



HAL
open science

Looking (for) empirical evidence of data-driven learning at lower levels

Alex Boulton

► **To cite this version:**

Alex Boulton. Looking (for) empirical evidence of data-driven learning at lower levels. Barbara Lewandowska-Tomaszczyk. *Corpus Linguistics, Computer Tools, and Applications: State of the Art.*, Frankfurt: Peter Lang. *Lodz Studies in Language.*, pp.581-598, 2008. hal-00326983v2

HAL Id: hal-00326983

<https://hal.science/hal-00326983v2>

Submitted on 17 May 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Looking for empirical evidence of data-driven learning at lower levels

Alex Boulton. 2008. In B. Lewandowska-Tomaszczyk (ed.), *Corpus Linguistics, Computer Tools, and Applications: State of the Art*. Frankfurt: Peter Lang, p. 581-598.

alex.boulton@univ-nancy2.fr

CRAPEL – ATILF/CNRS, Nancy-Université

3 place Godefroi de Bouillon

BP 3397

54015 Nancy – cedex

France

Abstract

One application of corpora involves language learning through direct contact with language corpora, often referred to as data-driven learning (DDL). It is often assumed that the approach is unsuitable for lower level learners, but this has rarely been empirically tested. This paper describes a simple experiment on phrasal verbs with 113 lower intermediate learners of English. They were tested on two items – *pick (up)* and *look (up)* – before and after exposure to raw concordance print-outs. This allowed us to test a specific research question: Can lower-level learners extract benefit from the raw data of a concordance print-out? The results are encouraging, suggesting that even lower level learners can derive some benefit from corpus data, and that DDL should be added to their armoury of techniques.

One problem with DDL is that, despite numerous papers citing theoretical arguments in its favour, empirical evidence is hard to come by. If DDL is to be taken seriously and break into mainstream teaching practice, then empirical testing is essential to understand the conditions where it may be of use and to convince the sceptics while not appearing as blind enthusiasts. The importance of a single experiment should not be overstated; what is needed is the weight of evidence from many empirical studies on different questions. The simple experiment reported here is deliberately modest in its design and aims in order to show that useful empirical results are not hard to obtain, and in the hope that others may therefore be encouraged to conduct their own empirical studies.

Keywords

Data-driven learning, concordances, low level learners, empirical research, phrasal verbs.

**PLEASE NOTE: THIS IS A PRE-PRINT VERSION AND MAY CONTAIN
DIFFERENCES WITH THE FINAL PUBLISHED VERSION**

1. Introduction

Data-driven learning (DDL) involves giving students access to sufficient language data so that they can discover and learn for themselves rather than being taught the language explicitly (Johns, 1991). Typical in DDL is the use of a large electronic corpus to which students have direct access, or indirectly for example when a teacher prints out information, often in the form of a concordance. DDL is argued to have many advantages, promoting learner autonomy, increasing language awareness and noticing skills through an inductive approach, improving ability to deal with authentic language, etc., and melds well with the “computer-age” approach of hands-on learning by discovery and doing (Frans, 2000).

If DDL does indeed have all these advantages, then one obvious question is, why isn't everyone doing it? There is unlikely to be one single answer to this, but there are perhaps three broad categories of explanation. Firstly, it may be that teachers and learners are simply unaware of the possibilities (Mukherjee, 2006). Secondly, some objections such as lack of resources, ICT skills, or training are apparently practical, and practical solutions can be proposed: many useful resources can be found free on line; these are often very user-friendly and should require little training prior to use. Finally, other objections to DDL are largely affective, and have received less treatment, presumably as solutions are more difficult to find. For example, learners and teachers may perceive any new approach as not “serious”; they may prefer the certainty (albeit illusory) of traditional rules to the “fuzziness” inherent in corpus work; they may dislike the new roles involved—teachers giving up their control and status as absolute language experts, and learners having to take on responsibility for their own learning. Contrary to popular belief, many learners (and teachers) dislike ICT, the approach may not be universally applicable (perhaps limited to advanced, sophisticated, mature learners), and training—as any learning-to-learn—can be perceived as a waste of time. These problems are particularly acute in traditional teaching and learning cultures such as those prevailing in France (Boulton, 2007b; Brown, 2004). It is difficult to counter such claims without evidence, but the evidence is sadly lacking. Indeed, a further possible reason why DDL has not had greater impact is quite simply that the arguments in favour are theoretical and that there's very little actual proof that it works: the majority of research articles to date are highly descriptive in nature and provide little empirical evidence either way (see Chambers, 2007; Boulton, 2007c). While the enthusiasts can be very enthusiastic indeed, it can be difficult to convince teachers and learners with a full schedule to invest in an approach which has yet to prove itself.

It seems to us that this is not only an unfortunate state of affairs, but also an unnecessary one: small-scale empirical studies can be quite simple to conceive, design and carry out, and do not require great corpus expertise or expensive tools. The main part of this paper describes a simple experiment which requires only a single computer and access to free tools on the internet; use of a print-out means that learners are not required to formulate requests or interact with a computer. The learners themselves are typical of many: although they are at university, they are not studying for a degree in English, are not advanced or especially sophisticated learners of English, are not particularly motivated to learn English, and have not been trained in DDL. The items tested are phrasal verbs (PVs), a staple of English language teaching materials and a “challenging area of English-language learning and teaching” (de Cock, 2005: LS16).

The aims of a single study must of necessity also be quite modest: specifically, we do not set out to prove that DDL is better than other methods, or even that it can enhance learning. The sole aim is to see if, faced with the complex raw data of a concordance print-out, learners are able to detect patterns in concordances and apply them appropriately (Scott & Tribble, 2006). This is not perhaps as anodyne as it seems, as much traditional teaching consists of the presentation of explicit rules, i.e. summaries of patterns which the teachers, materials writers or other experts have previously formulated and “digested” for the learners (Boulton, 2007a). While a single experiment cannot hope to prove the worth of DDL, the present study may provide additional

empirical evidence in addressing questions such as whether there can be benefits for lower-level learners, and not just those that are as “intelligent, sophisticated, and well-motivated” as those in many reported studies (Johns, 1991: 12).

2. Defining phrasal verbs

As PVs have been the subject of several studies in corpus linguistics as well as in DDL (e.g. Waibel, 2005), we begin with a discussion of them before moving on to the experiment itself. To do this, we first look at a sample of published materials, then discuss implications for language learning/teaching.

There is a surprising lack of agreement in defining exactly what PVs are, even among recent corpus-based reference works. One conclusion is that, if the experts find them difficult to pin down, it is only to be expected that PVs will constitute a “difficult” area for language learners. Furthermore, disagreement between experts suggests that there are many different ways to think of PVs, leaving room for each learner to create his or her own version (Allan, 2006: 15). As Aston has repeatedly pointed out with regard to DDL, “descriptions need not be fully accurate: since learning a language involves gradual approximation to the target system, then provided users are aware that their descriptions are partial and approximate, these may still be of value to them” (Aston, 2001: 13). In this view, learners are:

involved in acquiring partial—and only partially accurate—knowledge of patterns in the language, rather than in rivalling professional descriptive linguists... But for many of their purposes, this mattered relatively little: often it was enough for them to get an approximate idea of what an expression meant, or to see that a particular use or collocation occurred relatively frequently. The corpus was providing them with partial information, and provided they were careful not to overgeneralise, that information was often more useful than that available from other reference tools” (Aston, 1996: 187-188).

The following discussion is limited to five different resources for learners and teachers: a phrasal verbs dictionary, a usage manual, and three major grammars, all of which are heavily corpus-influenced. The fact that we can find information about PVs in all of these different types of reference works highlights the fact that PVs are neither purely lexical nor purely grammatical, but somewhere at the interface (c.f. Hunston & Francis, 1998).

Macmillan Phrasal Verbs Plus (Rundell, 2005) was “largely created using”¹ MacMillan’s own 200-million-word World English Corpus. PVs are treated as a verb and “a particle”, which can be either an adverb or a preposition (Potter, 2005: LS2); no further attempt seems to be made to distinguish them, and the entries include both together. This pragmatic approach may be deliberate to avoid bogging learners down in complicated and, perhaps, ultimately unhelpful distinctions. The dictionary also lists verbs which have both types of particles (“phrasal-prepositional verbs”), along with entries on nouns and adjectives formed from all three types. Surprisingly for a dictionary, there is no indication of the number of entries, but at over 500 pages, this volume seems to be more of a reference source than a learning aid.

Practical English Usage (Swan, 2005) is “checked against large electronic databases (‘corpora’) of authentic spoken and written English” (p. ix). This usage manual devotes two sections (a page and a half) to “two-part verbs”: PVs and prepositional verbs are thus kept separate. The distinction given is that the first are followed by adverb particles, the second by prepositions (three-part verbs are followed by both). As many words such as *down*, *in* and *up* can be both adverbs and prepositions, the user is referred to a separate section for further explanation. The main difference there seems to be that adverbs “have no objects” (p.15), although this might not help the learner much when faced with examples contrasting the preposition in *I jumped off*

¹ Background information from the dictionary’s homepage: <http://www.macmillandictionary.com/phrasal-verbs/about/index.html>, accessed April 2007.

the wall and the adverb in *I switched the light off*, as the latter could also be reformulated as *I switched off the light*.

The Cambridge Grammar of English: A Comprehensive Guide, Spoken and Written English Grammar and Use (Carter & McCarthy, 2006) is “informed by” the Cambridge International Corpus of 700 million words (p.11). The eight-page treatment of “multi-word verbs” (p. 429ff) distinguishes PVs from prepositional verbs, as well as phrasal-prepositional verbs. While the prepositional verbs consist of “a verb and a preposition” (p.434), PVs consist of “a lexical verb and a particle” (p. 431). No definition of “particle” is provided in this section, however, and as PVs may be transitive or intransitive, the distinction with (transitive) prepositional verbs may leave some learners confused.

The Cambridge Grammar of the English Language (Huddleston & Pullum, 2002) makes strong and repeated claims to a descriptive rather than a prescriptive approach, although mention of the corpora used is relegated to a foot-note: the Brown corpus, the London/Oslo/Bergen corpus, the Australian Corpus of English and the Wall Street Journal corpus (p. 11). There is a lengthy section on “special verb + preposition combinations and related types of complementation” (p. 272ff), but the authors seek to defuse the PV / prepositional verb contrast thus: “The view taken here... is that [such verb/particle combinations], despite their idiomatic interpretations, do not form syntactic constituents, any more than [others]. It is for this reason that we do not use the term ‘phrasal verb’ in this grammar” (p. 274).

The Longman Grammar of Spoken and Written English (Biber *et al.*, 1999) uses the 40 million words of the Longman Spoken and Written English Corpus, and claims to be “the first empirical corpus-based grammar of English, in that quantitative, empirical investigations of language use are found throughout every chapter” (p. 44). 24 pages are devoted to various types of “multi-word lexical verbs” (p. 403ff), including PVs (verb + adverbial particle), prepositional verbs (verb + preposition), phrasal-prepositional verbs (verb + particle + preposition), and other multi-word verb constructions. The authors describe in some detail the “number of semantic and structural criteria used to distinguish the various types of multi-word verb combinations” (p. 404). To simplify matters, they point out that two are sufficient to identify most cases of (transitive) PVs and prepositional verbs from free combinations: WH-question formation and particle movement. Specifically:

- transitive PVs allow for particle movement, unlike prepositional verbs or free combinations; compare *he picked the phone up* and *he picked up the note* with *I’m waiting for somebody* and **I’m waiting somebody for*);
- WH-questions with prepositional verbs are typically formed with *what* or *who*, while free combinations use *where* and *when*; compare *what are you listening to* with *where were they going [to]*.

To summarise: PVs can form a useful class even without a formal description (Rundell, 2005), or given minimal description (Swan, 2005; Carter & McCarthy, 2006); they can be dismissed as a fictional class of items (Huddleston & Pullum, 2002), or discussed in depth (Biber *et al.*, 1999). These different approaches perhaps underline the limits of a rule-based approach; the most that can be attempted, perhaps, is a description of typical behaviour. This is our approach here: in the spirit of DDL, we shall not attempt a formal description but merely refer to those features which are most salient in our corpus explorations.

3. Phrasal verbs in language learning and teaching

PVs are prominent in English language teaching, no doubt in part precisely because they are quite complex (Consigny, 2005)—there would be little point devoting considerable attention to a

feature which is not “difficult”. The obvious question arises: are PVs sufficiently important to justify the effort required to learn them? Again, corpora can help us here: although frequency is not the sole criterion for deciding what to teach, it seems uncontroversial that if learners are likely to encounter a language item frequently, they should at least be able to understand it.

Biber *et al.* (1999: 409) claim that in conversation and fiction, PVs occur over 1,800 times per million words (pmw), although they are less common in academic registers for example (800 pmw). In a posting to Corpora-List, Reynolds claims that “this seems lower than one would expect and would be evidence against the value of focusing on these [PVs] in TESOL, something that I would say happens quite a lot.”² On the other hand, Fidelholtz Doochin responds with the claim that same figures seem to represent “a very high frequency of occurrence, well worth stressing in an ESL class.”³ Clearly, then, the concept of frequency is a relative one—whether nearly 2,000 pmw is sufficiently frequent to warrant substantial teaching time or not. As so often, corpora can only provide the data, but these need interpreting. Given Biber *et al.*'s (1999: 39) own rule of thumb for an average speech rate of 120 words per minute and an average of 400 words per page of text, users can expect to hear a PV more than once every five minutes (4'43”) in conversation, or read one on every page and a half of fiction (although less in other registers). Such a perspective would seem to justify making sure learners are aware of typical behaviour of PVs as they will meet them on a regular basis in authentic communication.

Another way to apprehend frequency is to compare with a similar set of items, such as modals. Like PVs, these are most common in conversation where, according to Biber *et al.* (1999: 486), they occur 22,000 pmw. In other words, as a group, modals are over ten times as frequent as PVs in this register. Furthermore, there are only nine “central modals” in Biber *et al.*'s (1999: 486) study, compared with the 6,000 PVs described in Parkinson (2001). At the level of individual PVs, the most frequent is *come on*, which occurs nearly 350 pmw. Overall, Biber *et al.* (1999: 410) list only 17 PVs which occur over 40 pmw in conversation. At this rate, one can expect to meet any given one of these PVs only once every 200 minutes in conversation, and others are considerably more rare. Taken together, all of this suggests that PVs as a class might be worth while, but individual ones much less so.

If PVs do constitute a difficult area for learners, both semantically and syntactically, perhaps teaching for productive use could be confined to a small number of the most frequent PVs for lower level learners. Certainly, there are cases when “a phrasal verb is the most natural-sounding way of expressing a particular idea”, and learners may then be “encouraged to use phrasal verbs as and when they are appropriate” (Fletcher, 2005: LS13). But for less frequent items, there may be a case for working more on communicative strategies for dealing with them receptively. Many learners seem to have worked this out for themselves: de Cock (2005: LS16) claims that “avoidance” is the main problem area for learners (c.f. also Gabel, 2001: 284). If they are avoiding them, presumably they are using other, simpler items instead. The Oxford Phrasal Verbs Dictionary (Parkinson, 2001) explicitly claims to “give students help with synonyms, allowing them to decide whether a phrasal verb or a single-word verb is the more appropriate choice”.⁴ Marks (2005: LS12) provides a clue as to what these synonyms might be, drawing attention to “cases where a Germanic phrasal verb, noun or adjective has an exact equivalent word whose origins are Latin / French”, such as *put forward* / *propose*. Significantly, *propose* is morphologically very similar to *put forward*: the Latin-derived morphemes *pro-* and *-pose* correspond closely to the Germanic *forward* and *put* respectively. This gives the lie to the idea that PVs are somehow peculiarly “English”: “contrary to popular belief, many other languages also have vocabulary that is very similar to English phrasal verbs” (Marks, 2005: LS10). One might think first of the separable verbs in Germanic languages, but Marks (2005: LS11) also

² <http://torvald.aksis.uib.no/corpora/2003-1/0177.html> 6th March 2003.

³ <http://torvald.aksis.uib.no/corpora/2003-1/0420.html> 14th May 2003.

⁴ From the OUP promotional page: <http://www.oup.com/elt/catalogue/isbn/0-19-431721-8?cc=gb>, accessed April 2007.

draws attention to “very similar combinations in vocabulary that is derived from Latin and French too. In this case the order is particle+verb, and the spelling is always a single word”. The examples given include *progress / regress / digress / transgress*, and *compose / compound / compost / depose / deposit / expose / oppose*, along with their noun and adjective derivatives. Drawing learners’ attention to such patterns might help them to find more reliable equivalents, and perhaps even take some of the drama out of PVs.

However, given that there are at least some PVs which learners are likely to need, and that these remain difficult with traditional teaching, we might explore alternative methods. De Cock (2005: LS20) shows how corpora can be of use here:

in view of all the evidence of the difficulties that phrasal verbs can cause for learners, it is quite clear that these verbs ought to be treated as “chunks”—together with their syntactic, contextual, and collocational features, rather than in isolation. Providing learners with *lists* of phrasal verbs to learn by heart ought to be a thing of the past.

Corpus-based studies of phrasal verbs clearly show the need for a contextualized approach based on (semi-) authentic texts, as this will enable teachers to draw learners’ attention to: whether or not certain phrasal verbs are more typical of speech or of writing; the syntactic environment of phrasal verbs; the words that phrasal verbs tend to combine with.

It seems then that a DDL approach might be of particular benefit in this area of English language learning, especially suited to pattern-detection (Olivier *et al.*, 2007). As Cobb *et al.* (2001: 6) point out, while explanations and individual examples may prove insufficient, the point of DDL in the form of concordances is that “contextual inference can be substantially supported by multiplying the number of contexts available for a given word.”

4. Method

In the rest of this paper we describe a simple experiment to see if a sample of lower-level learners of English could detect patterns of PV use in concordance print-outs of authentic language, and apply their findings to new contexts. The corpus data for the experiment are taken from the British National Corpus (BNC), a large (100-million word) corpus of contemporary British English. It is available for purchase, but various derivatives are available entirely free on line. In particular, simple queries can be conducted from the BNC’s own site,⁵ while Mark Davies’ Variation In English Words and Phrases (VIEW) interface⁶ provides a number of quite sophisticated alternatives, especially for searching by register. Understandably, neither of these websites are as complete as the commercialised version, which inevitably leads to some ad hoc DIY queries and manual editing of data (Sinclair, 2005). However, they have the enormous advantage of being freely available, and their simplicity is no doubt an advantage for most learners.

As we wanted to keep the experiment simple, we selected only two PVs, different verbs with the same particle: *look up* and *pick up*. There were several reasons for this choice, based on frequency, meaning and grammar. Firstly, *up* is the most frequent adverb particle in VIEW, and is more frequently an adverb particle than a preposition or any other part of speech. *Look* is likely to be familiar to learners as it is one of the most frequent verbs in English, occurring 520.86 per million words—over eight times as frequent as *pick* at 57.92 pmw.⁷ Both PVs are among the ten most frequent in VIEW,⁸ although *pick up* is four times more frequent than *look up* (25.30 pmw

⁵ <http://sara.natcorp.ox.ac.uk/lookup.html>, accessed April 2007.

⁶ Previously at <http://view.byu.edu>, now at <http://corpus.byu.edu/bnc>

⁷ Kilgarriff’s lemmatised frequency list for the BNC ranks *look* as the 14th most frequent verb and *pick* as the 150th. <http://www.kilgarriff.co.uk/bnc-readme.html#lemmatised>, accessed February 2007.

⁸ Biber *et al.* (1999: 410) also list *pick up* and *look up* among the “activity transitive” PVs occurring over 40 pmw in at least one register, with *pick up* at number two and *look up* at number ten.

and 5.82 pmw respectively). Just over half (50.6%) of all occurrences of *pick* are as a PV, compared to only 7.4% of occurrences of *look*. The ranges of meaning are to an extent reflected by frequency: the Oxford Advanced Learner's Dictionary (Hornby & Wehmeier, 2007)⁹ lists six main meanings for *look*, but only three for *pick*. On the other hand, Macmillan Phrasal Verbs Plus (Rundell, 2005) gives 18 separate meanings for *pick up*, but only four for *look up*. Grammatically, both PVs can be separable, both can be used transitively, both are regular in form, and both occur overwhelmingly as verbs, making for easier and more reliable searching.

The experiment involved 113 students in their first year of university studies at a general engineering college in the north-east of France.¹⁰ The average age was just over 18, and unsurprisingly in this discipline, the majority were male (84%), and all but eight were native French speakers (four Chinese, four Arabic speakers). Despite an average of six and a half years of English at school, their motivations and efforts lie elsewhere than with English, which is perceived by many as a burden to be endured rather than a useful vocational skill for their later careers (Brown, 2004). The levels too are surprisingly low, as can be seen from the scores of the 103 who had sat a start-of-year English placement test. This was based on a full-length TOEIC¹¹ exam, as the school requires a minimum of 720 points out of 990 on this test by the end of the second year. In the placement test, these students' scores averaged around 450, towards the lower end of the TOEIC "intermediate" level. To give a better idea, the raw scores ranged from 33% to 79% (average 50.25%), and it should be remembered that this is an entirely multiple-choice test where completely random answers should obtain a score over 25%.

The experiment was conducted at the start of the second semester during normal class time and with the students' regular teachers. All the data used were taken from VIEW: we first downloaded 50 concordance lines for each of *looked*, *looked up*, *picked* and *picked up*. Only 35 occurrences of each were required for the entire experiment, so a pool of 50 allowed us some freeway to select the most appropriate choices (Todd, 2001: 92), especially by eliminating erroneous hits, items which were repeated or very similar, and so on. Even with the most advanced tools, this is the kind of manual manipulation and selection which is common in corpus linguistics (Sinclair, 2005), even more so in DDL where usable data are more important than rigorous precision. However, the concordance lines were left entirely unedited, and assigned at random to the different stages of the test, barring minor rearrangement to ensure similar examples did not occur together.

In the first stage of the experiment, a pre-test was distributed containing 10 concordance lines for *picked* / *picked up* and 10 for *looked* / *looked up*; half the PVs were separable. Each line provided a simple binary choice for the key item: half were for the PV, half for the non-PV usage (see examples in Figure 1). This task-type is arguably inauthentic, representing communicatively neither production nor reception; however, it has the advantage of being simple to produce and to score, and because it can be completed quickly there is less likelihood of numerous blank answers. It is also familiar to students, which might therefore go some way towards compensating for their unfamiliarity with the concordance format: apart from one brief encounter several months earlier, none had ever seen such data before. Students were reassured that a) each line constituted a separate context and the lines did not produce coherent text; b) each line was an extract with the key item in the middle and not a complete sentence.

" In the morning, after an early breakfast of freshly **picked mushrooms** / **picked up mushrooms**, Graham, myself and our hangovers ventured any other (and Goldberg, pausing in his typing, **picked his pen** / **picked up his pen** and put a small question mark in the margin year-old goalkeeper. "He was incredible. The Poles **picked an attacking side** / **picked up an attacking side** and had us under siege for a while from Addis Ababa had stopped at Dire Dawa but had **picked no passengers** / **picked up no passengers**. It had, however, collected eight and ground and we all fell on top of each other. Quickly, I **picked myself** / **picked myself up** and ran for my life through the trees. I

⁹ Available free on line, based on corpora including the British National Corpus.

<http://www.oup.com/elt/catalogue/teachersites/oald7/lookup?cc=global>, accessed April 2007.

¹⁰ Ecole Supérieure des Sciences et Technologies de l'Ingénieur de Nancy (Université Henri Poincaré – Nancy I).

¹¹ Test of English for International Communication: <http://www.toeic.eu/>, accessed: April 2007.

Fig 1. Sample test questions for *picked / picked up*.

Once the pre-test had been collected in, information sheets were distributed containing 25 concordance lines each of *picked*, *picked up*, *picked (something) up*; and *looked*, *looked up*, *looked (something) up*. The presentation was the same as for the test instrument, except that there was only a single item underlined in the middle of each line. After about 10 minutes for learners to make what sense they could of the concordances, the post-test was distributed; the presentation and task here were identical to the pre-test. This time students were allowed to consult the concordance sheets to help complete the test. In all, the whole session lasted just under 30 minutes.

5. Results and discussion

Table 1 below shows the overall scores as percentages in each test. Given that there were only two possibilities for each question, the results in each test should be 50% by chance alone. Column 3 shows the changes in scores (i.e. Test 2 minus Test 1), and column 4 shows change as a percentage (i.e. Test 2 divided by Test 1).

Table 1. Percentage scores and changes, test 1 > test 2

	Test 1	Test 2	change (T1-T2)	% change (T1/T2 x 100)
<i>look</i>	54.87	58.05	+ 3.19	+ 5.81
<i>look up</i>	50.62	60.53	+ 9.91	+ 19.58
<i>pick</i>	52.21	60.71	+ 8.50	+ 16.27
<i>pick up</i>	52.74	66.37	+ 13.63	+ 25.84
AVE	52.61	61.42	+ 8.81	+ 16.74

One immediately striking feature is the extremely low scores overall—only 57.01% over both tests. This is likely to be at least in part due to the learners' low levels of ability in English as a whole, as we have seen, but other possibilities need to be considered. Firstly, it could be that the target items themselves were difficult, and indeed informal feedback suggests that students were generally unfamiliar with *pick* and with both PVs. However, given the slight differences between these items and *look* as a base verb in Test 1, this too seems unlikely to explain the poor performance overall. It seems unlikely too that test design was the cause of the low scores, as students had only to select one of two answers given for each question, a familiar test format. If the instructions had been unclear, we would also have expected large numbers of blank responses, but in fact less than 2% of answers were left blank in either test. It might simply be that the students did not take the test very seriously, choosing answers almost at random. This is likely to be the case for at least some students, who do not see English as an important part of their engineering studies; but had this been the major cause of the low results, we would have expected there to be little difference in results between Test 1 and Test 2, which is clearly not the case. Finally, it cannot be discounted that students have difficulty working with concordances (Tribble & Jones, 1997), especially of such complex authentic data (Wible *et al.*, 2002). None of the students had received any training for this and can be considered novice users.

In the end, the low absolute scores are not in themselves a problem, as long as they allow us to distinguish between the different variables. It should be remembered that the point of the test was to see if learners could detect patterns through concordances, not to compare whether such an approach is more or less efficient than other methods. The fact that the overall scores on all items are above average even in Test 1, even though only slightly so in some cases, supports this hypothesis. More important is the improvement between the two tests: scores increased for all test items by an average of 8.81, which represents a 16.74% improvement overall ($p < 0.001$). The

only plausible explanation for such a significant difference is that learners were able to detect patterns in the concordance data they were exposed to after Test 1 and to apply those patterns to the questions in Test 2.

Table 2 shows the grouped scores for each verb, LOOK (i.e. *looked* plus *looked up*) and PICK (i.e. *picked* plus *picked up*); the improvement is significant for both verbs ($p < 0.001$). Intriguingly, the scores for each were virtually identical in Test 1, even though LOOK was presumed to be the more familiar verb; furthermore, PICK actually scored significantly higher than LOOK in Test 2. The most probable explanation for us is that students concentrated more on the less familiar item, assuming they already knew *look* but not *pick*. More research would be needed to check this, but one interpretation is that concordances are most useful when learners perceive that they are learning something new rather than refining existing knowledge. This sounds reasonable, but on reflection it runs counter to the idea that dictionaries and other reference tools are most suited for providing initial information about new items, while concordances are best left for increasing the depth of that knowledge. Of course, it could be that both tools are best for the initial information, which is easier to gain than more refined knowledge (Allan, 2006).

Table 2. Results: *look* vs. *pick*

	Test 1	Test 2	change (T1-T2)	% change (T1/T2 x 100)
<i>look + look up</i>	52.74	59.29	+ 6.55	+ 12.42
<i>pick + pick up</i>	52.48	63.54	+ 11.06	+ 21.08

Table 3 shows the grouped scores for the base verbs (i.e. *looked* plus *picked*) compared against the PVs (i.e. *looked up* plus *picked up*). Again, in Test 1 the results for the supposedly simpler items (base verbs) are slightly higher than those for the more difficult items (PVs); scores for each increase significantly in Test 2 ($p < 0.001$), but the improvement is greater for the PVs. The same explanation would seem to apply here as above, namely that students concentrated more on what they perceived as more difficult or more intriguing, or where they felt they were acquiring new knowledge.

Table 3. Results: base verbs vs. phrasal verbs

	Test 1	Test 2	change (T1-T2)	% change (T1/T2 x 100)
BASE VERB	53.54	59.38	+ 5.84	+ 10.91
PHRASAL VERB	51.68	63.45	+ 11.77	+ 22.77

These generally robust scores hide a certain diversity at the level of the individual learner. Scores in Test 1 ranged from 15% to 85%, in Test 2 from 30% to 95%. The correlation coefficient is not particularly high (0.32), although nearly two thirds of students increased their scores between the two tests (64.60%). However, this still leaves over a third who either scored the same (9.73%) or lower (25.66%) in Test 2. Unsurprisingly, most of the decreases were among the students who had scored highest in the first test—the only way was down, so to speak: over half of those whose score decreased had scored in the top fifth in Test 1.

To see if the general level of English was important, we divided the data into three sets according to the students' scores at the start-of-year placement test. The highest band (group A) had an average mark of 64% (at the lower end of TOEIC's "basic working proficiency"), group B 50% ("intermediate"), and group C 41% ("elementary"). As Figure 2 shows, the results are as would be expected, with group A achieving the best scores on all counts.

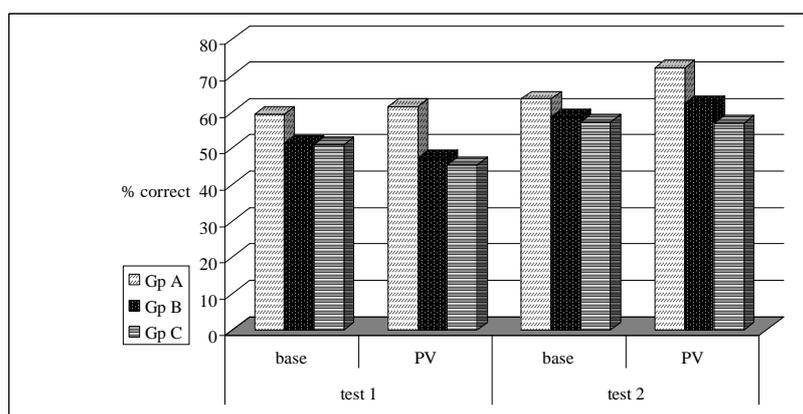


Figure 2. Results by level

Group A scored over 60% in Test 1, with the others scoring just under 50% (Table 4); Group A was also the only group to score higher on the PVs than on the base verbs in Test 1. They still had the highest scores in Test 2, although the improvement was less than for the other two groups. This is no doubt partly because, as we saw earlier, the more advanced levels of knowledge are more difficult to acquire, and the higher the starting score, the less room there is for improvement in any case. That said, it does suggest that DDL might be beneficial to less advanced learners: our students are not what would normally be considered “advanced” learners of English, and the middle group marked up the best improvement between tests.

Table 4. Results by level

		Test 1		Test 2		change (T1-T2)	% change (T1/T2 x 100)
Gp A	base	60.44	59.41	67.79	63.53	7.35	12.17
	PV		61.47		72.06		
Gp B	base	49.41	51.47	60.44	58.53	11.03	22.32
	PV		47.35		62.35		
Gp C	base	48.14	50.86	57.00	57.14	8.86	18.40
	PV		45.43		56.86		

6. Conclusion

This paper described a simple experiment designed to see if learners were able to detect patterns of meaning and use among raw concordances. The results are encouraging: although the higher-level learners scored highest, all levels showed improvement, suggesting that even lower level learners may be able to derive some benefit from data-driven learning. Furthermore, the language items studied, phrasal verbs, are traditionally considered difficult for French learners, but greater improvement was recorded on these than on the simpler items (base verbs). This supports the idea that pattern-detection may be more useful where rules and explicit instruction are inadequate, especially if they allow a “broader perspective” on the use or meaning of language items (Levy, 1990).

We argue that far more empirical research is needed on all aspects of DDL if it is to convince a wider audience and break out of its current research environment (c.f. Cobb, 1997). Useful evidence can be gathered from very simple experiments which do not necessarily require extensive resources or great expertise. The present study is one of a series of experiments (Boulton, 2006; 2007a) using only free software, and which do not require extensive learner-

training or access to a computer laboratory in class. Although this experiment is not designed to test whether DDL leads to efficient learning, the results do provide further evidence of learners' ability to detect patterns, which is an absolute prerequisite for such an approach.

REFERENCES

- Allan, R. (2006). "Data-driven learning and vocabulary: Investigating the use of concordances with advanced learners of English". *Occasional Paper 66*, Centre for Language and Communication Studies. Dublin: Trinity College Dublin.
- Aston, G. (1996). "The British National Corpus as a language learner resource". In Botley, S., Glass, J., McEnery, A. & Wilson, A. (eds.) *Proceedings of TALC 1996: UCREL Technical Papers 9*: 178-191.
- Aston, G. (2001). "Learning with corpora: An overview". In Aston, G. (Ed.) *Learning with Corpora*. Houston: Athelstan: 7-45.
- Biber, D., Johansson, S., Leech, G., Conrad, S. & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. London: Pearson.
- Boulton, A. (2006). "Tricky to teach, easier to learn: Empirical evidence for corpus use in the language classroom". *American Association of Applied Corpus Linguistics (AAACL)*. Northern Arizona University, 20-22 October.
- Boulton, A. (2007a). "DDL is in the details... and in the big themes". *4th Corpus Linguistics conference*. University of Birmingham Centre for Corpus Research, Birmingham UK, 27-30 July.
- Boulton, A. (2007b). "Esprit de corpus: Promouvoir l'exploitation de corpus en apprentissage des langues". *5^{ème} Journée de la Linguistique de Corpus*. Université Bretagne Sud, France, 13-15 September.
- Boulton, A. (2007c). "But where's the proof? The need for empirical evidence for data-driven learning". *40th Annual Meeting of the British Association for Applied Linguistics: Technology, Ideology and Practice in Applied Linguistics*. University of Edinburgh, Scotland, 6-8 September.
- Brown, D. (2004). "Attitude, motivation et la théorie des objectifs dans l'environnement L2 à l'université française". *APLIUT*. Aix en Provence, 3-5 June.
- Carter, R. & McCarthy, M. (2006). *Cambridge Grammar of English: A Comprehensive Guide. Spoken and Written English Grammar and Use*. Cambridge: CUP.
- Chambers, A. (2007). "Popularising corpus consultation by language learners and teachers". In Hidalgo, E., Quereda, L. & Santana, J. (eds.) *Corpora in the Foreign Language Classroom*. Amsterdam: Rodopi: 3-16.
- Cobb, T. (1997). *From Concord to Lexicon: Development and Test of a Corpus-Based Lexical Tutor*. Montreal: Concordia University, Unpublished PhD thesis.
(<http://www.er.uqam.ca/nobel/r21270/webthesis/Thesis0.html>, accessed March 2006)
- Cobb, T., Greaves, C. & Horst, M. (2001). "Can the rate of lexical acquisition from reading be increased? An experiment in reading French with a suite of on-line resources". In Raymond, P. & Cornaire, C. (eds.) *Regards sur la Didactique des Langues Secondes*. Montréal: Editions Logique: 133-135. (<http://www.er.uqam.ca/nobel/r21270/cv/BouleE.htm>, accessed July 2006)
- Cock, S. de (2005). "Learners and phrasal verbs". In Rundell, M. (ed.) *Macmillan Phrasal Verbs Plus*. Oxford: Macmillan: LS 16-20.
- Consigny, A. (2005). "Chunks of meanings in English phrasal verbs". In Hamm, A. (ed.) *Language Chunks and Linguistic Units. Ranam 38*: 75-84.
- Fletcher, B. (2005). "Register and phrasal verbs". In Rundell, M. (ed.) *Macmillan Phrasal Verbs Plus*. Oxford: Macmillan: LS 13-15.

- Frand, J. (2000). "The Information-Age mindset: Changes in students ad implications for higher education". *EDUCAUSE Review* 35/5: 14-24.
<http://www.educause.edu/pub/er/erm00/articles005/erm0051.pdf>, accessed October 2006)
- Gabel, S. (2001). "Over-indulgence and under-representation in interlanguage. Reflections on the utilization of concordancers in self-directed foreign language learning". *Computer Assisted Language Learning* 14: 269-288.
- Hornby, A. & Wehmeier, S. (2007). *Oxford Advanced Learner's Dictionary*, 7th edition. Oxford: Oxford University Press.
- Huddleston, R. & Pullum, G. (2002). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Hunston, S. & Francis, G. (1998). "Verbs observed: A corpus-driven pedagogic grammar". *Applied Linguistics* 19/1: 45-72.
- Johns, T. (1991). "Should you be persuaded: Two examples of data-driven learning". In Johns, T. & King, P. (eds.) *Classroom Concordancing. English Language Research Journal* 4: 1-16.
- Levy, M. (1990). "Concordances and their integration into a word-processing environment for language learners". *System* 8/2: 177-188.
- Marks, J. (2005). "Phrasal verbs and other 'phrasal' vocabulary". In Rundell, M. (ed.) *Macmillan Phrasal Verbs Plus*. Oxford: Macmillan: LS 10-12.
- Mukherjee, J. (2006). "Corpus linguistics and language pedagogy: The state of the art—and beyond". In Braun, S., Kohn, K. & Mukherjee, J. (eds.) *Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods*. Frankfurt: Peter Lang: 5-24.
- Olivier, N., Brems, L., Davidse, K., Speelman, D. & Cuyckens, H. (2007). "Pattern-learning and pattern-description: An integrated approach to proficiency and research for students of English". In Hidalgo, E., Quereda, L. & Santana, J. (eds.) *Corpora in the Foreign Language Classroom*. Amsterdam: Rodopi: 221-235.
- Parkinson, D. (2001). *Oxford Phrasal Verbs Dictionary for Learners of English, new edition*. Oxford: Oxford University Press.
- Potter, E. (2005). "The syntactic behaviour of phrasal verbs". In Rundell, M. (ed.) *Macmillan Phrasal Verbs Plus*. Oxford: Macmillan: LS2-4.
- Rundell, M. (ed.) (2005). *Macmillan Phrasal Verbs Plus*. Oxford: Macmillan .
- Scott, M. & Tribble, C. (2006). *Textual Patterns: Key Words and Corpus Analysis in Language Education*. Amsterdam: John Benjamins.
- Sinclair, J. (2005). "Corpus and text-basic principles". In Wynne, M. (ed.) *Developing Linguistic Corpora: A Guide to Good Practice*. Oxford: Oxbow Books: 1-16. (<http://ahds.ac.uk/linguistic-corpora/>, accessed February 2006)
- Swan, M. (2005). *Practical English Usage*, 3rd edition. Oxford: Oxford University Press.
- Todd, R.W. (2001). "Induction from self-selected concordances and self-correction". *System* 29/1: 91-102.
- Tribble, C. & Jones, G. (1997). *Concordances in the Classroom*, 2nd edition. Houston: Athelstan.
- Waibel, B. (2005). "Phrasal verbs and the foreign language learner: Results from a pilot study based on the International Corpus of Learner English". In Hamm, A. (ed.) *Language Chunks and Linguistic Units. Ranam* 38: 65-74.
- Wible, D., Chien, F., Kuo, C-H. & Wang, C. (2002). "Toward automating a personalized concordancer for data-driven learning: A lexical difficulty filter for language learners". In Kettemann, B. & Marko G. (eds.) *Teaching and Learning by Doing Corpus Analysis*. Amsterdam: Rodopi: 147-154.