entirety whatever the treatment, and to check at the end as the whole group returned to the original questions before proceeding to the following item. The intention was thus that all students should come away with essentially the same final information – the main differences lay in the way it had been arrived at and the materials it was derived form. It took between five and ten minutes to go through each item, although all teachers were surprised that the DDL treatment did not take substantially longer than teaching from the dictionary-based materials.

Finally in this session, the questionnaire was distributed for students to report on their impressions of the different materials and approaches.

Results

Test Results

The overall scores are fairly low – the mean score in Test 1 was less than 50% – even though the test questions were based on clear examples and offered only four possible answers. The highest scores were 25 out of 30 in Test 1 and 27 in Test 2; the lowest were 6 and 8 respectively. As Table 1 shows, the mean scores increased from 14.56 in Test 1 (SD = 5.06) to 17.31 in Test 2 (SD = 5.26), an improvement of 2.82 points (SD = 2.80), or 19.38%.

Table 1 Descriptive data for mean test results

There are two main ways to compare the data: changes between the two tests, and differences between the three treatments. A two-way ANOVA for repeated measures shows there to be a significant improvement overall between tests (F = 17.79, $p < .0001$). One possibility is that there may simply have been a "test effect", with students scoring higher the second time simply as they became more used to the test design and what was required. There was indeed a small improvement in the control items of 5.32%, but this is not significant ($p > .5$). This means that the significant improvement must derive from the other items: a 22.2% increase in score for the traditional items ($p < .01$), 31.6% for the DDL items ($p < .0001$). The first conclusions therefore are that the test effect is minimal, while both kinds of treatment do have a significant effect.

The key question now is whether DDL was significantly more effective than traditional teaching. The results shown in Table 1 above seem promising, as do the figures showing the number of students who increase or decrease their scores for each treatment, or show no change between tests (Table 2).

Table 2 Number of students scoring higher or lower in Test 2, by treatment

Despite these apparently encouraging figures, a Tukey test derived from a one-way ANOVA conducted on the increase in scores between tests for each treatment (Table 3) shows minimal significant difference between the dictionary items and the corpus items (p = .15). In other words, although the DDL treatment was more effective than the traditional treatment, there is a 15% likelihood that this could be due to chance alone. The Tukey test shows a significant difference between the dictionary items and the control items (p = .013), and between the corpus items and control items (p = .0003).

Table 3 One-way ANOVA and post hoc Tukey for improvement between tests

Table 4 shows the number of correct responses in Tests 1 and 2 (normalised to allow for the different numbers of students), followed by the difference in performance (Test 2 minus Test 1), which is then presented as a percentage change over Test 1. With the DDL treatment, the biggest improvement was recorded for *right/good* and *in the first part*; with the traditional treatment it was for *steal/rob* and *play… sport*. At the other end of the scale, the experiment seems to have produced confusion for some items, with students scoring better in Test 1 than in Test 2 for *home* and *hundred…* whatever the treatment, as well as for *want* with the traditional treatment only. Overall, the DDL treatment seems to have been more effective than traditional techniques for most items, with a comparative advantage of over 100% for *right/good* and *want*. The three exceptions are *play… sport*, *steal/rob* and *say/tell*.

Table 4 DDL vs. traditional treatments for individual items (normed)

Questionnaire Results

The first items on the questionnaire were closed questions, asking the students to compare the two approaches they had just experienced on a 5-point Likert scale. Looking only at the positive results (*agree* or *strongly agree*), 30 students found the dictionary work easy compared to 54 for the corpus work; 31 found the dictionary work useful, compared to 59 for the corpus work. 37 thought the dictionary work would help them avoid certain errors in the future (suggesting they felt they had learned something from the activities), rising to 58 for the corpus work. These encouraging results for DDL are reinforced in the final pair of questions: only 28 students would like to do more dictionary activities in the future, while 51 would like to pursue the DDL work. A general comparison of the "pro-DDL" results (the *b* questions in Table 5: mean = 3.98; *SD* = 0.79) with the "pro-traditional" results (the *a* questions: mean = 3.24; *SD* = 0.97) shows that DDL was significantly better received on the whole ($p = .049$ on a two-tailed paired *t*-test), with the lower standard deviation an interesting reflection of more coherent attitudes.

Table 5 Questionnaire results for 71 students following the experimental session

The final closed question asked the students if they would prefer to explore corpora on their own on computer rather than via the intermediary of paper-based materials. Although they had had no experience of hands-on computer-based DDL, they showed comparatively little enthusiasm to try it: 21 students agreed, and 21 disagreed, while 29 claimed no opinion (Figure 1; mean = 2.94; *SD* = 0.97). Many of the students took the opportunity to explain why: nearly half (25 of the 55 who responded) believed that prepared exercises would get straight to the point and

avoid time-wasting, and teacher guidance would be essential to avoid drawing wrong conclusions from the mass of data. As two students pointed out, they would need to try hands-on DDL first, but two others simply found the possibility "unattractive". Two felt that talking about things was a useful part of the activity rather than just sitting in front of a computer, while eight thought that "doing it themselves" would be more relevant, motivating and lead to more effective learning. More generally, many stressed the importance of context and felt that the numerous samples would help to "visualise" or get a "feel" for the items under study, whether via prepared materials or on their own.

Figure 1 Percentage of students who would like to explore corpora on computer rather than through prepared paper-based exercises

Two open questions allowed the students to say what they felt were the respective advantages of dictionaries and corpora. Dictionaries were considered most useful for new or unknown words (26 students had comments to this effect) and for meanings or definitions (26), or simply for translations (19). Many were interested in usage information (20): for some this was best presented in the form of "rules", while others preferred looking at the examples – although one particularly wanted meanings "independent of any context". Several were interested in checking their intuitions or doubts, though less frequently for formal aspects such as spelling (8) or pronunciation (2).

Corpora, on the other hand, were felt to be most useful for the contexts and "concrete examples" which highlight usage and grammar (58), and to represent "practical English", "frequent usage", the "language of today". Little mention was made of "formulae" or "idiomatic expressions" as such (6), though allusion to context and, more specifically, "words that go

together" reveals a certain sensitivity to this. Most responses seem to refer to corpus use for productive purposes, although some explicit reference was also made to comprehension (13). Similarly, the word "learning" was only mentioned on occasion (4), but is implied in many more responses; thus corpora would help one to "gain an intuitive knowledge of usage", another to develop "my own vocabulary", and a third specifically mentioned that the analysis itself was good "practice". Relatively few made any mention of meanings (7), form (2) or translation (1) compared to the same question for dictionaries.

A final chance to add other comments elicited comparatively few responses, but included some noteworthy remarks. In particular, four students suggested a combination of traditional and DDL work, implying that they see them as filling different purposes or being efficient for different things – one specifically mentioned that corpora helped with "[things] you don't find in a dictionary". Only one was in any way negative, from a student who thought that "a little more context would have been useful" in the concordances. On the other hand, some answers were extremely enthusiastic, including the following:

- Very interesting, an experience to repeat several times with other usage difficulties.
- I'd never heard of corpora. Thank you!
- It's the first time I'd done this type of exercise – but none too soon! Thank you! I'll assimilate things better this time! (Now go and kick out the teachers in high school!!!)

Discussion

The main finding is that scores improved significantly between tests following both traditional and DDL treatments, and that this improvement cannot be attributed to a test effect, as untreated control items did not improve significantly. Although the greatest improvement in Test

2 was for the DDL items, the difference was not highly significant (p = .15), a result which merits further discussion.

The first point to be made is that it is perhaps unreasonable to expect experiments of this type, even given the most favourable conditions, to lead immediately to significantly greater success than traditional methods. This is not intended as an *ad hoc* excuse for disappointing results, but applies similarly to any language teaching/learning approach or methodology: no single experimental study is likely to "prove" the superiority of, say, the Silent Way or Suggestopaedia, or even the Communicative Approach. Indeed, one might be inclined to scepticism if such results were to be obtained. More specifically here, it seems likely that an approach such as DDL, involving enormous amounts of language, is likely to engender considerably wider "incidental" learning than the target items focused on: "What is 'taught' is often not learnt, and learners often 'learn' things which have not been taught at all" (Willis, 2003, p. 1). Ilse (1991) found that "the factual information which each student took away at the end of the [DDL] lesson was less than it would have been had the lesson been traditionally didactically taught" (p. 107), but defends this on the grounds that they will have learned much more besides. This is part of the attraction of DDL, as it might "help [students] to become better language learners outside the classroom" (Johns, 1991b, p. 31) by encouraging noticing and consciousness-raising, leading to greater autonomy and better language learning skills in the long term. O'Sullivan (2007) provides an impressive list of cognitive skills which DDL may be supposed to promote, many of which presumably also apply to paper-based materials: "predicting, observing, noticing, thinking, reasoning, analysing, interpreting, reflecting, exploring, making inferences (inductively or deductively), focusing, guessing, comparing, differentiating, theorising, hypothesising, and verifying" (p. 277). While it is difficult to see how such skills can be measured in an experiment like this, a glimpse of empirical support is given in Cobb (1999): both

concordances and dictionary information brought short-term benefits, though only the group using concordances retained their knowledge – and "if anything increased theirs with time" (p. 354). Another is hinted at by Allan (2006), where the experimental DDL group recorded significant improvement in both target and non-target items:

> The fact that this was true both for words explored through concordances and those not included on the concordance tasks may be seen as an indication that the benefits of the approach extended beyond reflection on individual word learning to strategy use. (p. 44)

The range of scores for the DDL items (*SD* = 2.1 out of 10 in Test 2, compared to 1.7 for each of the other sets of items) suggests also that the experiment conceals considerable variation, and that different learners react to the approach very differently. This no doubt contributes to the lack of significant difference viewing the participants as a whole, and it leads to a number of crucial questions. Firstly, what type of learner takes to DDL most readily, and is it possible to provide some kind of profile? We shall return to this point with regard to level of language ability below. Secondly, is it possible to increase the benefits for those who initially find little gain, and if so how – what kind of training or alternative materials or introduction might be useful? Thirdly, if DDL is found to be simply unsuitable for some types of learners, it leaves open the possibility of having different parts of a class doing different things, with the DDL element based on individual discovery and small group work rather than whole-class teaching. Certainly, there is no question of throwing out all past good practice in language teaching.

Crucially, while the DDL treatment was not found to be significantly more effective than the traditional treatment at the 95% level, it was nonetheless more effective[ii]. Indeed, though the difference may only have had a comparatively low level of statistical significance (85%), it did nonetheless lead to greater improvement. Furthermore, scores improved more under the DDL treatment for seven of the ten language items covered, and more learners increased their scores

for the DDL treatment than for the traditional treatment. All of this suggests that DDL at least has its uses alongside traditional teaching, and perhaps has wider applications than usually assumed – for comparatively low levels of language ability, for regular teachers with little or no background knowledge of corpus linguistics, in normal classrooms outside a research environment.

With appropriate materials for suitable language items, DDL can be exploited immediately and with virtually no training for either teacher or learner. With only a five-minute introduction, the learners seem to have had little difficulty grasping the basic approach – detecting relevant patterns from data rather than being taught more abstract rules. Given that this briefest of introductions did lead to significant improvement, there is every reason to suppose that, with a little more training and experience, DDL would be considerably more effective. Frankenberg-Garcia (2005a) comes to a similar conclusion in her comparison of dictionaries and corpora in a translation course:

> The students in the present study were no better at using dictionaries, even though they had been using them all their lives, than at using corpora, search engines and comparable texts for language research… These findings suggest that the amount of training needed to use unmediated resources does not seem to be any greater than the amount of training needed to use dictionaries and other resources mediated by lexicographers and terminologists. (p. 352)

DDL is usually proposed mainly for more advanced learners, while those in the present study are intermediate and below. The overall scores are fairly low in both tests, suggesting that the items in question are, as intended, problematic for these learners. Cobb (1999) shows that DDL is best suited for depth of knowledge (extending knowledge of known items) rather than breadth (adding new items), and it seems axiomatic that less advanced learners are more likely to be preoccupied with the latter. However, this is only a relative state of affairs and does not preclude DDL on many usage points, as the overall improvement here shows.

One point of comparison can be made between students' level, as measured by the start-of-year TOEIC scores, and the test results for these 62 participants. Pearson's correlation coefficient shows a strong positive correlation with both Test 1 (r = .82) and Test 2 (r = .77; p < .0001 for all correlations here). The slight decrease is no doubt due to the effects of teaching, whether traditional or DDL, as prior knowledge (as reflected by level) was all the participants had to draw on in Test 1. It is also possible to compare levels against the performance on the three types of items in Test 2. Unsurprisingly, the correlation was strongest (r = .76) for the control items: as for Test 1, the students could only draw on their previous knowledge of the language for control items which were not explicitly covered in class. The correlation is lower but still substantial for the traditional items in Test 2 (r = .54); in other words, it can be inferred that the comparatively advanced students gained greater benefit from traditional teaching and dictionary entries. Again, this seems reasonable, in that overall level was mainly the result of traditional teaching in the past, so that those who had been most successful then would also be most successful now using similar techniques. On the other hand, the correlation with level is considerably lower for the DDL items (r = .13), suggesting that all levels benefited as much as each other from this type of information and approach. Again, the implication is that DDL can be of use at lower levels, where learners are capable of detecting patterns at least for some items and applying them to new contexts, contrary to popular assumptions. Support for this can be found in a comparison of the learners' levels and their *reactions* to the different approaches expressed in the questionnaires (as opposed to their performance in Test 2), as these turn out to be negligible (r = .33 for DDL, r = .019 for the traditional treatment, and not significant at the 95% level). In other words, level of language ability was no more a relevant factor in these learners' receptivity to DDL than it was to their actual performance.

The questionnaire results are generally favourable to both approaches, several students explicitly mentioning that they could complement each other rather than be considered as mutually exclusive, though it is clear that DDL was rated significantly higher overall. In particular, these learners found the corpus data *easier* to use (mean score 3.82 out of 5) as well as more *useful* (3.92) than the dictionary information (3.10 and 3.16 respectively).

They further claimed that the DDL treatment would help them to avoid the errors concerned in the future (4.04) better than the traditional treatment (3.39). However, for the 59 participants present at the test and who answered this question, the correlations between these claims and their actual improvement in each case was very small (r = .21 and r = .14 respectively) and not significant at the 95% confidence level. While learners' representations and perceptions are interesting in themselves, this discrepancy highlights once again the dangers inherent in supposing that self-reports represent any more concrete reality.

These learners are keen to continue with corpus data and DDL activities in class (mean score 3.86), but not so keen to do so on computer (2.94, with equivalent numbers agreeing and disagreeing, and 40% having no opinion). It is worth noting that there was a slight although not significant *negative* correlation between the participants' level and their strength of response to this question; in other words, it is not necessarily the more advanced learners who are keenest to get their hands on the computers. As these students need to use sophisticated software for their architectural studies, this reluctance is unlikely to be simply a matter of the technophobia which one might find with surprising numbers of humanities students (Seidlhofer, 2000, p. 208; Bernardini, 2002, p. 169; Mukherjee, 2004, p. 249). Hands-on concordancing may be rejected as time-consuming, boring, laborious, frustrating, as Whistle (1999) found with his students, and it may be that some are happy to escape the computer for once, although many highly computer-literate users are more likely to reject the use of paper-based materials outright (Flowerdew,

2008). The questionnaires do provide a number of clearly-formulated reasons for this, underlining the fact that many learners at this level do not have confidence in their own abilities to reach appropriate conclusions, and feel that the teacher's job in preparing materials is to reduce time-wasting. This corresponds to Granath's (1998) finding that two thirds of students liked the teacher-prepared activities, while less than half were enthusiastic about formulating queries themselves. As our participants completed the questionnaires without the benefit of having experienced hands-on DDL or being informed of its potential, it seems reasonable to propose that they should at least try it before making their minds up. However, it would seem essential to remain sensitive to the local culture, as Seidlhofer (2002) eloquently argues:

> there are no global solutions to motivational problems, no generally valid answers and truths. FL [foreign language] pedagogy, and presumably any pedagogy, has to be local, designed for specific learners and settings. This means that any supposedly general principles have to be interpreted with reference to local settings, or otherwise they are doomed to remain meaningless. (p. 220)

In this case, it entails sensitivity to the learners' relative reluctance to let go of the teacher and take charge of their own learning, to abandon the safety of being taught for the risky business of active discovery. Rather than imposing hands-on DDL on the assumption that "teacher knows best", a gentle lead-in would seem desirable, from paper to computer, from pre-set exercises to more open-ended exploration. The brief experience of paper-based exercises here is clearly not sufficient in itself, but further work might increase the learners' confidence, and they might become more receptive to the possibilities of hands-on corpus exploration after further exposure to the inevitably limited and limiting nature of prepared exercises. However, the case against hands-on DDL should not be overstated on this basis alone: the learners in this study overall were not *against* the use of hands-on concordancing; rather, they were divided – some for, some against, with the largest group undecided. This again raises the questions of whether DDL is

suitable for all types of learner profile, and what can be done to promote its use among the remainder, if indeed there are benefits in their case (cf. Flowerdew, 2008).

Conclusions

The experiment reported here with learners at lower levels of language ability found that, with no prior training, they managed to benefit significantly from prepared, paper-based DDL materials. In particular, they expressed considerable enthusiasm for DDL, and performed better with this approach than they did using dictionary entries and traditional teaching methods, although the difference here was only significant at the 85% level. The present study does not attempt directly to compare the benefits of paper-based materials against hands-on DDL work, a question sorely in need of empirical research as others have mentioned (Aston, 2000, p. 14; Chambers, 2005, p. 121).

Much current DDL research seems to throw learners in at the deep end, requiring them to master the concept of corpora as well as the software and DDL techniques all at once. If paper-based materials have any place at all, it is only as an under-reported introduction before learners move on to more autonomous corpus exploration: clearly learners have to understand the nature of corpus data and analysis before they can explore on their own, and need guidance in their use – autonomy does not come automatically to all (O'Sullivan, 2007). As Sun (2003) points out, "the learning curve… is arduously steep, in that students tend to get confused easily about the concordancer outputs; thus, they need either a stronger degree of teacher involvement, or to learn in a more structured environment" (p. 609); using prepared materials would seem to provide one obvious and "convenient way of introducing concordance-based methods and as preparation for using a full concordancer" (Johns, 1997, p. 113). The findings reported here would seem to support this use, but also go much further, suggesting that the benefits of paper-based materials

are not dependent exclusively upon long-term goals of independent concordancing, but also have short-term pay-offs: students are learning English efficiently while they are learning to use corpora.

We have argued that eliminating the computer from the equation, far from fatally undermining the conceptual basis of DDL, can in fact make the learners' task considerably easier. In particular, it alleviates a number of methodological difficulties by, among other things, reducing the amount of data and limiting the range of possible answers (Thompson, 2006) – not to mention technical, logistical and financial obstacles for the teacher. Learners such as ours may also initially feel paper-based resources are more relevant or efficient and, as Whistle (1999) puts it, simply have difficulty understanding "why the concordances could not be prepared in advance and handed out in class" (p. 77). Paper-based materials also provide something tangible for students to take away and consult at a later date.

The received wisdom seems generally to be that DDL is less suitable at lower levels of language ability. However, many of the most frequent objections concern the hands-on exploitation of corpora, and do not necessarily extend to the use of paper-based DDL materials. Certainly, the positive results reported here suggest that, at least when materials are prepared in advance and presented on paper, DDL can be appropriate for lower-level learners, thus supporting previous research of our own (Boulton, 2007; Boulton, 2008c; Boulton, 2009) and a small number of other studies (e.g., Yoon & Hirvela, 2004).

Further work is clearly needed, but results such as these might help to convince a wider public that surprisingly little investment is required for rapid and substantial returns, and is within the capabilities of regular teachers with little expert knowledge of corpus linguistics. One major problem remains, namely that DDL materials are extremely time-consuming to prepare: each of the items here required half a day's work; Johns (1991a) similarly reports spending eight hours

on a single handout. Clearly such investment cannot be expected in normal teaching contexts, but if it is possible to produce "materials that can be used more than once, for items that a teacher knows from experience cause problems in many students' reading and writing, then the time and effort needed to make those materials will be richly rewarded" (Warren, 1998, p. 214). For the moment, though, downloadable worksheets remain scarce, are not necessarily transferable to new contexts, and are dependent on researchers' goodwill; Johns' (2000) kibbitzers are frequently cited in this area. Furthermore, there are virtually no published materials available (Boulton, 2008b): of the eight empirical studies using paper-based materials reported in Boulton (2008a), all but one relied on home-produced materials. Although corpora have been used to inform coursebooks, this generally concerns only the content; and even here their presence is often invisible – deliberately so, with McCarthy (2004) repeatedly remarking for the *Touchstone* series (e.g., McCarthy, McCarten & Sandiford, 2005) that the authors attempt to conform to traditional presentations as far as possible. Only two truly DDL books are currently available (Tribble & Jones, 1997; Thurstun & Candlin, 1997), neither of which has apparently been tested empirically beyond the authors' own pilot studies, although Hadley (2002) has done this with the set exercises from a *COBUILD Concordance Sampler* (Goodale, 1995). Greater research interest producing positive results might inspire publishers to produce DDL books or paying Web sites; to integrate DDL activities into more general materials; or to include corpora and interactive tools on Web sites or DVD-ROMs which accompany their publications.

References

Allan, R. (2006). Data-driven learning and vocabulary: Investigating the use of concordances with advanced learners of English. *Centre for Language and Communication Studies, Occasional Paper, 66*. Dublin: Trinity College Dublin.

Aston, G. (1997). Involving learners in developing learning methods: Exploiting text corpora in self-access. In P. Benson & P. Voller (Eds.), *Autonomy and independence in language learning* (pp. 204-263). London: Longman.

Aston, G. (1998, May). Learning English with the British National Corpus. Paper given at 6th Jornada de Corpus, Barcelona: UPF. Retrieved August, 2008, from http://sslmit.unibo.it/~guy/barc.htm

Aston, G. (2000). Corpora and language teaching. In L. Burnard & T. McEnery (Eds.), *Rethinking language pedagogy from a corpus perspective* (pp. 7-17). Frankfurt: Peter Lang.

Bernardini, S. (2001). Spoilt for choice: A learner explores general language corpora. In G. Aston (Ed.), *Learning with corpora* (pp. 220-249). Houston: Athelstan.

Bernardini, S. (2002). Exploring new directions for discovery learning. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 165-182). Amsterdam: Rodopi.

Boulton, A. (2007). DDL is in the details… and in the big themes. [Electronic version]. In M. Davies, P. Rayson, S. Hunston, & P. Danielsson (Eds.), *Proceedings of corpus linguistics 2007*. Retrieved June, 2008, from http://www.corpus.bham.ac.uk/corplingproceedings07/

Boulton, A. (2008a, March). Evaluating corpus use in language learning: State of play and future directions. Paper presented at the *American Association of Corpus Linguistics*, Brigham Young University, Provo, UT.

Boulton, A. (2008b, September). 'Off-the-peg' materials for data-driven learning. Paper presented at the *New Trends in Corpus Linguistics for Language Teaching and Translation Studies: In Honour of John Sinclair*, University of Granada & University Jaume I, Granada, Spain.

Boulton, A. (2008c). Looking for empirical evidence of data-driven learning at lower levels. In B. Lewandowska-Tomaszczyk (Ed.), *Practical applications of language and computers: PALC07*. Frankfurt: Peter Lang.

Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *ReCALL, 21*(1).

Bowker, L. (1999). Exploring the potential of corpora for raising language awareness in student translators. *Language Awareness, 8*(3-4), 160-173. Retrieved August, 2008, from http://www.multilingual-matters.net/la/008/0160/la0080160.pdf

Breyer, Y. (2006). My Concordancer: Tailor-made software for language learners and teachers. In S. Braun, K. Kohn & J. Mukherjee (Eds.), *Corpus technology and language pedagogy: New resources, new tools, new methods* (pp. 157-176). Frankfurt: Peter Lang.

Brown, D. (2007). Language learner motivation and the role of choice in ESP listening engagement. *ASp, 51-52*, 159-187.

Chambers, A. (2005). Integrating corpus consultation in language studies. *Language Learning & Technology, 9*(2), 111-125.

Chambers, A. (2007a). Popularising corpus consultation by language learners and teachers. In E. Hidalgo, L. Quereda & J. Santana (Eds.), *Corpora in the foreign language classroom* (pp. 3-16). Amsterdam: Rodopi.

Chambers, A. (2007b). Language learning as discourse analysis: Implications for the LSP learning environment. *ASp, 51-52*, 35-51.

Chambers, A., & Kelly, V. (2004). Semi-specialized corpora of written French as a resource in language teaching and learning. *Teanga, 21*, 114-133.

Chambers, A., & O'Sullivan, I. (2004). Corpus consultation and advanced learners' writing skills in French. *ReCALL, 16*(1), 158-172.

Ciezielska-Ciupek, M. (2001). Teaching with the Internet and corpus materials: Preparation of ELT materials using the Internet and corpus resources. In B. Lewandowska-Tomaszczyk (Ed.), *PALC 2001: Practical applications in language corpora* (pp. 521-531) [Special issue]. *Lodz Studies in Language, 7*. Frankfurt: Peter Lang.

Cobb, T. (1999). Breadth and depth of lexical acquisition with hands-on concordancing. *CALL, 12/4*, 345-360.

Cobb, T. (1997). *From concord to lexicon: Development and test of a corpus-based lexical tutor.* Ph.D dissertation, Concordia University, Montreal. Retrieved August, 2008, from http://www.er.uqam.ca/nobel/r21270/webthesis/Thesis0.html

Cobb, T. (2003). Do corpus-based electronic dictionaries replace concordancers? In B. Morrison, G. Green & G. Motteram (Eds.). *Directions in CALL: Experience, experiments, evaluation* (pp. 179-206). Hong Kong: Polytechnic University. Retrieved August, 2008, from http://www.er.uqam.ca/nobel/r21270/cv/replace_conc.htm

Collins COBUILD English dictionary for advanced learners. (4th ed.). (2003). [Electronic version]. London: HarperCollins. Retrieved January, 2008, from http://dictionary.reverso.net/english-cobuild/

Collins English French electronic dictionary. (2005). [Electronic version]. London: HarperCollins. Retrieved January, 2008, from http://dictionary.reverso.net/french-english/ and http://dictionary.reverso.net/english-french/

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Oxford: Oxford University Press.

Cresswell, A. (2007). Getting to 'know' connectors? Evaluating data-driven learning in a writing skills course. In E. Hidalgo, L. Quereda & J. Santana (Eds.), *Corpora in the foreign language classroom* (pp. 267-287). Amsterdam: Rodopi.

Davies, M. (n.d.). *BYU-BNC: British National Corpus*. Retrieved January, 2008, from http://corpus.byu.edu/bnc/x.asp

Estling Vannestål, M., & Lindquist, H. (2007). Learning English grammar with a corpus: Experimenting with concordancing in a university grammar course. *ReCALL, 19*(3), 329-350.

Farr, F. (2008). Evaluating the use of corpus-based instruction in a language teacher education context: Perspectives from the users. *Language Awareness, 17*(1), 25-43.

Flowerdew, L. (2008). Pedagogic value of corpora: A critical evaluation. In A. Frankenburg-Garcia (Ed.), *Proceedings of the 8th teaching and language corpora conference* (pp. 115-119). Lisbon: Associação de Estudos e de Investigação Científica do ISLA-Lisboa.

Frankenberg-Garcia, A. (2005a). A peek into what today's language learners as researchers actually do. *International Journal of Lexicography, 18*(3), 335-355.

Frankenberg-Garcia, A. (2005b). Pedagogical uses of monolingual and parallel concordances. *ELT Journal, 59*(3), 189-198.

Gabrielatos, C. 2005. Corpora and language teaching: Just a fling or wedding bells? *Teaching English as a Second Language – Electronic Journal, 8*(4), 1-35. Retrieved August, 2008, from http://tesl-ej.org/ej32/a1.htm

Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for writing errors? *System, 32*(3), 301-319.

Gavioli, L. (2005). *Exploring corpora for ESP learning*. Amsterdam: John Benjamins.

Goodale, M. (1995). *Concordance samplers 2: Phrasal verbs*. London: HarperCollins.

Granath, S. (1998). Using corpora in teaching English syntax to EFL students at the university level. In L. Burnard (Ed.), *Proceedings of teaching and language corpora: TaLC98* (pp. 87-92). Oxford: Keble College.

Granger, S. (2003). The International Corpus of Learner English: A new resource for foreign language learning and teaching and second language acquisition research. *TESOL Quarterly, 37*(3), 538-545.
Hadley, G. (2002). Sensing the winds of change: An introduction to data-driven learning. *RELC Journal, 33*(2), 99-124.
Higgins, J., & Johns, T. (1984). *Computers in language learning* (pp. 88-93). London: Collins.
Ilse, W. (1991). Concordancing in vocational training. In T. Johns & P. King (Eds.), *Classroom concordancing* (pp. 103-113) [Special issue]. *English Language Research Journal, 4*.
Johns, T. (1991a). Should you be persuaded: Two examples of data-driven learning. In T. Johns & P. King (Eds.), *Classroom concordancing* (pp. 1-16) [Special issue]. *English Language Research Journal, 4*.
Johns, T. (1991b). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. In T. Johns & P. King (Eds.), *Classroom concordancing* (pp. 27-45) [Special issue]. *English Language Research Journal, 4*.
Johns, T. (1997). Contexts: The background, development and trialling of a concordance-based CALL program. In A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles (Eds.), *Teaching and language corpora* (pp. 100-115). Harlow: Addison Wesley Longman.
Johns, T. (2000). Kibbitzing. Retrieved August, 2008, from http://www.eisu2.bham.ac.uk/johnstf/timeap3.htm#revision
Johns, T. (2002). Data-driven learning: The perpetual challenge. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 107-117). Amsterdam: Rodopi.
Koosha, M., & Jafarpour, A. (2006). Data-driven learning and teaching collocation of prepositions: The case of Iranian EFL adult learners. *Asian EFL Journal Quarterly, 8*(4), 192-209.
Lamy, M-N., & Klarskov Mortensen, J. (2007). Using concordance programs in the modern foreign languages classroom. module 2.4. In G. Davies (Ed.), *Information and communications technology for language teachers (ICT4LT)*. Slough, Thames Valley University. Retrieved August, 2008, from http://www.ict4lt.org/en/en_mod2-4.htm
Mauranen, A. (2004). Speech corpora in the classroom. In G. Aston, S. Bernardini & D. Stewart (Eds.), *Corpora and language learners* (pp. 195-211) Amsterdam: John Benjamins.
McCarthy, M. (2004). *Touchstone: From corpus to coursebook*. Cambridge: Cambridge University Press. Retrieved August, 2008, from http://www.cambridge.org/us/esl/Touchstone/teacher/images/pdf/CorpusBookletfinal.pdf
McCarthy, M., McCarten, J., & Sandiford, H. (2005). *Touchstone 1: Student's book*. Cambridge: Cambridge University Press.
Mukherjee, J. (2004). Bridging the gap between applied corpus linguistics and the reality of English language teaching in Germany. In U. Connor & T. Upton (Eds.), *Applied corpus linguistics: A multidimensional perspective* (pp. 239-250). Amsterdam: Rodopi.
Mukherjee, J. (2006). Corpus linguistics and language pedagogy: The state of the art – and beyond. In S. Braun, K. Kohn & J. Mukherjee (Eds.), *Corpus technology and language pedagogy: New resources, new tools, new methods* (pp. 5-24) [Special issue]. *English Corpus Linguistics, 3*. Frankfurt: Peter Lang.
Osborne, J. (2004). Top-down and bottom-up approaches to corpora in language teaching. In U. Connor & T. Upton (Eds.), *Applied corpus linguistics: A multidimensional perspective* (pp. 251-265). Amsterdam: Rodopi,.
O'Sullivan, I. (2007). Enhancing a process-oriented approach to literacy and language learning: The role of corpus consultation literacy. *ReCALL, 19*(3), 269-286.

Oxford advanced learner's dictionary (7th ed.). (2007). [Electronic version]. Oxford: Oxford University Press. Retrieved January, 2008, from http://www.oup.com/elt/catalogue/teachersites/oald7/lookup?cc=global

Scott, M., & Tribble, C. (2006). *Textual patterns: Key words and corpus analysis in language education*. Amsterdam: John Benjamins.

Seidlhofer, B. (2000). Operationalizing intertextuality: Using learner corpora for learning. In L. Burnard & T. McEnery (Eds.), *Rethinking language pedagogy from a corpus perspective* (pp. 207-223). Frankfurt: Peter Lang.

Seidlhofer, B. (2002). Pedagogy and local learner corpora: Working with learner-driven data. In S. Granger, J. Hung & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 213-234). Amsterdam: John Benjamins.

Stevens, V. (1991). Concordance-based vocabulary exercises: A viable alternative to gap-filling. In T. Johns & P. King (Eds.), *Classroom concordancing* (pp. 47-61) [Special issue]. *English Language Research Journal, 4*.

Sun, Y-C. (2003). Learning process, strategies and Web-based concordancers: A case-study. *British Journal of Educational Technology, 34*(5), 601-613.

Tayfoor, S. (2004). *Common mistakes at first certificate… and how to avoid them*. Cambridge: Cambridge University Press.

Thompson, P. (2006). Assessing the contribution of corpora to EAP practice. In Z. Kantaridou, I. Papadopoulou & I. Mahili (Eds.), *Motivation in learning language for specific and academic purposes*. Macedonia: University of Macedonia. Retrieved August, 2008, from http://www.reading.ac.uk/internal/appling/thompson_macedonia.pdf

Thurstun, J., & Candlin, C. (1997). *Exploring academic English: A workbook for student essay writing*. Sydney: CELTR.

Thurstun, J., & Candlin, C. (1998). Concordancing and the teaching of the vocabulary of academic English. *English for Specific Purposes, 17*(3), 267-280.

Todd, R. (2001). Induction from self-selected concordances and self-correction. *System, 29*(1), 91-102.

Tribble, C., & Jones, G. (1997). *Concordances in the classroom*. (2nd ed.) Houston: Athelstan.

Turnbull, J., & Burston, J. (1998). Towards independent concordance work for students: Lessons from a case study. *ON-CALL, 12*(2), 10-21.

Warren, E. (1998). The use of concordance lines in self-access grammar materials. In L. Burnard (Ed.), *Proceedings of teaching and language corpora: TaLC98* (pp. 213-221). Oxford: Keble College.

Whistle, J. (1999). Concordancing with students using an 'off-the-Web' corpus. *ReCALL, 11*(2), 74-80.

Willis, D. (2003). *Rules, patterns and words*. Cambridge: Cambridge University Press.

Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2. *Journal of Second Language Writing, 13*(4), 257-283.

APPENDIX A. Essay topics for collection of language items

a) Carbon-neutral: housing for the future.
b) Living in the country: a better quality of life.
c) Housing estates: renovate, or knock them down and start again?
d) Out-of-town shopping centres are killing inner cities.

APPENDIX B. Language items selected, plus sample errors

1. dozen, hundred, thousand, million, billion
   10000 of habitants; ten millions
2. want
   the governement want knock them down; do not want that they come
3. person, people
   to meet new persons; persons who live in the city
4. in the first part
   in first time; in a second time
5. right, good
   the good answer; what is the good solution
6. depend
   it depends of the people; it's depending of; it's depend of you
7. home
   the moment where I arrive at home; they come at home
8. play, practice, do, go (+sport / activity)
   to practise activities like bike; you can practice many sports
9. say, tell
   many people tell that you can't; my teacher say me that
10. steal, rob
    people don't steal each other; someone robbs my bicycle
11. agree
    every doctor are agree to say that; I am not agreed
12. allow
    that allows to get fresh ideas; proximity allows to shop
13. lose, loser, loss, loose
    you loose your identity; we can lost time in the car
14. near, far, close
    near from the place where we are living; too far of the civilisation; closer of your job
15. only, alone, lone, lonely
    the alone sound which you hear; in isolation you feel alone

APPENDIX C. Sample test items

1. The company employs no more than _____ people.
   A) two hundreds
   B) two hundred
   C) two hundred of
   D) two hundreds of

2. They don't want _____ in the restaurant.
   A) that people smoke
   B) that people smoking
   C) people to smoke
   D) people smoke

3. I had a letter from _____ who used to live next door.
   A) the people
   B) the persons
   C) the peoples
   D) a people

4. We will discuss a number of problems, but _____ we begin with definitions.
   A) in a first part
   B) in the first part
   C) in a first time
   D) in the first time

APPENDIX D. Materials[iii]

RIGHT (group A: DDL).

Introduction:
The adjective *right* has several meanings in English.

a) How would you translate *la bonne réponse* into English?
b) How would you translate *the right side* into French?

Main materials:
c) Below are some of the most common nouns which follow *right*, 1) e.g. *right hand* or 2) *right time*.
For each one, decide how you would probably translate it.

1.	hand	_____	7.	place	_____
2.	time	_____	8.	angles	_____
3.	side	_____	9.	wing	_____
4.	thing	_____	10.	arm	_____
5.	way	_____	11.	hemisphere	_____
6.	direction	_____	12.	leg	_____

d) How many translations did you have for *the right side*?
Why do you think we often say *the right <u>hand</u> side*?

e) Look at the words to the left of *right answer* and *good answer* in the concordance lines below.
What do you think is the difference in meaning and use?
Would you translate them differently in French?

1. way. Corporate strategists point out that there is no single **right answer**. The correct strategy will be industry specific. This is what the
2. are made. These answers should not be viewed as the " **right answer**" and we would suggest that wherever possible the students are
3. , and avoid making them feel that they have to search for the **right answer**, hidden somewhere in the teacher's head. Open questions
4. in a number of respects although he may have reached the **right answer** by the wrong route. I do not agree. His conclusion that there
5. a set of facts and techniques --; in which questions have one **right answer** and prescribed methods of solution. Hand-in-hand with this
6. much children have taken in. In drama there is rarely a single **right answer**, and it's often more appropriate to phrase questions so that it is
7. the right question is usually more difficult than to find the **right answer**. The questions which are tackled at Advanced level reflect the
8. argument and "yes" more quickly when "yes" is the **right answer**," he said. Of the 40,000 asylum applications made last year,
9. "Closer study shows that there is no simple solution, no one **right answer**, no single "management style" that delivers better results." He
10. answer. However, I think we are more likely to find the **right answer** if we ask the right question. We should not ask "Why are
11.
12. don't have any meetings. Well yeah, that's a very **good answer** to that, yeah, but when, how would you do research? Yeah
13. there was next. To see what animals came next. Yes. **Good answer**. And what about you. You'd live in? You'd live in
14. that answer your question? I mean, it's not a very **good answer**, because frankly we don't know, th the full reasons for this,
15. it that a causal circumstance makes an effect happen? A **good answer** is that we regard the causal circumstance as leaving no room
16. I've got a right." Nutty could not think of a **good answer** and nor could Mr Sylvester, so Nails was allowed to come.
17. : the Godfather's name is that. Will-power supreme" "A **good answer** to a question that has no simple answer. A reverend answer. A
18. . "Hanging about waiting to die is not my idea of a **good answer**!" "Visual sighting of parasite at grid mark four by five,"
19. more grass," said the aunt. It was not a very **good answer**, and the boy knew it. "But there is lots of grass in
20. Susan smiled. "Well …" But she couldn't find a **good answer** to Karen's question. A week later, when the workers had to
21. so angry with him?" Lydia was evasive, not having a **good answer** ready. "Mmm," she said shiftily, "I just thought it

Look at questions (a) and (b) again. How would you answer them now?

RIGHT (group B: Traditional).

Introduction:
The adjective *right* has several meanings in English.

a) How would you translate *la bonne réponse* into English?
b) How would you translate *the right side* into French?

Main materials:
c) What common nouns can you think of which frequently follow *right*? For each one, decide which translation is the best.

right	hand	*main droite*	right		
right	time		right		
right			right		
right			right		

d) How would you translate *the right side*? Why do you think we often say *the right <u>hand</u> side*?

e) Look at the dictionary entries below. What do you think is the difference in meaning and use between *right answer* and *good answer*? Would you translate them differently in French?

right
adj
(not left) **droit**(e)
(=correct) [answer, road, direction, address, number] **bon(bonne)**
(=accurate) [time] **juste**
(=most suitable) [moment, choice] **bon(bonne)**
(=morally good) **bien** *inv*
(in normal or satisfactory condition) I don't feel right today. Je ne me sens pas bien aujourd'hui.
(=socially acceptable) the right people les gens bien placés
(British) * *(=total)* **sacré**(e) *

right adj
1 equitable, ethical, fair, good, honest, honourable, just, lawful, moral, proper, righteous, true, upright, virtuous
2 accurate, admissible, authentic, correct, exact, factual, genuine, on the money (U.S.) precise, satisfactory, sound, spot-on (Brit. informal) true, unerring, valid, veracious
3 advantageous, appropriate, becoming, comme il faut, convenient, deserved, desirable, done, due, favourable, fit, fitting, ideal, opportune, proper, propitious, rightful, seemly, suitable
4 all there (informal) balanced, compos mentis, fine, fit, healthy, in good health, in the pink, lucid, normal, rational, reasonable, sane, sound, unimpaired, up to par, well
5 conservative, reactionary, Tory
6 absolute, complete, out-and-out, outright, pure, real, thorough, thoroughgoing, utter **adv**
7 accurately, aright, correctly, exactly, factually, genuinely, precisely, truly
8 appropriately, aptly, befittingly, fittingly, properly, satisfactorily, suitably
9 directly, immediately, instantly, promptly, quickly, straight, straightaway, without delay
10 bang, exactly, precisely, slap-bang (informal) squarely
11 absolutely, all the way, altogether, completely, entirely, perfectly, quite, thoroughly, totally, utterly, wholly
12 ethically, fairly, honestly, honourably, justly, morally, properly, righteously, virtuously
13 advantageously, beneficially, favourably, for the better, fortunately, to advantage, well **n**
14 authority, business, claim, due, freedom, interest, liberty, licence, permission, power, prerogative, privilege, title
15 equity, good, goodness, honour, integrity, justice, lawfulness, legality, morality, propriety, reason, rectitude, righteousness, truth, uprightness, virtue
16 by rights equitably
17 to rights arranged **vb**
18 compensate for, correct, fix, put right, rectify, redress, repair, settle, set upright, sort out, straighten, vindicate **adj**
1 bad, dishonest, immoral, improper, indecent, unethical, unfair, unjust, wrong
2 counterfeit, erroneous, fake, false, fraudulent, illegal, illicit, inaccurate, incorrect, inexact, invalid, mistaken, questionable, uncertain, unlawful, untruthful, wrong
3 disadvantageous, inappropriate, inconvenient, undesirable, unfitting, unseemly, unsuitable, wrong
4 abnormal, unsound
5 left, leftist, left-wing, liberal, radical, right-on (informal) socialist **adv**
7 inaccurately, incorrectly
8 improperly
9 incompletely, indirectly, slowly
13 badly, poorly, unfavourably **n**
15 badness, dishonour, evil, immorality, impropriety **vb**
18 make crooked, topple

Look at questions (a) and (b) again. How would you answer them now?

WANT (group A: Traditional).

Introduction:
Verbs often don't behave the same way in sentences in different languages.

a) Is *want* used syntactically the same way as *vouloir*?
b) How would you translate: *Qu'est-ce que tu veux que je fasse?*

Main materials:
c) What are the main patterns of usage for *want* in the dictionary entries below?

vouloir
vt
(=exiger, désirer) **to want**
[destin, circonstances] le hasard a voulu que ... as fate would have it, ...
(souhait vis-à-vis de qn) vouloir qch à qn to wish sth for sb
(=consentir) je veux bien (bonne volonté) I'll be happy to, (concession) fair enough, (en acceptant une proposition) I'd love to
en vouloir à qn (rancune) to be angry at sb
en vouloir à qch (convoitise) to be after sth
(formules de politesse) veuillez patienter please wait
(autres locutions) oui, si on veut *(=en quelque sorte)* yes, if you like
nm
le bon vouloir de qn sb's goodwill

want
vt
(=desire, wish for) **vouloir**
(=need) **avoir besoin de**
if you want si vous voulez
what do you want? Qu'est-ce que vous voulez?
(=should) sb wants to do sth * qn devrait faire qch
n
(=poverty) **besoin** *m*
(=lack) **manque** *m* **wants**
npl
(=requirements) **besoins** *mpl*

want (**wants** *plural & 3rd person present*) (**wanting** *present participle*) (**wanted** *past tense & past participle*)
1 **verb** If you **want** something, you feel a desire or a need for it. I want a drink... *V n*
2 **verb** You can say that you **want** *to* say something to indicate that you are about to say it. Look, I wanted to apologize for today. I think I was a little hard on you. *V to-inf*
3 **verb** You use **want** in questions as a way of making an offer or inviting someone to do something. Do you want another cup of coffee?... *V n*
4 **verb** If you say to someone that you **want** something, or ask them if they **want** *to* do it, you are firmly telling them what you want or what you want them to do. I want an explanation from you, Jeremy... *V n*
5 **verb** If you say that something **wants** doing, you think that it needs to be done. (mainly BRIT) INFORMAL *(=need)* Her hair wants cutting. *V -ing*
6 **verb** If you tell someone that they **want** *to* do a particular thing, you are advising them to do it. INFORMAL *(=ought)* You want to be very careful not to have a man like Crevecoeur for an enemy... *V to-inf*
7 **verb** If someone **is wanted** by the police, the police are searching for them because they are thought to have committed a crime. He was wanted for the murder of a magistrate. *be V-ed for n*
♦ **wanted adj** He is one of the most wanted criminals in Europe.
8 **verb** If you **want** someone, you have a great desire to have sex with them. Come on, darling. I want you. *V n*
9 **verb** If a child **is wanted**, its mother or another person loves it and is willing to look after it. Children should be wanted and planned... *be V-ed*
10 **n-sing** A **want of** something is a lack of it. FORMAL *(=lack)* ...a want of manners and charm...
11 **n-plural** Your **wants** are the things that you want. Supermarkets often claim that they are responding to the wants of consumers by providing packaged foods.
12 If you do something **for want of** something else, you do it because the other thing is not available or not possible.
♦ **for want of phrase** Many of them had gone into teaching for want of anything better to do... **want out phrasal verb** If you **want out**, you no longer want to be involved in a plan, project, or situation that you are part of. INFORMAL

d) Is it possible to say *want that*?
If so, when? If not, why not?

e) *Want* is a regular verb: *want, wants, wanting, wanted*.
Which one of these is rare, and why?

Look at questions (a) and (b) again. How would you answer them now?

WANT (group B: DDL).

Introduction:
Verbs often don't behave the same way in sentences in different languages.

a) Is *want* used syntactically the same way as *vouloir*?
b) How would you translate: *Qu'est-ce que tu veux que je fasse?*

Main materials:
c) The following concordances all come from the same novel.
Can you identify the pattern of usage in each group?

1. Harsnet (typed Goldberg), does she entice him there or does he **want** to come? Does she entice him or does he appear by chance? Ovid
2. wrote Harsnet, it is this which excites me, this which makes me **want** to go on. As if my whole life, he wrote, had been spent
3. , it had all become too complicated, too extreme, I did not **want** to know any of it until it was all over, until I had made
4. What about the right hand side? I said. That's what I **want** to know. But he couldn't help. Only kept repeating it was quite
5. , he wrote, this is a message from the past. I just **want** to tell you. Goldberg, pushing aside pad and pen, drew the little
6. and less and less well of course, but the truth was I did not **want** to wake up. How the days, instead of each being distinct from each
7. story, you can have one. If not, not. If you **want** to walk round it, you can do so. If you want to get
8. you with it, he said. My best picture. You didn't **want** to win their lousy prize, did you? I said. You did too

9. a room. And so on. Otherwise perhaps too dangerous. I don't **want** anyone to get lost in it, I said. To be unable to get
10. now? I want it to be as though I had never been. I **want** it to be as though I had never taken that turning. But that can not
11. difficult to know if idea is really valid. Will really yield what I **want** it to. You can dream and dream but only what happens in this room
12. Goldberg too: Help me, I'm so unhappy. What do they **want** me to do? Blow their noses for them? And why me? Why
13. ? Go where? wrote Harsnet (typed Goldberg). Where did they **want** me to go? Mushrooms grow in the dark, I said to Hilda.
14. to stifle doubts by crushing you with sheer bulk, he wrote. I **want** my doubts to play and dance. And Goldberg, drawing his pad
15. glass to be seen, I want it placed in a morgue and I **want** people to come in and see it, pay money and come in and see
16. to be done. I myself am guilty, he wrote, in that I **want** the glass to be seen, I want it placed in a morgue and I

d) There are no examples of *want that* in our novel, but below is a sample of the 140 occurrences in BNC fiction.
Would you translate *want that* back as *vouloir que* here?

17. whole face drooped and she blushed with chagrin. "I think you **want that**, living under their feet, and their scolding and groaning. Why
18. a goddam. I want to be your lover. And I believe you **want that** as well. I know you're scared of it, fuck it, you
19. to Matlock. But I won't make the phone call. If you **want that** done, you'll have to come down and do it yourself. Tomorrow
20. didn't want Bella to say anything more about herself. She didn't **want that** door to open any further. But Bella did not notice the tone:
21. if our dad'adn't pulled'er off. You sure you don't **want that** other sausage roll, love?" "I ain't got the strength
22. round the mean, narrow, metal window. "Where do you **want that** put?" McLeish advanced on the cabinet and shifted it
23. too far, they'd have to move lodgings again. Dot didn't **want that**. "Why can't he stay in this hospital? He likes it
24. ? Would I become a burden to everyone again? I really didn't **want that**. Above all else I didn't want that. They had become my
25. he'd come and take me to Combe Court, and I don't **want that**. And anyway, I ain't got my yellow card on me.
26. , still too middle-class, too much her father's daughter, to **want that**. This was leafy north Oxford, this was the pace of security.

e) There are no examples of *want<u>ing</u>* in our novel either.
How many examples can you find of the present or past continuous in the following lines
from BNC fiction?

27. Building Larkin feared. He was right, I said, suddenly cold and **wanting** home; cold later, too, in bed, listening to wind and rain
28. and technique --; not seeing the only point of being there was **wanting** each other more than the rest for that short time --; lying till sex
29. The reply is a blur of distress the only words I catch (not **wanting** to) but --; doctor --; I don't know what to do with myself
30. worries the night sky where the hillside consumes itself. Those **wanting** compensation tie a burning brand to a trapped bird's foot so
31. means of transport, air-routes and sea-routes, and found them **wanting**. And every road he chose led back to Rome, a Rome not quite
32. prow. "Milord, you are, yes, yes, yes, **wanting** hotel on Capri? Yourself and the pretty lady, yes? One
33. only things." "Yeah --; too right, love! Are you **wanting** to take any of it now?" "No, I need to have
34. night, Christ, that is, and woe betide you if he finds you **wanting**.) Thou shalt do everything that the church says thou shalt not.
35. raised my eyes to the sky --; where God was --; beseeching Him, **wanting** Him to see my fear and my contrition. But I couldn't see the
36. thickly around them. Afterwards they stood up slowly, not **wanting** to part from such intimacy, but beginning to feel chilly now the

Look at questions (a) and (b) again. How would you answer them now?

APPENDIX E. Questionnaire

First name: _____ Surname: _____ Group: _____

Pour chaque question, entourez le chiffre qui correspond le mieux à votre perception, ou bien répondez avec vos propres mots.

	Pas du tout d'accord	Plutôt pas d'accord	Je ne sais pas	Plutôt d'accord	Tout à fait d'accord
1a) Pour moi, le travail avec le dictionnaire a été facile.	1	2	3	4	5
1b) Pour moi, le travail avec le corpus a été facile.	1	2	3	4	5
2a) Pour moi, le travail avec le dictionnaire a été utile.	1	2	3	4	5
2b) Pour moi, le travail avec le corpus a été utile.	1	2	3	4	5
3a) Grâce aux dictionnaires, je pense que j'éviterai certaines erreurs à l'avenir.	1	2	3	4	5
3b) Grâce aux corpus, je pense que j'éviterai certaines erreurs à l'avenir.	1	2	3	4	5
4a) J'aimerais faire d'autres activités de dictionnaire en classe.	1	2	3	4	5
4b) J'aimerais faire d'autres activités de corpus en classe.	1	2	3	4	5
5a) J'aimerais explorer un corpus moi-même sur ordinateur plutôt qu'à travers des exercices préparés.	1	2	3	4	5

5b) Pourquoi ?

…….

…….

6a) Selon moi, un dictionnaire est surtout utile pour :

…….

…….

6b) Selon moi, un corpus est surtout utile pour :

…….

…….

7) D'autres commentaires :

…….

…….

…….

…….

Merci pour votre participation.

Table 1 Descriptive data for mean test results

	DDL /10	traditional /10	control /10	total /30
TEST 1	4.85	4.65	5.06	14.56
TEST 2	6.39	5.68	5.32	17.39
difference	+1.53	+1.03	+0.26	+2.82
change	+31.56%	+22.22%	+5.10%	+19.38%

Table 2 Number of students scoring higher or lower in Test 2, by treatment

	DDL	traditional	control
higher	43	38	26
no change	12	12	17
lower	7	12	19

Table 3 One-way ANOVA and post hoc Tukey for improvement between tests

items	*n*	mean	*SE*	pooled *SE*	*SD*
DDL	62	1.5	0.27	0.24	2.1
traditional	62	1.0	0.22	0.24	1.7
control	62	0.3	0.21	0.24	1.7

source of variation	sum squares	*DF*	mean square	*F* statistic	*p*
groups	51.1	2	25.6	7.38	0.0008
residual	633.2	183	3.5		
total	684.3	185			

Tukey contrast	difference	95% *CI*	contrast		
DDL vs. traditional	0.5	-0.3	to 1.3		(*p* = 0.1539)
DDL vs. control	1.3	0.5	to 2.1	(significant)	(*p* = 0.0003)
traditional vs. control	0.8	0.0	to 1.6		(*p* = 0.0134)

Table 4 DDL vs. traditional treatments for individual items (normed)

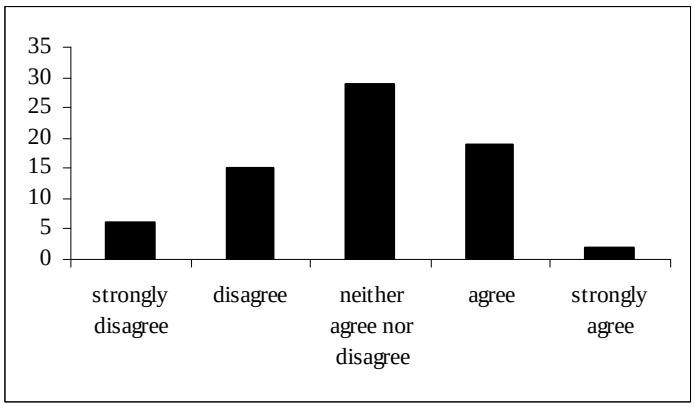
	DDL				traditional				comparative
	T1	T2	diff (T2-T1)	change (diff/T1)	T1	T2	diff	change (diff/T1)	DDL advantage
1. right. good	0.60	1.44	0.84	140.0%	1.08	1.40	0.32	29.6%	110.37
2. want	0.96	1.64	0.68	70.8%	1.32	0.80	-0.52	-39.4%	110.23

3.	in the first part	0.52	1.12	0.60	115.4%	0.88	1.44	0.56	63.6%	51.75
4.	person. people	0.60	0.96	0.36	60.0%	1.12	1.36	0.24	21.4%	38.57
5.	depend	0.89	1.19	0.30	33.3%	0.68	0.68	0.00	0.00%	33.33
6.	hundred…	1.20	0.88	-0.32	-26.7%	1.64	1.12	-0.52	-31.7%	5.04
7.	home	1.65	1.51	-0.14	-8.2%	1.08	0.97	-0.11	-10.0%	1.80
8.	play… sport	0.97	1.59	0.62	63. 9%	0.54	1.14	0.59	110.0%	-46.11
9.	say. tell	1.32	1.35	0.03	2.0%	0.86	1.32	0.46	53.1%	-51.08
10.	steal. rob	0.68	0.97	0.30	44.0%	0.54	1.27	0.73	135.0%	-91.00

Table 5 Questionnaire results for 71 students following the experimental session

	strongly disagree	disagree	neither agree nor disagree	agree	strongly agree	mean	SD
1a. I found the dictionary work easy.	1	26	12	25	5	3.10	1.05
1b. I found the corpus work easy.	1	8	6	34	20	3.93	0.99
2a. I found the dictionary work useful.	1	20	16	29	2	3.16	0.94
2b. I found the corpus work useful.	0	2	7	42	17	4.09	0.69
3a. Thanks to the dictionary information, I think I will avoid certain errors in the future.	0	9	25	30	7	3.49	0.84
3b. Thanks to the corpus data, I think I will avoid certain errors in the future.	0	0	11	44	14	4.04	0.60
4a. I would like to do other similar dictionary activities in class.	3	15	23	22	6	3.19	1.02
4b. I would like to do other similar corpus activities in class.	0	5	15	36	15	3.86	0.83

Figure 1 Percentage of students who would like to explore corpora on computer rather than through prepared paper-based exercises



[i] The school requires all students to obtain a minimum score of 700 points in the TOEIC test by the end of their third year as a prerequisite to graduation, a common practice vocational higher education establishments in France.
[ii] Estling Vannestål & Lindquist (2007) exhibit rare integrity in deliberately publishing disappointing results, with the DDL groups scoring lower than the control group and having less positive attitudes towards grammar at the end of the course. Negative results can help to pin down the limits of DDL and avoid unwarranted enthusiasm in assuming that it works for all learners in all contexts on all language items (Yoon & Hirvela, 2004, p. 279).
[iii] This appendix includes both DDL and traditional materials for the two most successful items: *right* and *want*.