Corpus consultation for ESP: A review of empirical research
Alex Boulton

To cite this version:
Alex Boulton. Corpus consultation for ESP: A review of empirical research. Alex Boulton, Shirley Carter-Thomas


HAL Id: hal-00683784
https://hal.archives-ouvertes.fr/hal-00683784
Submitted on 7 Aug 2018

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.
Abstract
Corpus linguistics is commonly concerned with language description, but the tools and
techniques can also be used by language learners themselves in what has come to be known as
data-driven learning or DDL (Johns & King 1991). This paper reviews 20 studies representing
empirical evaluations of corpus use by second or foreign language users within the context of
ESP. The wide variety of tools, corpora, course objectives, experiment designs and research
questions makes a thorough meta-analysis impossible. However, analysis of the individual
papers shows that students can use corpora successfully for ESP and are generally favourable to
the approach, whether as a learning tool (especially for vocabulary and lexico-grammar) or as a
reference resource (especially for writing).

Key words: data-driven learning; corpus consultation; language learning; empirical evidence;
learning outcomes

1. Corpora for ESP learners
Corpus linguistics has contributed considerably to the arsenal of tools available for language
description, with methodologies applicable to virtually any area of language study. Yet like most
academic disciplines, corpus linguistics is generally limited to researchers and other
professionals; in the fields of foreign/second language teaching and learning, this involves
considerable mediation upstream (by linguists, lexicographers, publishers, decision-makers,
materials writers, etc.) before the results filter down to the end users – teachers and learners.
This is nothing new: in the field of languages for specific purposes, as long ago as 1987
Hutchinson and Waters remarked that “in its development up to now, ESP has paid scant
attention to the question of how people learn, focusing instead on the question of what people
learn” (p. 2). The corpus approach to ‘what’ questions have contributed significantly to our
understanding of genre and text type as well as discipline-specific language, not least the
Academic Word List (Coxhead 2002) and more recently the Academic Formulas List (Simpson-
Vlach & Ellis 2010). An alternative would be for corpus data to feature directly in the learning
process, with learners acting as ‘researchers’ to explore the language themselves. This clearly
goes beyond language description to impact on methodology and is commonly referred to as
data-driven learning (DDL), an idea particularly associated with the work of Tim Johns (e.g.

Whether as a teaching aid, learning tool or reference resource, corpora are of particular
relevance to ESP by its very nature; for Bennett (2010: 11), “ESP is probably one of the most
obvious and pointed applications of corpus linguistics.” On a practical level, there are
innumerable resources for English for ‘general purposes’ as the market world-wide is potentially
immense, but the more specific the need, the more difficult it is to develop generic materials
and resources that are financially worthwhile for publishers to invest in. Traditional resources (coursebooks, dictionaries, grammar books, usage manuals, style guides, and so on) are unlikely to be able to cover the highly specific language that many ESP users may need, or to keep up with changes as the terminology evolves in specific areas.

Corpus linguistics allows precisely this, providing a framework to highlight the highly conventionalised language used in specialist disciplines, especially where the focus is on a specific genre or text type. As corpus linguistics tools and techniques become increasingly available and user-friendly, teachers and learners can use them to uncover and highlight such regularities at a variety of levels, from discourse to collocations to lexical bundles to individual specialist terminology. Small ESP corpora are especially suited to this, as they are comparatively easy to compile and manipulate even for the end-users – non-native speaker teachers and learners. Compilation of ad hoc or disposable corpora allows even greater specialisation for occasional needs, and can be complemented by a web-as-corpus approach to take into account the ever-changing terminology encountered in many ESP domains. Given the potential for corpus use to spread to such end-users, it becomes increasingly important to investigate what they can do with the corpora themselves and how they can interact with the wealth of data thus afforded. This, of course, is data-driven learning in its wider sense.

The objective of this chapter is to see how effective or efficient the approach is when actually applied in ESP. The rest of this section looks at empirical research in DDL for general purposes, before homing in on issues related to evaluating DDL in ESP contexts. Twenty empirical evaluations of DDL for ESP are identified, and form the basis for the rest of the chapter.

1.1. *Empirical research in data-driven learning*

As corpus linguistics has shown time and again, popular ideas about language are often wrong: facts need checking, and simple repetition does not make something true. The same is true of tools and techniques, approaches and methodologies in language teaching and learning: theoretical arguments need putting to the test. The last 20 years has seen substantial research interest in DDL, with countless scholarly articles arguing its various merits, though it is frequently remarked that empirical research is disturbingly lacking even today (e.g. Conrad 2005, Cresswell 2007, Estling Vannestål & Lindquist 2007, O’Keeffe *et al.* 2007, Johansson 2009). However, an ongoing collection has uncovered over 80 studies to date which attempt to evaluate some aspect of DDL.¹ This is a surprisingly large number given the repeated lamentation of the lack of empirical studies. As fully three quarters of them appear in well-known journals or volumes by international publishers, the discrepancy merits some comment.

Firstly, we might be looking for different things, as there is certainly a lack of agreement on exactly what counts as DDL: those who claim to be implementing it make use of an extraordinary variety of activities and techniques, including Johns himself in his own work. In particular, one might wonder whether the *data* have to be in the form of a corpus or can include a single text or extract; whether *driven* refers only to serendipitous hands-on computer work or can include prepared materials on paper; whether *learning* excludes the use of corpora

This is a pre-publication version. For the version of record, please see DOI: 10.1075/scl.52.11bou
(or email me at alex.boulton@atilf.fr)
as a reference resource for writing or translating, among others. There certainly seems to be a kind of ‘prototype’ form of DDL referring to “the hands-on use of authentic corpus data (concordances) by advanced, sophisticated foreign or second language learners in higher education for inductive, self-directed language learning of advanced usage” (Boulton in press a). But none of these aspects is necessary or sufficient to characterise DDL to the satisfaction of all concerned. For our purposes, DDL is taken in a broad sense to refer to any use of language corpora by second or foreign language users.

Similarly, there are no hard-and-fast criteria for what counts as ‘empirical evaluation’. One might restrict the survey to quantitative assessment of learning outcomes (i.e. discarding qualitative studies, or those evaluating learners’ behaviours, reactions, or uses of corpora as a reference resource), though this still leaves 29 individual studies, a far from negligible number. Again, the definition adopted here is fairly general in order to ensure maximum coverage, and is taken to include any studies which “subject some aspect of DDL to observation or experimentation with some kind of externally validated evaluation other than the researchers’ own intuition” (Boulton 2010a: 130).

Finally, of course, it is inevitably a subjective question as to whether the existing quantity of research constitutes a significant ‘lack’, and corpus linguists may, by the nature of their field, be more demanding of evidence than in other areas of applied linguistics. Nunan (2007: 9) claims that the “designer methods movement” of the 1970s was largely “‘data-free’, drawing sustenance from rhetoric rather than empirical support”, but that this era is now largely over as teachers have become “less inclined to embrace pedagogical proposals without some kind of evidence.” However, many currently highly-cherished ideas do seem to be based more on argument and zeitgeist than on a solid base of empirical evidence – at least, if one is to accept critiques by researchers such as Truscott (1998) on noticing, Matthews (2003) on constructivism, Mayer (2004) on discovery learning, Swan (2005) on task-based learning, Kirschner et al. (2006) on induction, Gilmore (2007) on authentic materials, and so on. But these are generally lone voices which do little to limit the spread of the approaches concerned. On the one hand, progress would be far slower if teachers had to wait for the evidence to be in before adopting new practices; on the other, evidence is important, empirical research should not be neglected, and the results deserve to be more widely known.

One thing seems sure: empirical studies are extremely difficult to conduct in applied linguistics, as indeed in any other area of social sciences or the humanities, given the enormous number of different factors which can impact on any given study. For this reason, a common approach is to try to break larger issues down into smaller, more manageable hypotheses in the form of specific research questions (cf. Ellis 1998). This piecemeal approach should allow the gradual accumulation of a larger picture by combining the results from many studies – such is indeed the major purpose of the present paper. Larsen-Freeman and Cameron (2008) take issue with the underlying philosophy of this, as they see language as a complex system which is immune to such a reductionist approach: “The unknowableness and interconnectedness of systems makes it much more difficult, if not impossible, to isolate independent variables that act in causal...
For them, because of the importance of individual differences it is not possible
to control all variables, and therefore not possible to make complete and accurate predictions in
any study. While this much seems certain, few individual studies do actually make strong claims,
and researchers are at pains to hedge their findings (cf. Boulton 2009b). What is important is,
again, the overall picture emerging from numerous studies conducted in different conditions to
identify general tendencies which may be worth exploring further.

1.2. From DDL to ESP
The first academic publication to discuss uses of corpora in language teaching and learning
comes from America, produced by McKay in 1980 working in San Francisco. However, earlier
roots of DDL can be found in Birmingham: McEnery and Wilson (1997: 12) trace it back to the
work of Peter Roe at Aston in 1969. It is particularly associated with the work of Tim Johns at
Birmingham University, who had been working in this direction for years before he made the
first brief mention of it in Higgins and Johns in 1984, followed by a more substantial treatment
in Johns 1988. The term ‘data-driven learning’ in this context first appears in print in Johns
(1990), and the field really took off the following year with the publication of the seminal
collection of papers in the ELR Journal special issue on Classroom Concordancing edited by
Johns and King (1991). Since then there have been literally hundreds of papers given over to
corpus uses for language teaching and learning, in addition to a number of recent full-length
book treatments, such as Reppen (2010), Bennett (2010), Anderson and Corbett (2009),

Previous reviews of empirical evidence in DDL include Chambers (2007), who covered 12
papers, and Boulton (2010a), who found 27 studies evaluating learning outcomes from DDL, a
subset of the 80 studies in the updated list. The present chapter has a very specific focus,
looking only at empirical evaluations of corpus use by second or foreign language users within
the context of ESP. It thus does not attempt to cover the bulk of work in DDL, which remains
largely descriptive or argumentative (cf. Granger 2009). The same is true even of some
important work in ESP, including book-length treatments by Bowker and Pearson (2002) and
Gavioli (2005) – indeed, the “lack of quantitative evidence” was one of the main reservations
“texts on corpus analysis that ignore quantitative evidence are doomed to endless instances of
phrases such as ‘it seems to me’.” It also omits studies featuring corpora in languages other
than English, though there are not many of these (and most feature European languages,
especially French, German and Italian). Also neglected are studies featuring native speakers of
English, even though novice researchers, for example, may benefit from introduction to the
language appropriate in different genres of their discipline; but again there are very few studies
in this area, and including them might give an unbalanced presentation. One regrettable
omission is Bowker (1998, 1999), who reports using ESP corpora for Irish students translating
into their mother tongue.
More difficult to isolate is whether students are using corpora for specific purposes or not. As Hutchinson and Waters (1987: 19) point out, “ESP is not a matter of teaching ‘specialised varieties’ of English.” Rather:

ESP must be seen as an *approach* not as a *product*... It is an approach to language learning, which is based on learner need. The foundation of all ESP is the simple question: Why does this learner need to learn a foreign language? ... ESP, then, is an approach to language teaching in which all decisions as to content and method are based on the learner’s reason for learning. (Hutchinson & Waters 1987: 19)

Firstly, the nature of the corpus is no guarantee, as general corpora can also be used for ESP (cf. Maniez, this volume). Secondly, it is also misleading to rely only on the type of students participating: students majoring in English may follow an ESP course and thus be included here; conversely, non-majors may use corpora for general English only and not ESP (e.g. Chan & Liou 2005; Landure & Boulton 2010). Similarly, studies featuring both majors and non-majors will be excluded if they do not have an ESP focus, such as Tian (2005). Varley (2009) is particularly worth mentioning in this respect: although his focus was on general English, the course included both majors and non-majors; in the follow-up questionnaire, it was the latter who were more receptive and who expressed greater interest in using corpora in the future. Similarly excluded are studies such as Maia (1997) and Boulton (in press b): although the majors in these studies are able to compile their own corpora to pursue their own research topics, at times highly specific, this is rather incidental to the main focus of these studies. Finally, ESP is taken to include English for academic purposes (EAP), insofar as “EAP can be seen to be a subset of English for Specific Purposes” (Thompson 2006: n.p.), even if this is not necessarily discipline-specific. On the other hand, the students involved in studies such as Wu et al. (2010) and Allan (2006) did have a specific academic objective – namely to pass IELTS and CAE examinations – but are not covered here as the exams themselves are general English and not ESP.
suggests that “corpus-based courses on academic writing are still at an exploratory stage” even today.

All of these studies are article length; two of them are described in more than one paper, giving a total of 22 publications. Two of these appeared in published conference proceedings, three in books, but academic journals are the outlet for all the others. What is perhaps surprising is the variety of different journals – twelve in total – which may contribute in part to a fragmented appearance, and indeed to the perception that little in the way of empirical research into DDL has been carried out – a frequent claim as we have seen. Some of the journals are predictable (Language Learning & Technology, Computer Assisted Language Learning and System account for seven papers between them), but one might have expected more than a single study each in the Journal of English for Academic Purposes and English for Specific Purposes.

Overall, six of the studies were conducted in Asia, six in North America, five in Europe, but only one each in Africa, Australia and the Middle East. Seven were carried out in inner-circle (Kachru 1992) English speaking countries (Australia, Canada, the UK and the USA), two in outer-circle countries (Hong Kong and Zimbabwe), but the majority (eleven) are in the expanding circle. Four of the studies in inner-circle countries included learners of mixed language backgrounds; the other three featured a single ESL population; in all other cases, the learners’ L1 was the language of the country where the study was conducted.

Most of the learners had relatively high levels of English language ability: ten studies involved advanced learners, four upper-intermediate, the others intermediate, though this should be taken as a rough guide only as the tests and definitions vary considerable. High levels are probably to be expected for logistical reasons, since all of the studies without exception took place in universities – eight of them with postgraduate students – as that is where researchers work and where many ESP studies are to be found. More speculatively, there is also a common perception that only relatively advanced level learners can usefully move on to ESP and successfully use corpora, despite findings from a number of empirical studies such as Boulton (2009a) which suggest otherwise.

The participants in four studies were majoring in English, in two cases using corpora for EAP, as well as one for ESP translation, and one for an option in law. Surprisingly, perhaps, only six of the studies are oriented towards a single discipline; this is in part because half of the studies featured students from mixed disciplinary backgrounds, with the focus thus on EAP as a common element. However, in seven of the studies focusing on academic writing, the students were able to use corpora directly for their own writing purposes rather than an imposed task. The disciplines covered are tremendously varied, from hard sciences (e.g. physics, biology, geology) to applied sciences (e.g. engineering, technology, business, computing) to human and social sciences (e.g. philosophy, history, law, tourism).

Not all the papers provide details concerning the corpora and software used. In some cases the corpus and software come as a package, especially for on-line corpora (such as the Collins
COBUILD Corpus Concordance Sampler,² the British National Corpus,³ and the corpora in the Compleat Lexical Tutor⁴, but also in MicroConcord, a popular concordancer in the 1990s (see Johns 1986). WordSmith Tools⁵ is the successor to this, cited explicitly in three studies; most other mentions are of in-house tools, though there is also occasional use of general web tools (AlltheWeb⁶ and Google Custom Search⁷). There is even greater diversity in the types of corpus used, both large (100m tokens) and small (2k), public and in-house, general and specialised, monolingual and parallel, written and spoken, fixed or with student input, and so on.

The focus in many studies is on academic vocabulary and writing, so it comes as no surprise that most of the studies are based on written corpora (though some of the larger public corpora also include spoken components: the BNC and COBUILD). Exceptions include the Michigan Corpus of Academic Spoken English⁸ in one study and 25k of spoken business reports and product reviews in another. However, no use is made of sound or multimodal corpora, and parallel corpora only on two occasions. Large public corpora do feature in several studies: the full BNC in one (100m words), COBUILD in three (45m words currently available), the Brown Corpus (1m words, available in the Compleat Lexical Tutor) in another. Such general corpora can satisfy a variety of needs, especially for EAP in courses for learners of mixed disciplinary backgrounds. Sometimes however sub-components are explored individually, such as BNC Written (1m words) or BNC Academic (2.7m). Some of the other existing corpora adopted are similarly general in nature, including TANGO (40m words) and the Independent newspaper (size not given), while others are more specialised: MICASE (1.7m words) and Hyland’s research article corpus (1.3m) (see Hyland 2005).

Many studies rely on smaller in-house corpora or create their own, sometimes from quite diverse sources, such as the 75k words in Aston (1997) comprising newspapers, the texts from MicroConcord, popular science articles from the New Scientist and simplified readers. More usually, however, in-house corpora tend to be built precisely so they can concentrate on very specific text types relevant to the participants’ needs. Students’ own textbooks constituted one popular source in the 1990s: several chapters of unspecified length in Stevens (1991); 7k words of economics, 22k words of geology, 33k words of philosophy for the three different populations in Mparutsa et al. (1991). Other ESP corpora include 35k words from 12 medical articles in Aston (1997); 1m words from six areas of business English (general business, management, marketing, law, management information systems, finance) in Curado Fuentes (2002); 800k words from 114 legal cases in Hafner and Candlin (2007); 350k of academic articles in Park and Kinginger (2010); and 40k words of tourist advertisements in Curado Fuentes (2007).

In several cases students had the opportunity to work on texts they chose themselves, especially where the group included participants from different disciplinary backgrounds, thus allowing them to focus on their own speciality. In these cases the Internet is the obvious source of texts, usually collected whole but occasionally as far smaller extracts: AlltheWeb for a set of concordances in Todd (2001), examples of usage for collaborative WordNet entries in Horst and Cobb (2001) and for another collaborative writing aid in Sun (2007). In Lee and Swales (2006)
and Turnbull and Burston (1998), students used their own written productions as the basis for corpus work.

2.2. Research design
The size of the population under study varies from the small case study with only one learner in Park and Kinginger (2010) to a large scale study of 300 students in Hafner and Candlin (2007). However, often such larger studies feature a qualitative analysis of a smaller number of students, especially through interviews, observation or analysis of written output. Population size varies throughout some studies and is not given in others; where the information is available, the average number of participants is 45, allowing quite a rigorous statistical analysis in nine cases; raw figures and percentages are given in a further six studies not backed up by statistical analysis; only five rely exclusively on qualitative data with no figures given at all. The duration of the programmes is highly variable, from only one or two sessions to a full academic year, most being measured in hours or weeks. Clearly this depends on the objective, with the shorter time span often reflecting experimental design to capture learning of specific items, longer studies being more qualitative in nature to gather insights from complete courses.

The quantitative studies frequently featured control groups or control items (5 cases) or (pre- and post-) tests for given language items (7). Some looked at what the learners actually did, mainly by automatically tracking their searches (4) or by observation (3); others were interested in uptake in various types of writing or self-correction (7). More often, however, the focus was on learners’ perceptions of what they did and how they felt about corpus use. In five cases this was built into the final task, where learners (individually or in groups) were required to present spoken or written reports; but the most frequent tools were questionnaires (8), individual interviews (9) and class discussions (6). Other occasional tools include written diaries, blog postings, email surveys and ‘informal’ feedback (2).

The tools chosen tended to depend on what the research was designed to find out: only seven studies attempted to evaluate ‘learning’ as such, in other cases focusing more on what learners did with a corpus and how they reacted to it. Only Stevens (1991) made use of paper-based materials in the main experiment design; four studies presented specially-designed CALL programs for the students, and 15 provided direct access to the corpora on computer, thus reflecting the explicit or implicit aim in the majority of cases to increase learner autonomy by preparing them for possible future corpus consultation (cf. Boulton 2010b).

The language focus was frequently “on the ‘collocational border’ between syntax and lexis, [where...] DDL methods seem to be most effective” (Johns 2002: 109). Two studies are interested in vocabulary from the Academic Word List (Coxhead 2002), others with vocabulary items specific to the students’ own field, whether imposed or chosen by the students themselves. The focus is generally on use in context, including collocations, clusters, and compounds, rarely on meaning as such. Though the analysis was occasionally of spoken performance, only Curado Fuentes (2004) seems particularly concerned to ensure the use of spoken corpora for this purpose. Three papers are concerned with reading, but production is a
3. Qualitative discussion of individual studies

The discussion so far can only bring out a certain number of typical characteristics of the studies, but for more detailed understanding it is necessary to describe each in rather more detail. This section divides the studies into two groups, the first with the main emphasis on vocabulary, the second on writing.

3.1. Vocabulary studies

The first papers to evaluate some aspect of ESP in classroom use appeared in Johns and King (1991). The first is by Stevens (1991), who was interested in using multiple gapped concordance lines for testing purposes. To this end, he created a small corpus of physics textbook chapters for 54 first-year science students in Oman. Following normal classes, students were tested on vocabulary either in a traditional text gap-fill or by a set of four gapped concordance lines (all on paper). Although training was minimal, the students performed significantly better on the concordance test in the second session, suggesting that truncated concordance lines are not in themselves an insurmountable problem (despite later findings to the contrary by e.g. Cheng et al. 2003; Yoon & Hirvela 2004), and that four short authentic contexts are more amenable than a longer stretch of discourse designed specially for the purpose.

Another early study was geared towards student reactions to using a corpus to help with reading. Mparutsa, Love and Morrison (1991) each contributed a case study description of using a corpus compiled from textbooks for their Zimbabwean students in economics, geology and philosophy (21k, 33k and 7k respectively). The studies were conducted in difficult conditions with very early tools and only four computers for up to 27 students who had never worked with computers before. The initial sessions were extremely controlled – “instructions were very much of the ‘press this key’ type” (p. 120) – to independent work in small groups by only the third session; this suggests that substantial ICT skills are not necessary for successful corpus work, as long as students are receptive (cf. Bernardini, 2002). The focus was mainly on reading and conceptual meaning of specialist terminology, with students in all three classes showing remarkable enthusiasm and claiming greater confidence, for a more active and reflective involvement in their reading by the end.

Aston (1997) also wanted to investigate how students got to grips with electronic corpora, here in the context of a translation course. He began with a general corpus of 75k words comprising the texts from MicroConcord and simplified readers along with popular science articles from New Scientist for his translation students in Italy. First use suggested that learners seek to find generalised language rules in corpus data, but do not pick up on patterns and tendencies, from which it is concluded that “it was clearly pedagogically inadequate just to tell learners to look for regularities while browsing, without any basis for deciding what regularities to look for, and without any clear motivation for doing so” (p. 208-209). Medical articles were added in the second stage (12 articles, 35k words), and students encouraged to list three specific questions.
of their own to explore. Working in small groups, they successfully pursued these questions in browsing mode, reported their findings clearly to others, and derived procedures for tackling texts in the future.

The next group of studies attempt to evaluate vocabulary learning quantitatively. Cobb and colleagues (Horst & Cobb 2001; Horst et al. 2001; see also Gaskell & Cobb 2004) report an experiment with learners from different disciplines needing to substantially improve their reading and vocabulary for EAP in Canada, the goal for the semester being to learn 650 new words. The 33 intermediate-level students were required to use dictionaries and a concordancer (mainly the 1m-word Brown Corpus via LexTutor) to help learn sets from the University Word List (Nation 1990); they also had to choose domain-specific or subtechnical vocabulary in groups, and read two ‘academic’ texts in order to select more words every week for inclusion in a collaborative WordBank. The definitions and examples in the latter were deemed generally very satisfactory, and the vocabulary gains significant – but generally small, which the authors attribute to problems in using a general vocabulary pre/post-test design to measure learning of specific items. A post-course questionnaire showed that concordancers were used less frequently than bilingual, monolingual or computer dictionaries, but that concordancing was the strongest predictor of vocabulary learning.

An eight-week course on AWL vocabulary was developed for 25 third-year English majors in Taiwan, presented in Lin (2008). The tool (TANGO) is based on a parallel corpus of 40m words from the English/Chinese cultural magazine Sinorama, more fully described in other papers not included here as they do not have a specifically ESP orientation (Chan & Liou 2005; Liou et al. 2006; Yeh et al. 2007). In each lesson the target words were first presented explicitly (in the final lessons by peer-teaching) and then in context; students then had to access to an on-line Moodle resource incorporating concordances to explore collocates. Lin encountered a similar problem to Horst and Cobb (2001) above in using a general vocabulary pre-test: very high scores did not allow for breadth of knowledge to increase, but depth improved significantly in the post-test. A VocabProfile analysis of student essays showed a substantial increase in productive use of the target items, declining only slightly in the final delayed post-test four weeks later; questionnaires also showed fairly positive attitudes to the course overall.

Kaur and Hegelheimer (2005) also attempt not only to quantify AWL vocabulary learning, but also to examine productive uptake of target items in student academic writing. 18 ESL students in the US were introduced to concordancing over two weeks and then completed vocabulary exercises on the target items (cloze and sentence-building) in class, and a writing exercise out of class. The control group was allowed to use an on-line dictionary, while the experimental group additionally had access to the BNC written section using LexTutor. The post-test results show the concordance group performing better, though not significantly so, with no apparent correlation between results and concordance use. The concordance group still used the dictionaries more frequently (cf. Horst & Cobb 2001), perhaps because the time allowed for the test was insufficient for the less familiar corpus tools. On the other hand, they used the target items significantly more frequently and more accurately in the written assignment.
Cresswell (2007) had similar aims in investigating how a general corpus, the Independent newspaper, could be used for Italian students majoring in English. The 126 participants needed help with academic writing, the experimental group being introduced to corpora to investigate the meaning and usage of various connectors. About half attempted an inductive approach, starting with the data and comparing back to traditional references; the others started with the references in a deductive approach. Learners of both types (especially inductive) generally succeeded in the task of formulating rules of use, with some qualification, although their findings proved difficult to convey to others. Furthermore, the overt knowledge was not found to translate well into use, as the experimental group performed only very slightly better than the control group on use of connectors in essays.

Curado Fuentes provides three independent ESP studies in Spain, all featuring an experimental / control group design and the emphasis also on productive use. In the first (2002, 2003), ten third-year business science students were trained in using a local 1m-word business corpus over two weeks, concentrating on clusters, collocations and compounds. The data were obtained from videos of five-minute oral presentations on a business topic, which showed the experimental group making rather more mistakes than the ten control students, but also considerably more effective use of the target points. A follow-up survey suggests the learners found the corpus tools more useful for semi-technical than technical items, grammar or discourse elements.

The next study (Curado Fuentes 2004) used a local ‘professional’ corpus of 25k words taken from spoken business reports and product reviews from the Internet, and an equivalent ‘academic’ corpus comprising discussions on socioeconomic topics from MICASE. These were explored by third-year tourism students to help work on both an individual oral report (prepared, but delivered without notes) and a group discussion on an unprepared topic. The experimental group made considerably more effective use of the target items from the corpus in Task A, as well as more effective marking elements in Task 2; they also showed “greater confidence than the control group in the spontaneous speech task” (p. 22). As in the previous study, the experimental group apparently made more errors, but these can be attributed to their faster pace of delivery and longer presentations.

Tourism students also participated in Curado Fuentes’ third study (2007), spending three hours exploring register with six tourist advertisements before using a concordancer for two hours on the rest of the 60 adverts (40k words). The set tasks were mainly for register and text type, vocabulary and constructions. In the reading comprehension post-test, the experimental group performed significantly better on all five question types in the post-test than the control group, who had worked with text book and texts in a traditional manner.

3.2. Writing studies
Park and Kinginger (2010) used corpus queries to provide insights into the processes involved in (ESP) writing, triangulating the query logs with recordings of the computer screen, along with

retrospective accounts by the single participant as she reviewed these and commented on her thoughts. The student was an advanced Chinese learner enrolled in an academic writing course as part of her first year of business studies in the US, using a 350k-word corpus comprising academic articles, accessed using Google CSE. 194 searches were conducted in 109 minutes for 118 ‘transactions’ (i.e. individual problems) which can be described as interrelated narratives; 26 were complex, featuring clusters of related points. The analysis showed that planning, writing and revising occur simultaneously and not in discrete steps.

The rest of the studies in this section allow the participants to use the corpora and techniques on their own real writing tasks, for reasons given succinctly by Hafner and Candlin (2007: 308): “it was felt that the corpus tools would best serve students as reference tools in the ordinary course of their legal writing and drafting.” In this way, the process becomes “authentically heuristic” (Kaur & Hegelheimer 2005: 289), even in the many cases where the emphasis is on ‘errors’ and other “genuine problems frequently encountered during the process of writing and revising” (Yoon & Hirvela 2004: 265).

One of the more controlled environments is provided by Sun (2007), who presents the Scholarly Writing Templates program developed for postgraduate students in Taiwan who needed to write research articles in English. The 20 participants each chose papers to include from their own specialist field, inputting items related to structural moves along with the wider context, then using the SWT to help with both the language and structure of articles via corpus searches. Learning outcomes were not examined in detail, although the paper reports more effective writing overall. Tracking showed that lower levels tended to use the template more but less strategically, and were more accepting of proposals for both article structure and language use, while more advanced learners were more strategic and critical in their use of the tool. The feedback was generally positive, especially among those currently writing papers and particularly for information structure, though previous publishing experience was not seen to affect use.

Error-correction was the main focus in Gilmore (2009), who used general corpora – the BNC and COBUILD – to help 45 second-year Japanese students from mixed disciplinary backgrounds needing English for EAP writing. They first received a 30-minute introduction to corpora, then spent an hour browsing them on their own before using them out of class to improve an earlier written assignment with errors highlighted. Four native-speaker reviewers then rated each version, and found the new version more ‘natural’ in 61% of cases, equivalent in 33%, less natural in only 6%. Questionnaire responses showed 95% of students found the corpora useful, generally preferring the COBUILD as more user-friendly, though some appreciated the larger size of the BNC.

Todd (2001) was also interested in error-correction: specifically, whether students could use a general search engine (AlltheWeb) to generate concordances from the Internet and induce rules, the searches focusing on two errors highlighted in their own work. The 23 postgraduate students were working in science and engineering at a technological university in Thailand. In 20
out of 23 cases their rules matched the concordances, and most also matched traditional reference sources, resulting in 18 valid corrections. Items with many meanings or patterns of usage were found to be the most difficult; more surprisingly, adjectives were found easier than verbs, in turn easier than nouns. Overall, the study lends strong support to learners’ ability to induce patterns from self-generated concordances.

The final study explicitly on error-correction comes from Turnbull and Burston (1998), who offer a detailed case study of two students enrolled in master’s degrees in Australia, one in applied Japanese linguistics, one in public policy and management. The corpora involved comprised the students’ own texts annotated for errors, but their radically different motivations and learning styles meant that one student (field independent in her “ability to focus on the similarities and differences of linguistic patterns in the data” [n.p.]) used the concordancer frequently and effectively and felt she learned a great deal, while the other (field dependent with “difficulty in identifying and analysing linguistic detail in the data” [n.p.]) showed less interest and was less successful, and found it a waste of time.

Similar detail is offered by Yoon and Hirvela (2004), who interviewed one receptive and one unreceptive student in each of two groups of differing levels of language ability, in addition to a detailed questionnaire for all students and feedback during class. Following an extensive four-week introduction to corpus techniques using COBUILD, 22 undergraduate and post-graduate learners in this study mainly worked in their own time on their own writing. Learners were generally positive, the overwhelming majority claiming they would use corpora in the future. They found corpora most useful for lexical usage and phrases (preferring dictionaries for meaning), as well as writing (the skill most focused on in class). The intermediate learners seemed more favourable than the advanced group and reported using corpora in other writing assignments, perhaps as they had received more guidance (which the authors consider crucial here). The study also reports some negative reactions, mostly of frustration with the technology, the time taken, repeated searches for the same items, and difficulty of interpreting truncated concordances.

Other individual variables were identified as important in a study by Yoon (2008), the most important of which concerned past and on-going writing experiences among postgraduate students and practising researchers. The 10-week EAP course at an American university was designed for Korean or Chinese L1 students in education, natural resources, aerospace engineering, history, nuclear engineering, molecular genetics. COBUILD was introduced to help the participants with their writing or checking their own productions out of class, with the teacher subsequently preparing handouts based on their emailed reports. A variety of tools was used to assess the behaviour and reactions of 6 of the 14 students, who reported increased confidence and autonomy in writing and improved writing procedures, as well as increased language awareness especially for checking existing knowledge. Corpus consultation was mainly perceived as a useful additional technique, more favourably received as time went on, although the paper reports mixed success depending on a variety of individual variables.
The only parallel corpus to be used for writing here was built by Fan and Xunfeng (2002), comprising 300k words of English and 500k characters of Chinese – legal texts, government documents, public speeches, minutes of meetings, annual reports and press releases, all taken from the Internet. Students majoring in English in an optional law course in Hong Kong could thus compare translations or click for concordances of key items. Questionnaires and videoed interviews showed positive reactions, with all but one of the 21 students finding the concordances and bilingual hyperlinks at least moderately useful. Tests of comprehension on two texts revealed mixed results, however, indicating that bilingual corpora present their own difficulties.

Another law corpus was built by Hafner and Candlin (2007) for a course on legal writing. Faculty suggested texts for the corpus, which consisted of 114 legal cases for a total of 800k, backed up by on-line tutorials and tools. Two cohorts of around 150 law students received a short introduction to corpus use, followed by work on their own writing tasks. 21% in the first cohort and 40% in the second used the corpus on their own initiative, but the main data were obtained from interviews with nine participants, all advanced learners and regular computer users. Initially four were identified as ‘adopters’ (and five non-adopters) as they used the corpus appropriately to help with their writing, but with no continued training they tended to abandon the new techniques (e.g. concordancing) and resort to practices familiar in their law studies, focusing on full texts for information search. This is seen as a problem, though Aston (1997: 210) suggests otherwise, specifically recommending that his students approach the corpus with “non-linguistic goals.”

Lee and Swales (2006) provide a detailed plan of their 15-week course to help Chinese research students in the US write academic papers. Work covered a number of language points decided by the teachers, using MlCASE (1.7m words), Hyland’s research article corpus (1.3m) and the BNC Academic (2.7m). The participants then compiled corpora of their own writing and texts in their field, and had freedom of choice in the final project, presented in the form of a conference paper. Those who attended on a regular basis coped well with the approach, and appreciated that the tools could increase their autonomy in working with language specific to their needs; this is reflected in the fact that most of them bought WordSmith Tools for their own future use.

4. Discussion

Having looked at the 20 studies, it is necessary to take a step back to see what they tell us about corpus use by second or foreign language users within the context of ESP as a whole. This is particularly difficult as one of the most striking things to emerge is the diversity of the various studies, in terms of research designs and questions, corpora and tools, aims and implementations, which inevitably makes a proper meta-analysis impossible. However, a number of patterns do begin to emerge, and it is the role of this section to relate these to other studies in the field.
4.1. Short-term and long-term benefits

Overall, it seems that the participants do manage to deal with corpus data quite successfully; in those studies that asked for feedback from the students, the reactions are generally positive – sometimes extremely so (Gilmore 2009) – though as we shall see, this depends on individual students. Inevitably, a number of problems do arise; these mostly concern technical aspects such as delays, crashes and going off-line, in addition to some problems of knowing what to look for, how to formulate a query, interpreting the results, and how to do this quickly and efficiently (e.g. Yoon and Hirvela 2004), while avoiding “concordancing burnout” (Lee & Swales 2006: 57). However, none of these problems are specific to ESP, and have been described elsewhere (see e.g. Boulton 2009a). In Bernardini’s (2001: 243) experience, “the difficulties should not be overestimated; learners should quickly acquire the skills needed.” Sinclair (2003) provides a comprehensive guide to all of this, the contents of which can be used by teachers to help their students. Specifically regarding reading concordances, most studies pass over this in silence, suggesting that it is not a major problem, the exception being (Yoon & Hirvela 2004). Stevens (1991) looked at this explicitly, and found that multiple gapped concordances were more effective than a longer stretch of cohesive gapped text, even after only a brief introduction to their use.

Comparatively few studies attempt a quantitative analysis of the effectiveness of DDL in terms of learning outcomes; and all those that do, focus on vocabulary and usage. Lin (2008) and Horst and Cobb (2001) both found significantly better uptake of vocabulary from the experimental group, but still less than expected; in both cases they attribute this to the generic vocabulary tests used which are insufficiently sensitive to specific items. Cresswell (2007) and Kaur and Hegelheimer (2005) also found an advantage for DDL, though not significantly so. Further, the latter study finds no correlation with corpus use, in contradiction with Horst and Cobb (2001) where concordancing is the best predictor of vocabulary learning. Some of the most encouraging findings in terms of learning come from Curado Fuentes (2007), where the DDL group perform significantly better than the control on all five test tasks for different types of language items.

At first sight, these results may not seem overwhelmingly positive. However, it seems likely that DDL will not be particular effective for learning large quantities of words, but rather for refining usage in context. Lin (2008) is a case in point, finding better results for depth than for breadth of vocabulary knowledge, in line with Cobb (1999). More relevant still is the effectiveness in actual use, whether from using the corpora as a learning aid or a reference resource. Cresswell (2007: 280) encountered “an observable (though slight) positive effect on actual use”, though most of the other studies are more encouraging in this respect. This of course makes the approach particularly useful for ESP, where the learners often have extensive referential knowledge of many technical words for their discipline – the difficulty lies in using them in context. Lin (2008) found substantial uptake of the target items in written work, especially in more complex and imaginative uses, in analysis of both immediate and delayed writing tasks, and the experimental group in Kaur and Hegelheimer (2005) likewise used the target items more frequently and more accurately in writing. Curado Fuentes (2003, 2004) is the only one to look for actual use in
spoken output, both planned and unplanned; he found the learners making slightly more errors, which he attributes to more ambitious productions (longer and faster), but in particularly using the target items significantly more often and more effectively.

Much of course depends not on immediate gains of target items but on subtle longer-term changes in behaviour. Native speakers rated Gilmore’s (2009) students as producing significantly more ‘natural’ language after using the corpus, suggesting more mature processes. Further evidence for this comes from Yoon (2008), with the students becoming more confident and autonomous in their work, resulting in more effective production; and also from Sun (2007: 339), where the “participants’ writing processes became more recursive and more reflective”, and their “writing strategies became more complex, more inductive, and more resourceful.” Where learners had the freedom to pursue their own goals, the almost unanimous verdict is that their sensitivity to language awareness is raised as they become more “confident” (Yoon 2008: 31) and more “empowered” (Lin 2008: 6) to “take control of their own learning” (Lee & Swales 2006: 72).

In most of the papers involving explicit error-correction, the errors are indicated by the teachers. However, many of the others involve ‘revision’ of the students’ writing, an integral part of the process (Park and Kinginger 2010) and one which includes the spontaneous correction of (unannotated) errors. In both types of case, it is clear that the vast majority of students are capable of transforming their questions into queries which they can usefully follow up in a corpus. Todd (2001) found students able to work out effective ‘rules’ from concordances they produced themselves and apply them to their own work. However, as Aston (1997: 208) points out, rules as such may not be best suited for corpus consultation. On a more metacognitive level, he required his students not only to formulate ESP translation questions and to look for patterns, but also to report back orally in such a way that the findings were of benefit to other students. They achieved these tasks well, in contrast with Cresswell (2007), also translation students in Italy, who managed to formulate rules more or less well, but did not manage to report back well to the rest of the group. In two other collaborative projects, the participants in Horst and Cobb (2001) chose their own items to work on, providing ‘very satisfactory’ examples and definitions in the on-line group word bank; and Sun’s (2007) students also successfully uploaded examples and usage notes to an on-line writing template.

Some students in Yoon and Hirvela (2004) already used the corpora in other courses, and the overwhelming majority of all students claimed they would continue to use them in the future. Results such as in Yoon (2008) are particularly encouraging as they found the participants to be increasingly favourable over time; although in several studies the introduction to corpus use is minimal (as little as 30 minutes in Gilmore 2009) the implication is that sustained usage may be more beneficial, but of course more longitudinal data are needed. The studies that asked the question found that many learners intended to continue using corpora after the end of the course, including in Lee and Swales (2006) where a number even invested money in the tools themselves. This is indeed one of the main advantages attributed to corpus use, that it can increase autonomy by providing the learner with a resource that can be used entirely
independently of teacher input. But for this to happen, the learners need to know how to use free, stable, accessible corpora and tools, especially those found on the Internet. For example, the students in Sun (2007) claim they will continue to use corpora in the future, but this depends on the Scholarly Writing Template being made available outside the course and off-campus.

4.2. Learner profiles
Kaur and Hegelheimer (2005) and Horst and Cobb (2001) both found that learners tended to use dictionaries in preference to corpus tools – unsurprisingly, as even master’s students in translation have been found to use dictionaries more frequently (Frankenberg-Garcia 2005). However, in the first this is attributed in part to the experiment design which encouraged use of familiar tools (dictionaries); and in the second study corpus use was found the best predictor of success overall. This might mean that concordancing is more effective, or perhaps that a particular type of (successful) learner is more likely to use concordancers. For example, Cresswell (2007) found some learners having difficulty coming to terms with an inductive approach (cf. also Lewis 2006), but concludes that DDL can be adapted to a deductive approach more appropriate for some learners. In Turnbull and Burston (1998), the field-independent learner was more successful and more favourable to corpus use, but learning styles in relation to corpus work represent almost virgin territory (see Boulton 2009b for a discussion and an experiment with DDL for general English).

Level of ability is another issue: though most studies here deal with advanced students, Yoon and Hirvela (2004) intriguingly found that the intermediate group appreciated the approach and used it more than the advanced group. However, a number of other factors come into play: the intermediate group had more training, and used the corpora less strategically. While a number of DDL experiments have been conducted with lower-level learners, more are needed in relation to ESP, especially as this reflects a potentially very large population. A further variable is identified by Yoon (2008), namely previous writing experience and current need. Finally, Hafner and Candlin (2007) also found results depended on individual learners: four of the nine adopted corpus queries easily; of the other five, two used other electronic tools, and three none at all (cf. the “adopters”, “minimal users” and “refusers” in Kaszubski 2008: 174). Though ICT literacy is apparently not an overriding barrier to DDL (witness the students in Mparutsa et al. (1991) who had never used a computer at all before), this inevitably depends on the complexity of the task at hand, perceived difficulty in relation to other students, and attitudes towards ICT in general – Bernardini (2002), Mukherjee (2004) and Seidlhofer (2000) are among those who report ‘technophobic’ students. The implication of all this is that DDL may not be appropriate for all students, and when implemented, the approach should be flexible to local cultures. Lee and Swales (2006: 62) adopt a “just in time” approach to cater for arising needs, and are surely right to warn that their course outline should not be adopted as it is for other populations.

5. Conclusion
The individual discussion divided the studies into two broad groups, those focusing on language and those focusing on skills, though inevitably the two overlap. The language focus is typically
fairly atomistic, concentrating on vocabulary usage and lexico-grammar, with comparatively little work on discourse, genre or text type (though see Curado Fuentes 2007; Sun 2007). As most of the researchers are also the teachers, this presumably reflects their experience of the kind of area where corpora are likely to be of most use: corpus queries can only focus on surface forms which are invisible in much grammar and discourse. With the exception of Curado Fuentes (2003, 2004), the skills focus is generally on writing, whether in production or revision (including error-correction), occasionally reading or translation. This no doubt reflects at least in part the fact that written corpora are easier to compile and more readily available, but it could also be that this is where corpora show their full potential: as a reference resource (as opposed to a learning tool) for ESP writing as they can provide far more information than any dictionary, specific to a given discipline or genre, thus providing relevant data for many highly individual queries. Nonetheless, there is considerable room for further work using spoken corpora, especially as speaking and listening are likely to be as important for many ESP learners as writing and reading.

Few of the learners surveyed here are majoring in languages; despite their sophistication in their chosen discipline, many of the others may have relatively little linguistic baggage with which to approach corpus work directly. While this will of course influence the choice of appropriate activities, it certainly does not exclude DDL outright – a point some studies are at pains to explore (e.g. Boulton 2010b). Even with the most advanced and sophisticated of learners, most of the researchers go out of their way to hedge their findings and are careful not to appear to proselytise (cf. Yoon & Hirvela 2004: 279), a prudent feature typical of much DDL research (e.g. Gilmore 2007). None would claim that DDL presents a panacea in ESP: it is not appropriate for all language questions, nor for all learners, and needs to be adapted to local conditions. Where it is successful, however, it is found overall to be highly motivating, as it fosters different cognitive skills and language awareness, promoting a learner-centred problem-solving approach to interacting with authentic language which is naturally inductive and constructivist, with all the potential for increasing autonomisation and thus life-long learning. While Johns (1993: 8) pointed out that DDL fitted the “zeitgeist” of the 1990s, all of these features are still very much “consistent with a variety of current principles in language learning theory” today (Conrad 2005: 402).

Several papers have lamented the difficulties that data-driven learning seems to be having in breaking into mainstream teaching – whether for ESP or foreign / second language teaching in general – and suggest a number of essential requirements, especially in the form of teacher training (e.g. Römer’s 2006 ‘wish list’). One particular feature which emerges here is the nature of the corpora and tools on offer. For DDL to have major impact on ESP, learners will need to have access to (and know how to use) simple, free, stable and accessible tools. These will of course include large principled corpora such as the *Corpus of Contemporary American English*, as well as tools to help users search their own carefully compiled corpora (e.g. *AntConc*). To this we might add a number of more “quick and dirty” solutions, with tools like *SketchEngine* allowing users to create highly specific corpora for their own immediate ends almost instantaneously, variously referred to as quick ‘n’ dirty, ad hoc, disposable, or ephemeral

corpora (cf. Corpas Pastor & Seghiri 2009). It will also no doubt develop to the use of the internet as a ‘corpus’ requiring no downloading, with tools such as WebCorp\textsuperscript{12} able to search specific sites or domains and provide KWIC output and collocation information. As these and other tools become more user-friendly and more widely known, the conditions will be more favourable for the spread of DDL techniques among ESP learners and other users.

Notes

References


Maniez, F. This volume. A corpus-based study of adjectival vs nominal modification in medical English.


This is a pre-publication version. For the version of record, please see DOI: 10.1075/scl.52.01bou (or email me at alex.boulton@atilf.fr)


**Appendix A: 20 empirical DDL studies in ESP**

Key:
- **Study**: the paper(s) reporting each study.
- **Country (L1)**: the country where the study was conducted, along with the mother tongue of the majority of participants.
- **Level**: the level of current studies as well as the level of proficiency in English – from (lower or upper) intermediate to advanced.
- **Speciality**: the participants’ major discipline, and whether the study focuses on English for specific or academic purposes.
- **Participants**: the number involved (including any control group, indicated in brackets).
- **Duration**: the time spent hours, weeks, semesters, etc.
- **Corpora / software**: source and size of the corpora used as well as the software, where the information is available; *own* = locally-built corpus.
- **Interface**: whether the participants used a hands-on concordancer, a CALL program including corpus data, or paper-based materials.
- **RQ (research question)**: whether the study is concerned with evaluating learners’ behaviour (B), attitudes (A), using corpora as a reference resource (R), or learning outcomes (L).
- **Aim**: the main point under study – usually a language item or a skill.
- **Research instruments**: the main tools used for data collection.
- **Data**: whether some kind of statistical analysis (S) is provided, or only raw figures and percentages, or no quantitative analysis at all (0).
<table>
<thead>
<tr>
<th>Study</th>
<th>Country (L1)</th>
<th>Level</th>
<th>Speciality</th>
<th>Participants</th>
<th>Duration</th>
<th>Corpora / Software</th>
<th>Interface</th>
<th>RQ</th>
<th>Aim</th>
<th>Research Instruments</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aston 1997</td>
<td>Italy (Italian)</td>
<td>uni (adv)</td>
<td>translation (ESP)</td>
<td>?</td>
<td>?</td>
<td>newspapers (75K, inc New Scientist), medical articles (35K, 12 articles), MicroConcord</td>
<td>hands-on</td>
<td>B</td>
<td>translation</td>
<td>class discussions, observation</td>
<td>0</td>
</tr>
<tr>
<td>Cresswell 2007</td>
<td>Italy (Italian)</td>
<td>uni 3 (adv)</td>
<td>English (EAP)</td>
<td>126 (c=65)</td>
<td>1 sem?</td>
<td>Independent newspaper, WordSmith</td>
<td>hands-on</td>
<td>B L</td>
<td>connectors</td>
<td>post tests, control population, informal feedback, interviews, oral presentations, writing</td>
<td>S</td>
</tr>
<tr>
<td>Curado Fuentes 2003 (also 2002)</td>
<td>Spain (Spanish)</td>
<td>uni 3 (int)</td>
<td>business (ESP)</td>
<td>20 (c=10)</td>
<td>2 wks</td>
<td>own: 1m business</td>
<td>hands-on</td>
<td>L</td>
<td>collocations, clusters, compounds</td>
<td>post tests (oral presentations), control population, e-mail surveys</td>
<td>raw n’s</td>
</tr>
<tr>
<td>Curado Fuentes 2004</td>
<td>Spain (Spanish)</td>
<td>uni 3 (int +)</td>
<td>tourism (ESP)</td>
<td>20 (c=10)</td>
<td>2 wks (5h+)</td>
<td>own (spoken): professional &amp; academic (50K), WordSmith</td>
<td>hands-on</td>
<td>L</td>
<td>(semi-) technical language in spoken use</td>
<td>oral reports, class discussions</td>
<td>raw n’s</td>
</tr>
<tr>
<td>Curado Fuentes 2007</td>
<td>Spain (Spanish)</td>
<td>uni (int +)</td>
<td>tourism (ESP)</td>
<td>20 (c=10)</td>
<td>5h</td>
<td>own: tourism (40K), WordSmith</td>
<td>hands-on</td>
<td>L</td>
<td>reading, genre awareness</td>
<td>post tests, control population, e-mail tasks</td>
<td>S</td>
</tr>
<tr>
<td>Fan &amp; Xunfeng 2002</td>
<td>Hong Kong (Chinese)</td>
<td>uni 3 (adv)</td>
<td>English (ESP)</td>
<td>21</td>
<td>1h40</td>
<td>own: parallel legal texts (300K English / 500K characters Chinese)</td>
<td>program</td>
<td>A R</td>
<td>reading</td>
<td>questionnaires, class discussions</td>
<td>raw n’s</td>
</tr>
<tr>
<td>Gilmore 2009</td>
<td>Japan (Japanese)</td>
<td>uni 2 (int)</td>
<td>mixed (EAP)</td>
<td>45</td>
<td>1h30</td>
<td>COBUILD sampler (45m?) &amp; BNC (100m)</td>
<td>hands-on</td>
<td>R A</td>
<td>self-correction</td>
<td>self-correction, questionnaires</td>
<td>raw n’s</td>
</tr>
<tr>
<td>Hafner &amp; Candlin 2007</td>
<td>Hong Kong (Chinese)</td>
<td>uni 4 (adv)</td>
<td>law (ESP)</td>
<td>300 (9 case studies)</td>
<td>1 yr</td>
<td>own: legal texts (800K, 114 legal cases)</td>
<td>hands-on</td>
<td>B A</td>
<td>writing</td>
<td>tracking, interviews</td>
<td>raw n’s</td>
</tr>
<tr>
<td>Study</td>
<td>Country (L1)</td>
<td>Level</td>
<td>Speciality</td>
<td>Participants</td>
<td>Duration</td>
<td>Corpora / Software</td>
<td>Interface</td>
<td>RQ</td>
<td>Aim</td>
<td>Research Instruments</td>
<td>Data</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------</td>
<td>-------</td>
<td>------------</td>
<td>--------------</td>
<td>----------</td>
<td>---------------------</td>
<td>-----------</td>
<td>----</td>
<td>-------------------------------</td>
<td>---------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Horst &amp; Cobb 2001 (also Horst et al. 2001)</td>
<td>Canada (mixed)</td>
<td>uni (int)</td>
<td>mixed (EAP)</td>
<td>33</td>
<td>12 wks</td>
<td>Brown (1m), LexTutor; S-created (WordNet)</td>
<td>program B</td>
<td>L</td>
<td>reading, vocabulary</td>
<td>post tests, control items, rating WordBank entries, questionnaires</td>
<td>S</td>
</tr>
<tr>
<td>Kaur &amp; Hegelheimer 2005</td>
<td>USA (mixed)</td>
<td>uni (int)</td>
<td>mixed (EAP)</td>
<td>18 (c=9?)</td>
<td>4 wks?</td>
<td>BNC written (1m), LexTutor</td>
<td>hands-on</td>
<td>R</td>
<td>L</td>
<td>AWL vocabulary</td>
<td>post tests, control population, writing</td>
</tr>
<tr>
<td>Lee &amp; Swales 2006</td>
<td>USA (Chinese)</td>
<td>uni 5+ (adv)</td>
<td>mixed (EAP)</td>
<td>4</td>
<td>13 wks (20h +)</td>
<td>Hyland’s research article corpus (1.3m), MICASE (1.7m), BNC academic (2.7m), students’ choice, WordSmith</td>
<td>hands-on</td>
<td>A</td>
<td>writing</td>
<td>interviews, class discussions, project reports</td>
<td>O</td>
</tr>
<tr>
<td>Lin 2008</td>
<td>Taiwan (Chinese)</td>
<td>uni 3 (adv)</td>
<td>English (EAP)</td>
<td>25</td>
<td>8 wks (13h20)</td>
<td>TANGO (40m-word Chinese-English, cultural magazine Sinorama, AWL Highlighter (Moodle)</td>
<td>hands-on</td>
<td>A</td>
<td>L</td>
<td>AWL vocabulary</td>
<td>pre+post tests, delayed writing, questionnaires</td>
</tr>
<tr>
<td>Mparutsa et al. 1991</td>
<td>Zimbabwe (Shona?)</td>
<td>uni (int +)</td>
<td>a) economics b) geology c) philosophy (ESP)</td>
<td>variable</td>
<td>variable</td>
<td>own: subject textbooks (7K, 22K, 33K), MicroConcord</td>
<td>hands-on</td>
<td>B</td>
<td>A</td>
<td>vocabulary, reading</td>
<td>informal feedback, questionnaires, class discussions, short reports</td>
</tr>
<tr>
<td>Park &amp; Kinginger 2010</td>
<td>US (Chinese)</td>
<td>uni 1 (adv)</td>
<td>business (EAP)</td>
<td>1</td>
<td>1h49</td>
<td>academic articles (350K), Google CSE</td>
<td>hands-on</td>
<td>B</td>
<td>writing</td>
<td>tracking, screen recordings, retrospection</td>
<td>raw n°s</td>
</tr>
<tr>
<td>Stevens 1991</td>
<td>Oman (Arabic)</td>
<td>uni 1 (int ?)</td>
<td>physics (ESP)</td>
<td>54</td>
<td>2 sessions</td>
<td>own: science textbooks (several chapters)</td>
<td>paper</td>
<td>B</td>
<td>vocabulary</td>
<td>post tests</td>
<td>S</td>
</tr>
<tr>
<td>Sun 2007</td>
<td>Taiwan (Chinese)</td>
<td>uni 4+ (adv ?)</td>
<td>mixed (EAP)</td>
<td>20 (3 interviews)</td>
<td>?</td>
<td>own: Scholarly Writing Template, students’ choice</td>
<td>program B</td>
<td>A</td>
<td>writing</td>
<td>tracking, questionnaires, interviews</td>
<td>S</td>
</tr>
</tbody>
</table>

This is a pre-publication version. For the version of record, please see DOI: 10.1075/scl.52.01bou
(or email me at alex.boulton@atilf.fr)
| Study                                | Country   | Location | Grade | Type   | Duration | Materials          | Methodology                  | Notes                                                                 |
|--------------------------------------|-----------|----------|-------|--------|-----------|--------------------|-----------------------------| Adam                         |
| Turnbull & Burston 1998              | Australia | uni 4+   | mixed | (EAP)  | 2         | 8 mths             | learners' own texts (2K),   | observations, interviews, questionnaires, class discussions, written evaluations |
| Yoon 2008                            | USA       | (Korean) | uni 4 | (adv)  | 6         | 10 wks (50h?)     | COBUILD sampler (45m?)      | observations, interviews, recall, tracking, assignments, reflection    |
| Yoon & Hirvela 2004                  | USA       | (mixed)  | uni   | 1/4+   | 22        | 20 wks (90h?)     | COBUILD sampler (45m?)      | questionnaires, interviews, observations, interviews                   |

This is a pre-publication version. For the version of record, please see DOI: 10.1075/scl.52.01bou (or email me at alex.boulton@atilf.fr)