

HEA/JISC Open Educational Resources case study: pedagogical development from OER practice

Openness in English for Academic Purposes

Alannah Fitzgerald

Durham University

Theme: OER and teaching quality

1. Outline	3
1.1 The opening landscape in formal EAP	3
1.2 The massification of higher education	4
Keywords	5
2. What are language corpora anyway?	5
2.1 Data-Driven Learning	6
3. Cascading openness in English for Academic Purposes	6
3.1 Openness and sharing on a shoestring	6
3.2 Wider motivations for developing competencies in openness for EAP	7
4. Open corpus-based resources for EAP	9
4.1 The Compleat Lexical Tutor (Lextutor)	9
4.1.1 The VocabProfile tool and the Academic Word List (AWL) in Lextutor	9
4.2 FLAX – the Flexible Language Acquisition project	11
4.2.1 What are collocations and why are they so important?	11
4.2.2 The Learning Collocations collection in FLAX	12
4.2.3 The Web Phrases and Web Collocations collections in FLAX	13
4.2.4 Designing Data Driven Learning resources for the masses	13
4.3 AntConc	14
4.3.1 DIY corpora with AntConc	14
4.3.2 File conversion issues in corpus building	14
4.4 WordandPhrase	15
5. Learning from open corpus-based resources for EAP	16
6. New directions	17
7. References	18
8. License and citation	19

1. Outline

Commissioned by the Higher Education Academy (HEA) in the United Kingdom, this case study has been written as an introductory guide for teachers and researchers working with international and home students for whom it would be beneficial to develop competencies with academic English as it is used across the disciplines. In particular, as per the HEA directives handed down in commissioning this case study, Open Educational Resources (OER) and teaching quality are the theme to be addressed in this report with respects to the open web-based tools, resources and practices in English for Academic Purposes (EAP) that will be introduced here.

A range of open corpus-based tools, resources and techniques will be demonstrated and discussed in section four, looking at different language projects from around the world that provide free access to valuable English language resources which are relevant for use both within and beyond traditional higher education. Because these tools and resources are openly available they can be used and shared by learners and teachers across a variety of contexts. For example, in language schools and in university language support centres, in open and distance education, and in independent or informal learning. They range from tools and resources that can provide diagnostic help for improving vocabulary, reading and writing to resources that can assist with identifying, retrieving, storing and managing useful words and phrases as they occur across a variety of authentic academic and general English language contexts.

Findings and resources will also be shared in section three, based on an OER cascade study that was carried out with EAP teachers and students at Durham University English Language Centre (DUELC). As part of the TOETOE project (TOETOE: Technology for Open English – Toying with Open E-resources), three corpus-based projects - [FLAX](#), the [Lextutor](#), and [AntConc](#) - were trialed for their efficacy in mainstream EAP teaching and learning practice. A fourth corpus-based project, [WordandPhrase](#), was introduced at one of the project dissemination events and will also be introduced here in this case study. None of the participants in the study had received any prior training with corpus-based resources for Data-Driven Learning (DDL) in language education. Initial findings from this study at Durham University on the design and usability of corpus-based resources have informed on-going research and development work with the TOETOE project in collaboration with the open-source FLAX project at the University of Waikato in New Zealand.

TOETOE began as a HEFCE-funded project (Higher Education Funding Council of England) managed by the Support Centre for Open Resources in Education (SCORE) at the Open University in conjunction with DUELC to explore the potential for openness in EAP resources and practices development. TOETOE entered a new phase in October 2012 with the University of Oxford as TOETOE International for the re-use of Oxford managed corpora and Oxford created OER in partnership with FLAX to make these resources more linguistically accessible to international audiences. There are three TOETOE case studies in this OER series with the HEA and the Joint Information Systems Committee (JISC). For a wider view on the development of the TOETOE project in collaboration with Durham University, the Open University and the University of Oxford, please refer to the following two case studies within this series:

Fitzgerald, A. (2013). TOETOE International: FLAX Weaving with Oxford Open Educational Resources. Open Educational Resources International Case Study. Commissioned by the Higher Education Academy (HEA) and the Joint Information Systems Committee (JISC), United Kingdom.

Comas-Quinn, A. & Fitzgerald, A. (2013). Open Educational Resources in Language Teaching and Learning. Open Educational Resources Case Study: Pedagogical development from OER practice. Commissioned by the Higher Education Academy (HEA) and the Joint Information Systems Committee (JISC), United Kingdom.

1.1 The opening landscape in formal EAP

To provide some context for this case study it is useful to reflect on the different types of open educational resources and practices in EAP which have been gathering momentum in terms of their online presence in parallel with the development of the TOETOE project. Until recently the sharing of EAP resources and practices in the UK had occurred mostly in face-to-face mode, either locally as part of continuing professional development in the workplace or nationally at [BALEAP](#) professional meetings and conferences for those EAP practitioners who were encouraged and

able to attend. Recent times have evidenced a steady opening up of practices for sharing expertise and resources in EAP with the EAP practitioner voice taking centre stage. For example, the [EAP Teaching blog](#).

Well-known online EAP resources that are freely available include those from dedicated websites such as [Using English for Academic Purposes](#) (UEfAP) by Andy Gillett and the [Online Writing Lab](#) (OWL) at Purdue University. By no means an exhaustive list, these popular online EAP resources are what is known as open gratis resources – they are free to access but their copyright remains All Rights Reserved. In recent years, with the UK government-funded OER programmes in higher education and OER mandates in the US, there has been a noticeable mushrooming of OER repositories and peer-to-peer online communities that encourage the sharing of open libre resources for teaching and learning; these are still copyright resources that have been published under Creative Commons licences, making them free to reuse and re-distribute, and in many cases re-purpose and re-mix. Open repositories based in the UK that are relevant to EAP include: the [HumBox](#), the [Language Box](#) and [LORO](#). Across the Atlantic with Joe Moxley at the University of South Florida is [Writing Commons](#): “a free, global, peer-reviewed, open-education resource for college-level writers, college faculty, and the everyday writer.” (Moxley, n.d.)

The [IRIS](#) project encourages the open sharing of research data and data collection instruments used for research into second language teaching and learning. Journal publications that are central to EAP remain subscription-based, however. These include the Journal of English for Academic Purposes (JEAP) and the English for Specific Purposes Journal. Both published by Elsevier, they offer hybrid Gold Open Access publishing options for authors at a cost and Green Open Access self-archiving options for authors to upload accepted manuscript versions of their articles to personal websites and institutional repositories. After an imposed embargo period has passed, these same accepted manuscript versions of articles can then be uploaded to high-traffic research sharing sites such as [Academia.edu](#) and [ResearchGate](#). Such EAP activity online provides a view onto a fast changing and opening landscape, with the examples offered here providing but a small window onto openness in EAP.

1.2 The massification of higher education

This case study has also been written to raise awareness around a growing number of learners worldwide, the current and arguably modest estimate being 100 million (Taylor, 2007; Uvalić-Trumbić & Daniel, 2011), who are qualified to pursue studies in higher education but for whom the cost of accessing traditional higher education to earn formal credentials is not realistic. Globally, English is the most widely adopted educational lingua franca in tertiary education and in international research and publishing. According to the UNESCO Education For All Global Monitoring Report (2008) there has been a boom in numbers for those students registered in traditional higher education programmes around the world with an estimated 138 million students in 2005, an increase of 45 million from 1999. The present mobile elite, estimated at 2.5 million (Altbach, Reisberg & Rumbley, 2009), of international students worldwide are in most cases also electing to study in English. This is having a competitive impact on the language of instruction elected for delivering higher education programmes internationally.

Continuing in the vein of massification, UNESCO's Institute for Lifelong Learning along with government and university initiatives worldwide are exploring new ways to scale widening access to higher education through different levels of openness, and are developing guidelines and research mandates for recognition, validation and accreditation systems in informal learning and in open and distance education. Technology has brought a diverse range of educational providers into the open, including those who are currently clamouring to participate in the provision of Massive Open Online Courses (MOOCs) with venture capital start-up funding. FutureLearn is the UK's response to the growing number of for-profit MOOC providers in the US, including [Udacity](#), [Udemy](#), and the largest MOOC enterprise of them all, [Coursera](#), which in February 2013 expanded its network to include 62 higher education institutions from 29 countries, offering MOOCs in 5 different languages.

Not-for-profit MOOCs also exist, such as edX from MIT and Harvard. edX develops and employs an open-source MOOC platform, [Open edX](#), which is shared with the wider community for the development of MOOCs outside of the edX consortium of handpicked elite university members. Other open education initiatives that plan to scale massively include the Open Educational Resources University ([OERu](#)), which will be launching in November 2013 and the Peer to Peer University ([P2PU](#)). Here with the OERu model, and unlike many of the for-profit MOOCs, all learning and teaching content will be open and licensed under Creative Commons. What is more, rather than

certificates of MOOC completion, the OERu global consortia of traditional open and distance universities will award university degrees. Once again, at no cost to the learner (or for a minimal fee for administering exams that can be waived through scholarships).

The jury is not yet out on MOOCs with their tens—or even hundreds—of thousands of learners in a single course but their emergence has garnered interest and a great deal of discussion around openness in higher education. Nevertheless, thousands of learners are registering for MOOCs. This speaks volumes to the nature of many non-formal online education offerings, especially MOOCs, the majority of which are offered in English with no or limited language support. For example, in the August 2012 edition of *Inside Higher Ed*, professors who had delivered Coursera MOOCs in literature and history subjects identified where they had experienced sudden drops in learner numbers when written English language components were introduced into their course assessments (Kolowich, 2012), raising questions about realistic and appropriate assessments for English-medium MOOCs that attract large numbers of learners from around the world.

At the time of writing this case study and of relevance to EAP was the *Critical Thinking* MOOC with Coursera delivered by Edinburgh University. In March 2013, there will be a further MOOC offering from Duke University – *English Composition 1: Achieving Expertise*. At present they have approximately 50,000 learners registered. It therefore stands to reason that there is an increasing need for flexible and high quality resources along with an expanded open infrastructure to support the teaching and learning of EAP across both traditional and informal modes of learning. The TOETOE project in collaboration with FLAX will be engaging in the research and development of flexible open corpus-based collections to support MOOCs to draw findings that can be generalised for EAP resources development in both traditional face-2-face teaching and in informal and non-formal open online learning.

Keywords

Academic Vocabulary Lists, AntConc, British Academic Written English (BAWE) Corpus, British National Corpus (BNC), Collocation, Compleat Lexical Tutor (Lextutor), Creative Commons, Data-Driven Learning, Design-Based Research, Durham University English Language Centre, English for Academic Purposes, English Language Teaching, Flexible Language Acquisition (FLAX), Google as a Corpus, MOOCs, Open Access, Open Educational Resources, Open Educational Practices, Oxford Text Archive, Teacher and Learner Cascade Training, Web as a Corpus, Wikipedia as a Corpus, WordandPhrase

2. What are language corpora anyway?

A language corpus (plural: corpora) is usually a collection of texts but the term can also refer to a collection of multimedia resources, for example audio-visual media that are linked to transcribed and tagged text files. These texts, which reflect features of written and spoken discourse, are enhanced by text analysis tools that enable us to discern noticeable language patterns to investigate the ways in which words combine with one another in different contexts of use.

Large collections of contemporary English that are also known as reference corpora provide a snapshot into how English is used in different writing and speaking contexts. Comparisons made between language as it is used in reference corpora and how it is used in academic sub-corpora, including published as well as student-generated academic writing corpora (for example, the British Academic Written English corpus) help us to identify which words and phrases occur more commonly in specific as well as in general academic contexts of use. Neither confined by the boundaries of the printed volume nor limited to the marketing enterprise of one-size-fits-all generic teaching resources from commercial publishers, these corpus-based e-resources are arguably more powerful than the average dictionary or EAP course book for insights into and practice with academic English for specific as well as general purposes. The large web-based projects presented in this case study open windows onto authentic collections of general and academic English with user interfaces for running text mining tools that are constantly undergoing new developments based on the latest research from applied corpus linguistics for supporting data-driven learning in language education.

In this case study, reference will be made to openly available linguistics resources such as dictionaries, thesauri,

collocations databases, and web-derived corpora from Google and Wikipedia that have been effectively linked to leading reference corpora, including the 100 million-word British National Corpus (see the [FLAX](#) project) and the 450 million-word Corpus of Contemporary American English (see the [WordandPhrase](#) project). In addition to this, a variety of learning support resources were developed by participating teachers in the OER cascade study for hands-on training with web-based text analysis tools for identifying and exploiting discipline-specific language from relevant published and student-generated text collections. These collections include the BAVE corpus of university student writing, and the sub-corpus of published academic writing within the Corpus of Contemporary American (COCA) English.

For a more extensive overview of corpus-based resources in English language teaching, please refer to the TOETOE project blog post, [Radio Ga Ga: corpus-based resources, you've yet to have your finest hour](#).

2.1 Data-Driven Learning

"...the attempt to cut out the middleman as far as possible and to give the learner direct access to the data" (Johns, 1991, p.30)

Importance is placed on empirical data when taking a corpus-informed and data-driven approach to language learning and teaching. Moving away from subjective conclusions about language based on an individual's internalized cognitive perception of language and the influence of generic language education resources, empirical data enable language teachers and learners to reach objective conclusions about specific uses of language based on corpus analyses. Tim Johns coined the term Data-Driven Learning (DDL) in 1991 with reference to the use of corpus data and the application of corpus-based practices in language learning and teaching (Johns, 1991). He also came up with the term English for Academic Purposes in 1974 (Hyland, 2006). The practice of DDL in language education was appropriated from computer science where language is treated as data and where "every student is a Sherlock Holmes", investigating the uses of language to assist with their acquisition of the target language (Johns, 2002:108).

Arguably, corpus developers have some of the greatest technological expertise in computational linguistics and the digital humanities, some of whom are no strangers to the meaning and practice of openness. More recently, there has been an identifiable effort among some corpus builders and concordancing text analysis software developers to engage in open practices with the development of, for example, open standards, open metadata, open collections, open tools and open communities (see Tom Cobb's Lextutor Facebook group and Laurence Anthony's AntConc Google group). Compare this technical expertise and knowledge of openness with that of the average language teacher who would not necessarily have received training in understanding the difference between open and proprietary resources and who would find different aspects of the digital infrastructure required in corpus building to pose insurmountable barriers.

Accessibility is a key issue with the design and uptake of corpus-based tools and resources for uses in mainstream language education where the focus is mainly on teaching and learning and not on research. There exist identifiable issues with the design of interfaces for text analysis tools that have been developed primarily by, and for, the corpus linguist rather than the language teacher and learner. As many of these corpus-based tools and resources stem from niche research and development projects, and because many of them are proprietary resources, there has been little in the way of formal training and support with DDL in EAP as well as in general language education.

3. Cascading openness in English for Academic Purposes

This case study presents insights from an OER cascade training and research project carried out at Durham University English Language Centre (DUELCC) with EAP teachers and students during the Michelmas (spring) term of 2012. Funding was awarded by Durham University from the *Enhancing the Student Learning Experience* award to visiting research and teaching fellow, Alannah Fitzgerald, to engage EAP teachers and students with the TOETOE project.

3.1 Openness and sharing on a shoestring

The OER cascade training project at DUELC is particularly relevant in the current climate of UK higher education, with many institutions forced to improve the quality of their teaching provision with continually diminishing resources and greater student expectations. Openly licensing and releasing resources online, beyond the realms of Durham University's password-protected Virtual Learning Environment (powered by Blackboard), would ensure that they can be shared with the wider EAP, open education and language teaching communities.

Programme leaders of the In-sessional, Louise Greener, and the Year-round Pre-sessional, Lesley Kendall, arranged for the OER cascade study to take place with three experienced EAP teachers at DUELC - Terri Edwards, Jeff Davidson and Clare Carr – who became “OER champions” at DUELC. The EAP teachers who became involved in the OER project did so for two main reasons. First, they wanted to see what open materials were available that would support the teaching of vocabulary, reading and writing. Second, they wanted to provide their students with tools that would help them to study beyond the classroom.

The teaching participants in this study volunteered to engage in open educational practices (OEP) for the design, development and delivery of open corpus-based workshops. Focus meetings were held once a week prior to the weekly workshops, which spanned a period of six weeks to discuss resources planning and the logistics of delivering the workshops as a team where the corpus resources were new to both the teachers and the students. Support OER in open file format were developed collaboratively with the project manager for teacher and learner training across two different EAP student cohorts (intermediate and proficient users of English). The focus of the OER cascade at DUELC was to develop teacher autonomy with open corpus-based tools and resources to assist the participating students with their English reading, writing and vocabulary acquisition in specific discipline areas. The training OER for EAP presented in this case study can be found online in the resources section of the [TOETOE project blog](#).

It was also anticipated that those teachers involved in the OER cascade would begin to develop the necessary competencies to engage with open corpus-based projects and resources for EAP teaching and materials development. EAP teachers are increasingly required to develop EAP programmes using authentic academic texts and assessments that address a range of students' needs across different academic disciplines.

BALEAP is trying to address a deficit in formal EAP training and has devised the Competency Framework for Teachers of English for Academic Purposes. Durham University's English Language Centre is one of a few BALEAP accredited institutions in the UK with teaching fellows who also work as BALEAP accreditors. It is also one of a few EAP university centres to have completely phased out the use of dedicated off-the-shelf published EAP teaching materials; fostering instead a keen focus on developing and re-visiting in-house materials for every EAP course they offer. DUELC has therefore been a good home for the TOETOE project in terms of developing and using OER, and for exploring issues around open licensing, open file formatting and open publishing to enable the sharing of OER across the wider EAP community.

3.2 Wider motivations for developing competencies in openness for EAP

For the participating teachers in this study, materials meetings in preparation of the cascade training sessions were held for discussing, understanding and planning the development of workshop OER for demonstrating and applying the selected open corpus-based tools and resources for training in EAP. The teachers also engaged in follow-up focus-group discussions for the perceived value of the selected corpus-based tools and resources for EAP and were encouraged to keep notes from observations of activities directed at both the teachers and students during the workshop training sessions.

To provide further evidence of the relevance of this wider-reaching form of collaboration, the HEFCE Online Learning Task Force released a comprehensive report entitled, 'Collaborate to Compete: seizing the opportunity of online learning for UK higher education' in January 2011. In this report, two out of the six recommendations made for increasing the competitiveness of HE in the UK reflect concerns for OER development and distribution. These comprise: recommendation number five - the realignment of training and development between technologists, learning support specialists and academics; and, recommendation number six - a call for investment in the development and exploitation of OER to enhance efficiency, quality and scale in learning and teaching. Practices and beliefs surrounding the value and impact of internationalising EAP curricular with OER for current and prospective international students

using open digital technologies have been explored in this project. Furthermore, being able to share and draw on findings with language teaching practitioners has also been an area of exploration with this project. The following interview discussion between the project manager, Alannah Fitzgerald, and one of the participating EAP teachers, Clare Carr, reflects how open educational practices were developed through the project and how they might be sustained beyond the project. A few days after the interview Clare presented her project findings at the EuroCALL CMC & Teacher Education SIGs Annual Workshop in Bologna where she shared best practice in OER with language teachers from around the world:

Alannah: How would you like to build on this training with OER for EAP in terms of the support that you need to go forward using corpus-based resources, especially open ones that you can access? What do you feel would be a useful way forward?

Clare: I think it's useful to work on this as more than one member of staff so you can then pull the ideas and find out what works and what didn't work. You've got to find what works within your own style of teaching but sometimes sharing that it worked if I did this - or have you tried this way, or this activity, or this exercise - was useful. And, also networking with other people in different institutions...it's useful to share with other institutions because you can get into a bubble where everyone's thinking the same way. It can be useful to find out how people elsewhere think about the resources.

By way of offering a preview into the corpus-based resources employed in the OER cascade at DUELC, which will be discussed in further detail in section four, the following webcast taken at the Cambridge 2012 conference in April 2012 provides insights from the project as presented by Terri Edwards and Alannah Fitzgerald.

Figure 1. [Webcast presentation](#) with Terri Edwards and Alannah Fitzgerald, 'Openness in ELT', at the Cambridge 2012 Conference.

The image shows a screenshot of a webcast presentation. At the top left, the logo for the 'OPEN COURSEWARE CONSORTIUM' is displayed, along with the 'gre' logo (Support Centre for Open Resources in Education) and a 'Login' link. The main title of the presentation is 'Openness in English Language Teaching', with the author listed as 'Alannah Fitzgerald' and the date as 'April 2012'. The license is 'Creative Commons Attribution (CC-BY)'. The central part of the image is a video frame showing two women in a presentation room. One woman is standing next to a large screen displaying the title 'Openness in English Language Teaching' and the word 'OPEN' in a stylized font. The other woman is seated at a desk with a laptop. Below the video frame, there are navigation buttons for 'Abstract', 'Slide contents', and 'Downloads'. At the bottom right, there are social media icons for Twitter, Facebook, Google+, and LinkedIn. The video frame also shows a close-up of a wooden board with the word 'OPEN' written in a natural, organic style using dried plant materials. The names 'Alannah Fitzgerald, Terri Edwards, Jeff Davidson & Lesley Kendall' are written below the board, along with logos for Durham University and The Open University.

The participating teachers actively contributed to dissemination events for cascading project outputs. All of the project outputs are openly available online and can be accessed via the [TOETOE project blog](#).

4. Open corpus-based resources for EAP

This section will take a look at the [FLAX](#), [AntConc](#), [Lextutor](#) and [WordandPhrase](#) projects in a little more detail, offering points for comparison and discussion of relevance to EAP. All of the projects offer powerful and valuable tools for language teaching and learning, however, on first view, the interfaces of these project websites may appear somewhat analytical and overwhelming. Nonetheless, it is well worth spending some time toying with these tools and their learning support resources for demonstrations and suggestions on how to use them in EAP and in general English language learning (or in French language learning as is the case with the Lextutor).

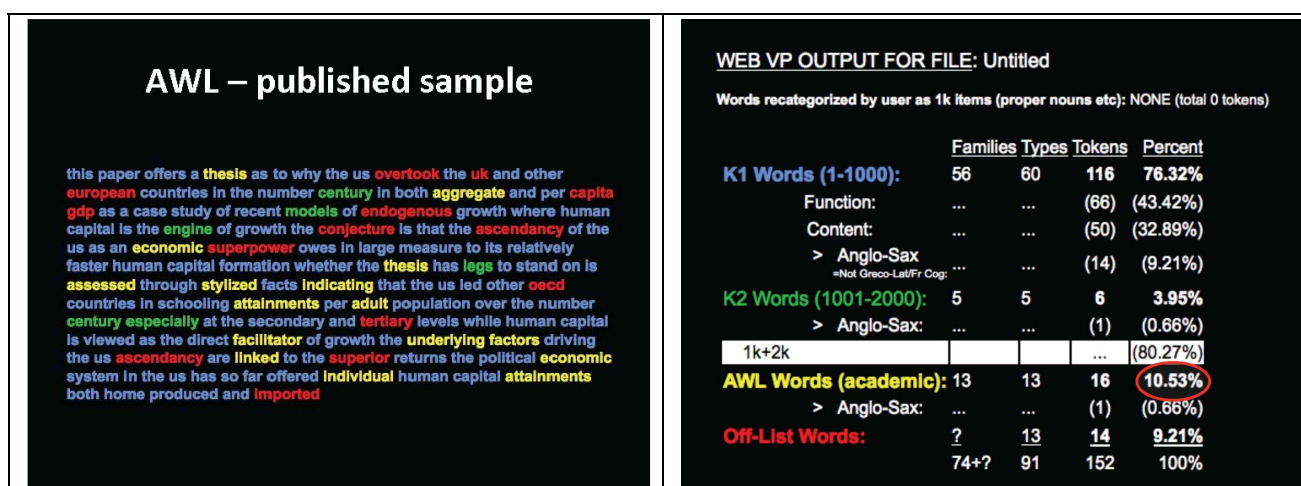
4.1 The Compleat Lexical Tutor (Lextutor)

The Compleat Lexical Tutor project, otherwise known as the Lextutor, was developed by Tom Cobb who is based at the Université du Québec à Montréal (UCAM) in Canada. Primary language resources in the Lextutor are for assisting with French and English language analysis. This interactive DDL website is possibly the most well known among English and French language “learners, teachers, course developers, and researchers worldwide, attracting 10,000+ visitors a week” (Lextutor website description). The Lextutor suite of tools and collections have been freely available on the web for well over a decade with the VocabProfile tool proving to be a popular choice among EAP teachers and students for analysing words in a text to identify what type of word list they belong to, for example academic or general word lists.

4.1.1 The VocabProfile tool and the Academic Word List (AWL) in Lextutor

The Academic Word List (AWL) was developed by Coxhead in 2000. Arguably, the AWL has been one of the most widely adopted corpus-based research outputs to be applied to EAP teaching and resources development. It is comprised of a group of 570 headwords that are common to published writing across the academic disciplines and builds on the work of the General Service List (GSL) of 2000 of the most frequently occurring words in English developed by West in 1953. It is important to bear in mind that this was groundbreaking research in 2000 and set a precedent for the use of corpus-based research in the development of EAP resources.

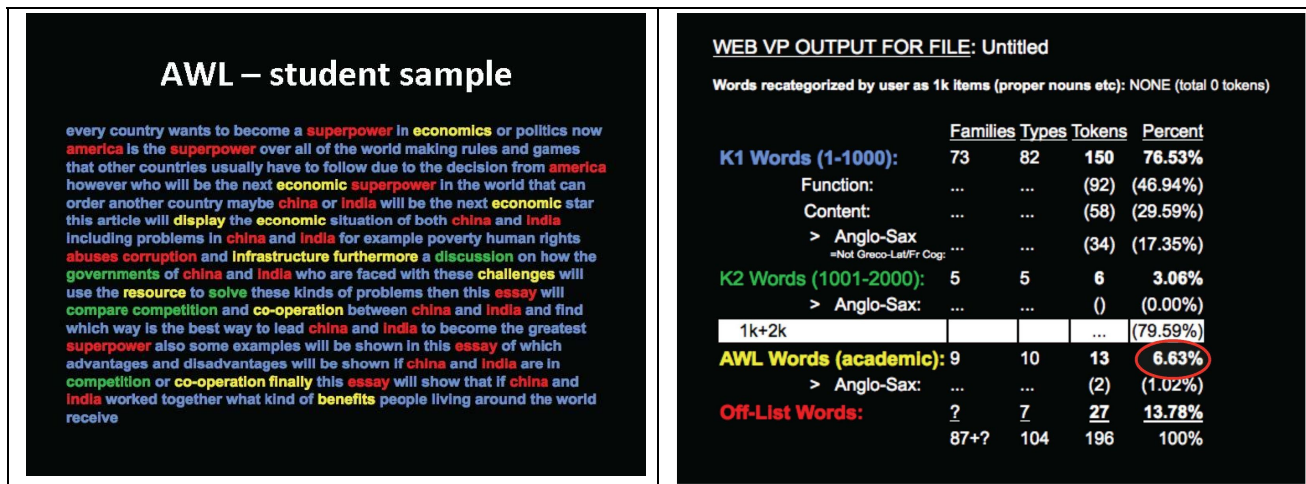
Figure 2. VocabProfile published text analysis output from the VocabProfile in the Lextutor



Coxhead’s research indicated that approximately 10% of the words generated from a broad sample of academic writing are represented by the AWL. In most cases a lot of ‘published’ writing, including writing from the genres of journalism and online magazines will also conform to Coxhead’s findings and return a count of at least 10% of words from the AWL. The VocabProfile in the Lextutor is therefore a quick and accessible online tool for EAP students to determine whether 10% of the words in an assignment or report they have written for their academic coursework come from the AWL. Figures 2 and 3 below show VocabProfile analyses of a published abstract for comparison with a

student essay abstract on the left-hand side of the tables (minus punctuation). An analyses of the language in both abstracts can be seen on the right-hand side of the tables according to the different word lists within the Lextutor they have been analysed against. The published abstract in figure 2 shows an AWL percentage of 10.53% whereas the student essay abstract in figure 3 shows an AWL percentage of only 6.63%.

Figure 3. Student essay text analysis from the VocabProfile in the Lextutor



In one of the OER cascade training activities the participating students were asked to bring along samples of their writing in soft copy format to the workshops so that they could analyse the percentage of words from the AWL that occurred in their writing. One student on the MA TESOL programme at DUELC who is a native speaker of English was surprised to learn that his essay when analysed only returned a percentage of 6.0% of words from the AWL. When studying the AWL or the newly developed Academic Vocabulary Lists, which will be introduced in section 4.4 on the WordandPhrase project, it is important to learn the derivatives (word families) of the headwords as well as the most frequent collocates that these words co-occur with in academic and more general published writing. A more detailed discussion on collocations will follow in section 4.2 with reference to the FLAX project. The following interview discussion between Terri Edwards and Alannah Fitzgerald reflects current concerns with the way the AWL has been taught and embedded into EAP teaching resources.

Alannah: I think the academic word list (AWL) was this amazing breakthrough but the awful thing is that people interpreted it as let's teach a list...and I'm sure that was not her [Avril Coxhead's] intention but it's the way that materials developers have taken it forward in many cases.

Terri: Yes, people always seize on something that's a very limited and quantifiable thing to use and deal with when language is not limited or quantifiable, that's the problem. And, when you've got students who are essentially instrumental learners, they will also seize on something that's limited and quantifiable...whereas, I think these tools [from the Lextutor and FLAX] show you that that's not the case and that's actually one of the beauties of them. I think one of the great things about these tools is that you can actually see language in action and how it works.

Returning momentarily to the VocabProfile tool in the Lextutor, it also has the added functionality of a RE-VP feature whereby you can edit your writing directly on the screen by trying to bring the number of yellow AWL items present in your writing up to and beyond 10% as can be seen in figure 4 below. In the same OER training cascade session with the Lextutor, we asked the students to consult the [AWL on the Victoria University website](#), where it is housed at the School of Linguistics and Applied Language Studies at Victoria University in New Zealand. From observations of the two classes who were assigned this task of trying to raise the percentage of AWL items using the RE-VP tool, we noted that the students were impeded by having to navigate back and forth to the AWL resource which exists on a different website (from the Lextutor) and is in .pdf format with the AWL presented as a static list.

Figure 4. VocabProfile tool in the Lextutor showing the RE-VP function for an updated AVL percentage of 9.88% based on a redrafted student essay abstract (Figure 3.) after consulting the AVL

VocabProfile edit & check

NEW - Edit-to-a-Profile

Profiled text Edit, Check with RE-VP and more

Current profile	
75.88%	Count
2.98	78.8
7.88	86.7
13.30	100

every country would like to establish themselves as a dominant superpower in the area of economics or politics now america is the superpower over all of the world making rules and games that other countries usually have to follow due to the decision from america however who will be the next economic superpower in the world that can order another country maybe china or india will be the next economic star this article will display the economic situation of both china and india including problems in china and india for example poverty human rights abuses corruption and infrastructure furthermore a discussion on how the governments of china and india who are faced with these challenges will use the resource to solve these kinds of problems then this essay will compare competition and co-operation between china and india and find

Every country would like to establish themselves as a dominant superpower in the areas of economics and politics. America's superpower status can be felt throughout the globe, and countries overseas are often required to adapt to American-style legislation that has been imposed as a result of foreign policy decisions made in America. Nevertheless, predictions about America's imminent global decline are evident and the question remains as to which country will become the next economic superpower.

4.2 FLAX – the Flexible Language Acquisition project

From its earliest inception the open-source FLAX project based at the Greenstone digital library lab at the University of Waikato's Computer Science Department in New Zealand has been envisioned and advanced with the language teacher and learner in mind. Since 2009 Alannah Fitzgerald has been engaged with the FLAX project for her PhD research to provide user feedback on the development of the language tools and collections in FLAX and to devise ways to promote the project resources within mainstream English language education and in open and distance education. A simplified and intuitive interface has been developed for presenting language collections with an endless supply of authentic language, support resources and interactive learning activities based on the powerful and complex handling of search queries from a range of interlinked tools and corpora.

In this case study we will be looking at different collections from the FLAX project, including the [Learning Collocations](#) collection, which is based on three corpora: the BNC, the BAWE and Wikipedia. We will also be looking at the [Web Phrases](#) and [Web Collocations](#) collections, which are based on a Google-derived English language corpus. A further series of collections in FLAX, which are highly relevant to openness in EAP, are the [BAWE collections in FLAX](#), which were developed for non-commercial educational and research purposes with the BAWE corpus of university student writing, distributed and managed by the [Oxford Text Archive](#), and now re-licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License](#). Training videos and a discussion on the development of the BAWE collections in FLAX, as part of the TOEFL International project with the University of Oxford IT Services, are covered in the *FLAX Weaving with Oxford Open Educational Resources* case study in collaboration with the HEA/JISC OER International Programme (Fitzgerald, 2013).

4.2.1 What are collocations and why are they so important?

The discussion will now turn to open language collections and activities in FLAX that have been designed to be helpful for learning collocations. Fernando in *Idioms and Idiomaticity*, defined collocations as “the company words keep” (1996). The main developer of FLAX, Shaoqun Wu, dedicated her entire PhD thesis to the development of open tools and resources for the advancement of a learning collocations system for English language learning (Wu, 2010). Without question, collocations are one of the most challenging aspects for students of a second or additional language to master due to the sheer size of collocational patterns as expressed by native speakers in any target language. However, research shows that the successful use of collocations in student writing and speaking supports not only improved levels of accuracy but also improved levels of fluency (Wray, 2002; Nesselhauf, 2003). Wu's research also identified a distinct lack of freely available and evidence-based online resources for effectively learning collocations.

4.2.2 The Learning Collocations collection in FLAX

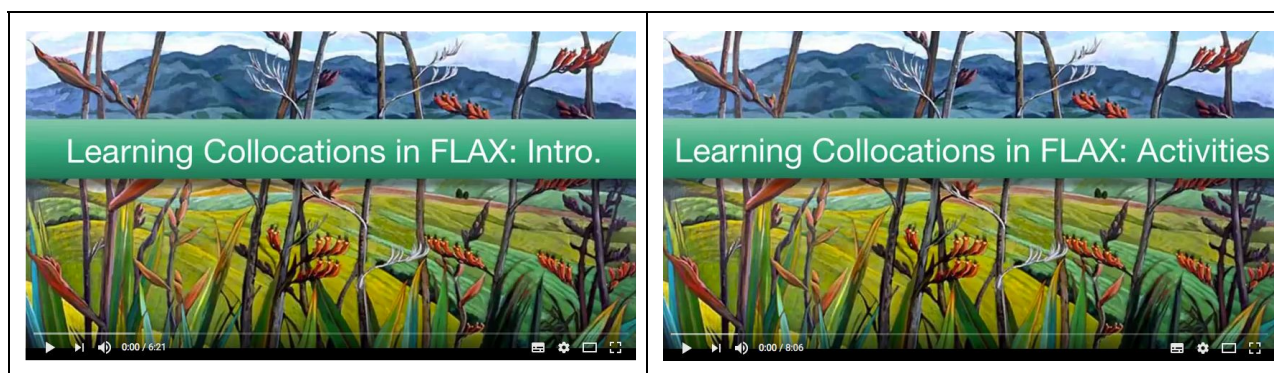
There are many definitions of collocation. At the FLAX project, we think of collocations in the same way as expressed by Benson et al.: "In any language, certain words combine with certain other words or grammatical constructions. These recurrent, semi-fixed combinations, or collocations, can be divided into two groups: grammatical collocations and lexical collocations." (Benson et al., 1986: ix)

The Learning Collocations collection in FLAX provides a dynamic user interface for searching for collocational patterns across three corpora: the Wikipedia collocations database derived from a crowd-sourced Wikipedia corpus of three million articles, the 100 million-word British National Corpus (BNC) reference corpus, and the British Academic Written English (BAWE) corpus of 2500 university student writing texts. The open source software in FLAX creates an easy-to-use experience for the search, retrieval, presentation and storage of phrases from the corpora, which are all authentic and highly contextualized examples of English in use. Comparisons for how these English language phrases are used in context from across the corpora and within their sub-corpora are further enhanced with open dictionary, thesaurus, and encyclopaedia resources that have been designed into the FLAX interface experience.

The FLAX Learning Collocations training videos in figure 5 focus on the following features:

- The Wikipedia Miner Toolkit based on artificial intelligence designs for searching across the corpora and presenting collocations in a user-friendly interface.
- The Cherry Basket for saving collocations from across the corpora using the Cherry Picking function, whereby collocations can be stored for reuse and printed or saved as export files.
- Game-based activities for dynamic interaction with collocations for advancing English language proficiency.

Figure 5. FLAX Learning Collocations Training Videos ([Introduction](#) and [Activities](#))



Participating teachers and students of the OER cascade viewed the affordance of being able to search for related collocations and definitions of key terms as assisting with an increase in the lexical range of student writing on a given topic. For an example of the employment of related collocations in student writing on the topic of the aging population, please see Fitzgerald (2012). These features for expanding collocational choices and linking to further resources can act as a gateway for students to check their understanding of key words and phrases as they are used in specific subject areas. Observations of student use of the Learning Collocations collection in the OER cascade study at DUELC led to the following discussion between Terri Edwards and Alannah Fitzgerald:

Terri: I mean this [Learning Collocations collection in FLAX] is so useful and this is the kind of thing our students massively lack. I mean, what was nice with [student] who was not having a very good day that day was when I sort of thought, well, I'll distract him and let's have a look at some geology stuff because he's a geologist. He was fascinated, absolutely fascinated and we just then started swapping words.

Alannah: *That's the great motivator, isn't it?*

Terri: *Yes.*

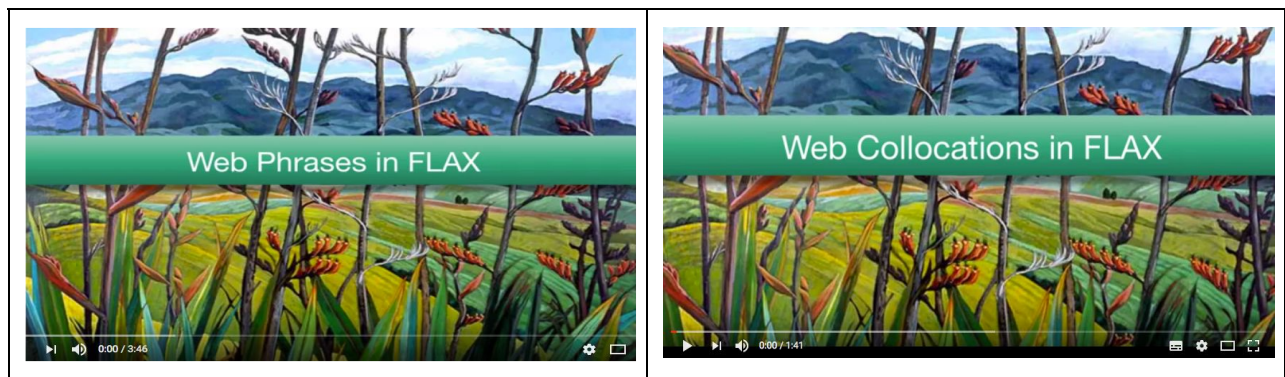
Alannah: *To get students to look in their own subject areas.*

Terri: *It was really nice for him to actually look at these to see how many words he actually knew and to then explain them and their collocates to me. He found that fantastically motivating. So, you know, I was milking it slightly... but it was good. And, again, I thought why don't we have each student looking at their subject-specific collocations sets?*

4.2.3 The Web Phrases and Web Collocations collections in FLAX

The [Web Phrases](#) and [Web Collocations](#) collections in FLAX are based on another extensive corpus of English derived from Google linguistic data. In particular, these Google-derived Web Phrases and Web Collocations collections allow users to identify problematic phrasing in writing by fine-tuning words that precede and follow keywords that they would like to use in their writing by drawing on this large database of English from Google. As shown in the training videos in figure 6, these collections allows users to substitute any awkward phrasing with naturally occurring phrases and collocations from the collections to improve the structure and the fluency of their writing.

Figure 6. FLAX Google [Web Phrases](#) and [Web Collocations](#) Training Videos on YouTube



4.2.4 Designing Data Driven Learning resources for the masses

One of the principle aims for the collaboration between the FLAX and TOETOE projects is to promote the understanding that the design of any technology user interface for uses in education has a better chance of success if it follows the design principles of simplicity, accessibility and functionality. Downes (2004) defines simplicity in educational technology design as those tools which are not only easy to use but those which have been designed to perform necessary functions only. In other words, if you enter a telephone box you are probably expecting to find a phone and possibly a phone book – you are not expecting to find Dr. Who's Tardis, although that could lead to a very interesting adventure, especially if Dr. Who has left some open training resources for you to work your way through to another dimension.

Research methods were employed in the OER cascade study to gauge experiences of using the newly introduced resources along with suggestions for further design iterations with the open corpus-based projects to achieve more beneficial and manageable outcomes for mainstream language teaching and learning needs. A typical feedback session around resource design is demonstrated in the following example from one of the focus discussions about the FLAX project website with Alannah Fitzgerald and Jeff Davidson, who has an engineering and design background:

Jeff: *Developers put in a lot of tools but they don't realise how a non-user can come to this homepage [FLAX] and*

think, well, how do I choose any of this? You've just given me a bag of tools and I don't know which one to use.

Alannah: That's right, and a lot of my work has been about trying to get people into the different project tools by demonstrating them. But that's a human layer...that's not a big enough resource, is it? Because I'm only touching so many people that I talk to or I might meet through my work. But if it's actually there on the project website it's going to be a lot better...it's going to reach a lot more people.

Jeff: I can just imagine people entering a kind of step one, step two – what would you like to do? I'd like to do this, this, or this, right? Which kind of collection would you like to look at and so on?

Alannah: I mean they [the FLAX software developers] do know this because the project is just in the research stage...but just to get these ideas from us is really good and that has become my role with giving feedback on the FLAX interface experience for teachers and learners of English. FLAX is all about radically transforming the user interface experience for those teachers and learners who want to use corpus tools but who get put off by a lot of the complexity with many of the corpus-based resources out there. And, you're probably right, I don't think DDL will ever become mainstream practice unless we design and develop the right interface experience for our teachers and learners – we're not corpus linguists after all.

Jeff: Aye...

Alannah: My feeling is, and this is only something anecdotal that I'll have to look into more thoroughly, is that the reason why most EAP teachers seem to only go to the VocabProfile tool in the Lextutor is because they've heard of the Academic Word List. And, the reason why most teachers would stop after the first few training videos with AntConc is that by the time you've reached training video number nine, it's like Clare said in the DIY corpora workshops - it's like going from beginners to advanced yoga in one go!

4.3 AntConc

AntConc is a multi-lingual freeware program developed by Laurence Anthony based at Waseda University in Japan for building language collections (corpora) and analysing texts. AntConc only requires light installation for use on any Mac, PC or Linux operating system.

4.3.1 DIY corpora with AntConc

Maggie Charles at the University of Oxford Language Centre has been carrying out studies with postgraduate EAP students, many of whom are working toward dissertation or thesis writing, for building DIY corpora using AntConc. Her recent paper in the *English for Specific Purposes Journal* (2012) points to AntConc's flexibility with student uptake due to it being freeware that can be installed on any personal computer or flash drive for portable use. The potential that AntConc offers for building select corpora to those students currently pursuing inter-disciplinary studies in higher education is also noted by Charles. Having said this, drawbacks with certain more obscure subject disciplines, for example Egyptology (Ibid.), that had not yet embraced digital research cultures and were still publishing research in predominantly print-based volumes or as image-based .pdf files made the development of DIY corpora still beyond the reach of those few students.

4.3.2 File conversion issues in corpus building

Converting .pdf files to .txt does prove to be one of the biggest challenges for someone who is faced with the task of building their first corpus as we observed in the OER cascade at Durham when introducing AntConc. 'Cleaning' the corpus to ensure that, for example, unnecessary formatting in the original documents is not carried over into the corpus is another issue to ensure its fast and smooth running with the text analysis tools available in AntConc. There are many .pdf conversion tools available online. For example, the free [Optical Character Recognition](#) (OCR) online service.

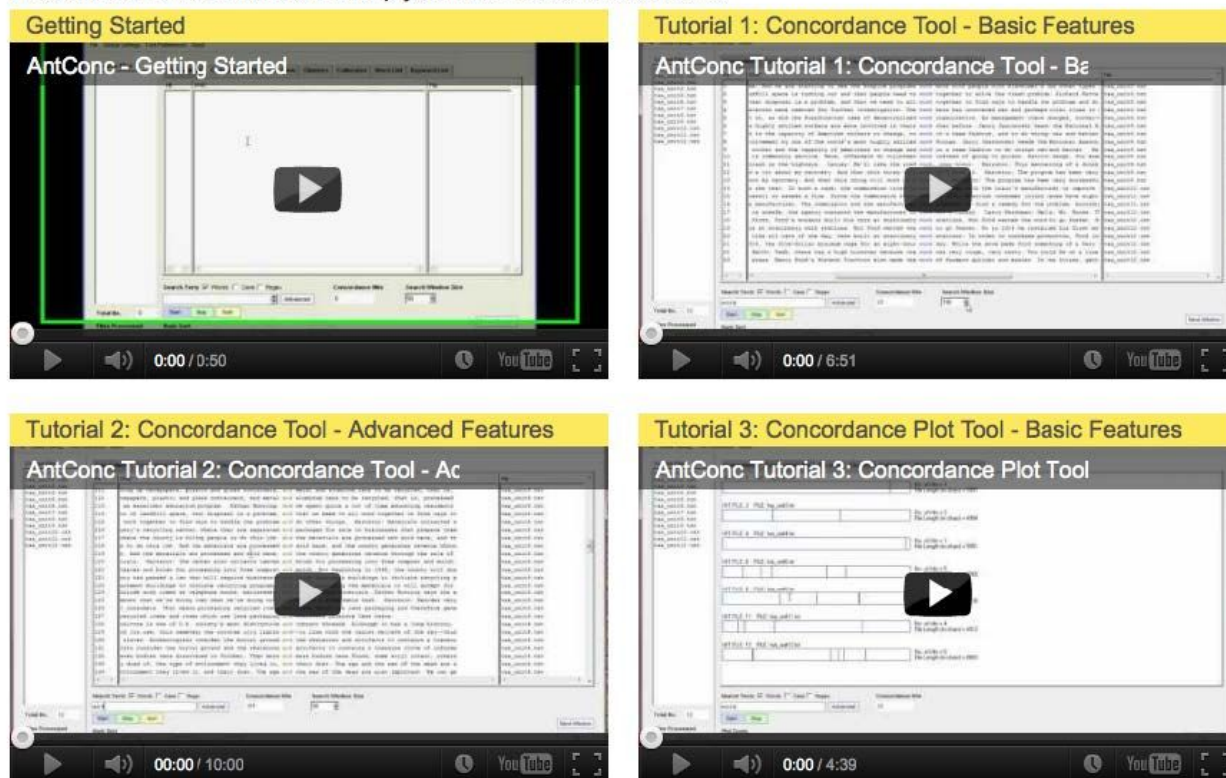
The AntConc training videos were used in the latter OER cascade training sessions where it was determined that the students had already progressed through basic concepts for data-driven learning and were ready to embark on building their own corpora using AntConc, which offers a more traditional concordancing (text analysis) interface as can be seen in the YouTube training videos in figure 7. The aim of this session was to build Do-It-Yourself corpora

based on required reading texts from the students' specific subject domains to identify key vocabulary and phrases. AntConc is also very useful for analysing a corpus of your own writing to highlight any overuse of certain words, phrases or forms.

Figure 7. [AntConc tutorial videos on YouTube.](#)

Video Tutorials

Here are some video tutorials to help you learn how to use AntConc.



All of the projects mentioned in this case study are continually undergoing upgrades for further development of tools and resources. Both Laurence Anthony of AntConc and Tom Cobb of the Lextutor are actively engaging with social networking media to sustain the active user groups of researchers and teachers who are interested in being involved in the development and testing of resources released on these project websites. You can subscribe to the AntConc-discussion group in Google Groups via the AntConc website. Laurence Anthony has developed many useful resources for building your own corpora which are available through his website, including a series of 9 detailed training videos which have been uploaded to YouTube. You can also join Tom Cobb's Lextutor Facebook group, which currently has around 200 members for assistance with using the Lextutor tools and resources, and you can also become involved in the resources development discussions with Tom Cobb and the other members.

4.4 WordandPhrase

Mark Davies and Lee Gardener at Brigham Young University (BYU) have endeavoured to develop Academic Vocabulary Lists (AVLs) rather than just one academic word list. In tandem with the development of new AVLs, Davies has developed a new resource-packed website interface, [WordandPhrase](#), for working with an academic sub-corpus of approximately 120 million words that exist within the Corpus of Contemporary American (COCA) English, also built by Davies. This site allows you to register and upload whole texts for analysis using the AVLs which have been derived from the 10,000 most frequent words in the COCA which is a large monitor corpus of approximately 450 million words, meaning that it is dynamically updating itself by continually pulling in new language

and growing at a rate of about 20 million words a year. Mark Davies released a series of [10 training videos on YouTube](#) for understanding and using the WordandPhrase website interface.

The completion of the OER cascade at DUELC roughly coincided with the BALEAP professional issues meeting in June 2012 on the EAP practitioner. An interesting discussion transpired from our workshop in response to other EAP practitioners' experience of using the AWL. Through this exchange we discovered that the new WordandPhrase project website for working with the academic sub corpus in COCA had the type of functionality we had been looking for with respects to supplying additional support resources all within the same user interface. With the most frequent collocates for words from the new AVLs (Davies usefully refers to collocates as 'nearby words'), along with synonyms and related words from the linked open data dictionary project, [WordNet](#), at Princeton, there is no longer any need to consult further resources such as wordlists, dictionaries and thesauri from other websites or offline resources.

Figure 6 WordandPhrase interface showing text analysis and word sketch outputs of a demonstration law text within the academic sub corpus of COCA.

The screenshot displays the Word and Phrase .INFO website interface. The main content area shows a text snippet: "After working more than 20 years in a court clerk's office, anyone would have a hugeI've-heard-it-all file. But that didn't stop Florence Prushan from feeling a shiver during one of her routine presentations about the Ventura County, Calif., Superior Court outreach program. The project is a leader in the movement to make the courts more user-friendly for the fast-growing numbers of litigants who do not have lawyers." The word "LAWYER" is highlighted in the text. Below the text, there are sections for "SYNONYMS", "DEFINITIONS", and "COLLOCATES". The "COLLOCATES" section shows a table of words and their frequencies, with "LAWYER" being the most frequent.

CLICK BAR TO LIMIT	HIS	EDU	SOC	LAW	HUM	PHIL	SCI	MED	BUS
PER MILL	0.5	0.1	0.2	7.2	0.1	0.4	0.1	0.1	0.8
SEE MORE	764	88	349	8565	150	521	256	101	935

5. Learning from open corpus-based resources for EAP

It is anticipated that the continued use and evaluation of the open corpus-based projects presented in this case study will in turn contribute to the growing number of open resources and practices in EAP. In particular, these open corpus-based projects offer great potential in the area of EAP materials development, whereby flexible corpus-derived resources can be developed for uses in both traditional EAP programmes and in informal and open online learning. The TOETOE project will continue to work at this interface for locating, sharing, communicating and promoting open educational practices with open corpus-based resources and their open learning and teaching derivatives for uses in EAP.

Observations from the training as well as the testing scenarios from this initial phase of the TOETOE project at DUELC were recorded for reflection and discussion on what worked well in the OER cascade. What was still needed from the corpus-based resources in terms of further development with learning support provision for novice corpus users, namely teachers and learners, specifically in the form of improved user interfaces, training videos and interactive learning activities was also reflected on by the participants in this study. Discussions around how best to support and develop EAP practitioners' competencies in Data-Driven Learning and the use of corpora were also raised in terms of what support resources would enable them to effectively embed these corpus-based resources into their language

teaching and materials development practices.

From the perspective of the participating teachers and students, the open corpus-based projects were much easier to use than had been anticipated and the flexibility of the resources for different uses in EAP teaching and learning far exceeded expectations. This points to a lack of exposure to and support with useful open corpus-based resources among EAP teachers and students in mainstream language education who are not yet exploiting the full potential of open corpus-based resources in their teaching and learning. Particular value was placed on those projects that successfully linked resources within one easy-to-use interface as with the FLAX Learning Collocations collections. The Academic Word List (AWL) was also perceived by participating students to be of high value and there were requests to have the AWL embedded within more corpus-based projects in addition to the VocabProfile tool in the Lextutor. A recommendation from this study would be to include the WordandPhrase project resources in subsequent training workshops for open corpus-based resources that offer the state of the art in Academic Vocabulary Lists. DIY corpora were viewed positively by the students and teachers but it was agreed that more support would be required with developing and using personalised corpora for independent learning over a longer period of time to achieve more substantial results.

6. New directions

This project work which began at DUELC has now moved onto the University of Oxford IT Services as the TOEFL International project. What was once an informal collaboration between TOEFL, FLAX, and the Oxford Text Archive, for the reuse and development of educational derivatives based on Oxford-managed research corpora has now become a formalised collaboration through the OER International programme with the HEA and the JISC (Fitzgerald, 2013).

Since the OER cascade at DUELC, TOEFL International has engaged with international stakeholders across eight different countries for the evaluation of Oxford content as it has been reused by the FLAX and TOEFL projects for uptake in English language education. User experiences of the FLAX tools and collections have been captured through various research methods, resulting in further design iterations with existing collections, and the development of new open source tools for enabling teachers to build interactive podcast corpora with Oxford university open lecture and seminar content.

The [BAWE collections in FLAX](#) have now become separate collections divided into corresponding sub corpora representing university student writing from the physical sciences, the life sciences, the social sciences and the arts and humanities. FLAX provides windows onto the full texts within the BAWE, revealing how the corpus has been organised by the developers of the corpus (Nesi et al, 2007) into different genres as well as into different disciplines. New automated features for the BAWE in FLAX include word lists, word keyness indicators, lexical bundles, links to the FLAX collocations database, and glossary features using Wikipedia to enable teachers and students to manage unfamiliar language that is specific to academic English from across the disciplines.

Further open language collections and interactive language activities are currently being planned for development in collaboration with the FLAX project. Based on a continuing exploration into educational and promotional channels for ELT, recommendations for cutting tracks through perceived barriers for the widespread adoption of DDL in EAP will be made by investigating current movements within open and informal education. It is hoped that the ongoing development, promotion and sharing of open language projects for use in higher education will reach all audiences who require support with EAP as it is used across international academic contexts and as it iterates toward openness.

7. References

Alexander, O., Bell, D., Cardew, S., King, J., Pallant, A., Scott, M., Thomas, D., & Ward Goodbody, M. (2008). *Competency framework for teachers of English for Academic Purposes, BALEAP*.

Altbach, P. G., Reisberg, L., & Rumbley, L. E. (2009). *Trends in Global Higher Education: Tracking an Academic Revolution*. A Report Prepared for the UNESCO 2009 World Conference on Higher Education. Retrieved from <http://unesdoc.unesco.org/images/0018/001832/183219e.pdf>

Anthony, L. (n.d.). Laurence Anthony's Website: AntConc. Retrieved from <http://www.antlab.sci.waseda.ac.jp/software.html>

Benson, M. Benson, E., Ilsen, R.F. (1986). *The BBI combinatory dictionary of English: A guide to word combinations*. Amsterdam/Philadelphia: John Benjamins.

British National Corpus, version 3 (BNC XML Edition). 2007. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: <http://www.natcorp.ox.ac.uk>

Cobb, T. (n.d.). Compleat Lexical Tutor. Retrieved from <http://www.lextutor.ca/>

Comas-Quinn, A. & Fitzgerald, A. (2013). *Open Educational Resources in Language Teaching and Learning*. Open Educational Resources Case Study: Pedagogical development from OER practice. Commissioned by the Higher Education Academy (HEA) and the Joint Information Systems Committee (JISC), United Kingdom.

Coxhead, A. (2000). *The Academic Word List*. Retrieved from <http://www.victoria.ac.nz/lals/resources/academicwordlist>

Charles, M. (2012). 'Proper vocabulary and juicy collocations': EAP students evaluate do-it-yourself corpus-building. *English for Specific Purposes*, 31: 93-102.

Davies, M. (1991-present). *The Corpus of Contemporary American English (COCA)*. Retrieved from <http://corpus.byu.edu/coca/>

Davies, M. & Gardener, D. (n.d.) *WordsandPhrases*. Retrieved from <http://www.wordandphrase.info>

Fitzgerald, A. (2012). *Openness in English Language Teaching*. Cambridge 2012: Innovation and Impact – Openly Collaborating to Enhance Education. Cambridge, UK.

Fitzgerald, A. (2013). *TOEFL International: FLAX Weaving with Oxford Open Educational Resources*. Open Educational Resources International Case Study. Commissioned by the Higher Education Academy (HEA) and the Joint Information Systems Committee (JISC), United Kingdom.

FLAX. (n.d.). *The "Flexible Language Acquisition Project"*. Retrieved from <http://flax.nzdl.org/>

Fernando, C. (1996). *Idioms and Idiomaticity*. Oxford: Oxford University Press.

Hoffmann, S., Evert, S., Smith, N., Lee, D. & Berglund Prytz, Y. (2008). *Corpus Linguistics with BNCweb - a Practical Guide*. Frankfurt am Main: Peter Lang.

Hyland, K. (2006). *English for Academic Purposes: An Advanced Handbook*. London: Routledge.

Johns, T. (1994). From Printout to Handout: Grammar and Vocabulary Teaching in the Context of Data-driven Learning. In Odlin, T. (ed.), *Perspectives on Pedagogical Grammar*: 27-45. Cambridge: Cambridge University Press.

Moxley, J. (n.d.) Writing Commons. Retrieved from <http://writingcommons.org/>

Nesi, H, Gardner, S., Thompson, P. & Wickens, P. (2007) The British Academic Written English (BAWE) corpus, developed at the Universities of Warwick, Reading and Oxford Brookes under the directorship of Hilary Nesi and Sheena Gardner (formerly of the Centre for Applied Linguistics [previously called CELTE], Warwick), Paul Thompson (Department of Applied Linguistics, Reading) and Paul Wickens (Westminster Institute of Education, Oxford Brookes), with funding from the ESRC (RES-000-23-0800)

Nesi, H. and Gardner, S. (2012) *Genres across the Disciplines: Student writing in higher education*. Cambridge: Cambridge University Press.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24(2): 223–242.

Report to HEFCE by the Online Learning Task Force. (2011) Collaborate to compete: Seizing the opportunity of online learning for UK higher education.

Taylor, J. (2007). *Open Courseware Futures: Creating a Parallel Universe*. e-Journal of Instructional Science and Technology (e-JIST), 10(1).

Tribble, C. (2012). *Teaching and Language Corpora Survey*.

UNESCO (2008). *Education For All. Global Monitoring Report 2008*. United Nations Education Scientific Cultural Organisation. Retrieved from www.efareport.unesco.org

Uvalić-Trumbić, S & Daniel, J. (2011). *Let a thousand flowers bloom!* UNESCO Global Forum on Rankings and Accountability in Higher Education: Uses and Misuses.

Wagner, C., & Wong, S. (2012). Unseen Science? Representation of the BRICs in Global Science. *Scientometrics* 90: 1001–1013.

West, M. (1953). *A General Service List of English Words*. Longman: London

Wray, A. (2002). *Formulaic Language and the Lexicon*. New York: Oxford University Press.

Wu, S., Franken, M. & Witten, I. H. (2009). Refining the use of the web (and web search) as a language teaching and learning resource. *Computer Assisted Language Learning* 22(3): 249-268.

Wu, S., Franken, M., & Witten, I. H. (2010). Supporting collocation learning with a digital library. *Computer Assisted Language Learning*. 23(1): 87-110.

Wu, S., Witten, I. H. & Franken, M. (2010). Utilizing lexical data from a Web-derived corpus to expand productive collocation knowledge. *ReCALL*, 22(1): 83–102.

Wu, S. (2010). Supporting collocation learning. *Doctoral Thesis, Computer Science Department, University of Waikato, New Zealand*. Retrieved from the Research Commons at the University of Waikato.

8. License and citation

This work is licensed under the Creative Commons Attribution License <http://creativecommons.org/licenses/by/3.0/>.

Please cite this work as: Fitzgerald, A. (2013). Openness in English for Academic Purposes. Open Educational Resources Case Study: Pedagogical development from OER practice. Commissioned by the Higher Education Academy (HEA) and the Joint Information Systems Committee (JISC), United Kingdom.