Title of Project:
A New Paradigm for Open Data-driven Language Learning Systems Design in Higher Education

Researcher:
Alannah Fitzgerald
alannahfitzgerald@gmail.com

Current Affiliation
The University of Waikato

Doctoral Institution
Concordia University

Research Supervisor:
Dr. Steven Shaw
Concordia University


Alannah Fitzgerald

Final Report

Motivation for the Research

This doctoral thesis presents three studies in collaboration with the open source FLAX project (Flexible Language Acquisition flax.nzdl.org). This research makes an original contribution to the fields of language education and educational technology by mobilizing knowledge from computer science, corpus linguistics, and open education and proposes a new paradigm for open data-driven language learning systems design in higher education. Furthermore, the research presented in this thesis uncovers and engages with an infrastructure of open educational practices (OEP) that push at the parameters of policy for the reuse of open access research and pedagogic content in the design, development, distribution, adoption, and evaluation of data-driven language learning systems.

A central proposition of this thesis with publications is that where language corpora have been deployed in the research for linguistic analyses by researchers, the knowledge generated has often failed to translate into the design of openly accessible pedagogical applications for data-driven learning (DDL). Instead what we have witnessed is corpus systems that have been designed and developed primarily by and for corpus linguists for research purposes. This failure in knowledge translation is due in no small part to the following issues: copyright restrictions with the texts in corpus building that inhibit text data mining and sharing; subscription costs with natural language processing (NLP) and text analysis software tools that restrict access; and complex user interface designs of NLP and text analysis tools that limit uptake and utilization by

non-expert users, namely language teachers and language learners.

As per the directives of the editors working for the various publishers in preparing these manuscripts, the two journal articles and one book chapter central to Studies 1-3 of the thesis are very concise documents, two of which have been submitted for publication and one of which has been published by Routledge with TIRF as part of the *Global Research on Teaching and Learning English Series* in the area of digital language learning and teaching: research, theory and practice for which this doctoral dissertation grant was awarded. In order to ensure that the thesis reads as a coherent whole, additional bridging material has been included with references to supporting research with the FLAX project that foreground each study and show the relationship between the central studies in the thesis. The main thrust of the thesis demonstrates the iterative nature of all three inter-related research studies and their implications for current as well as planned future research, in addition to implications for policy and pedagogy.

Research Questions

For Study 1, the following research questions were investigated:

1. To what extent can open access content foster open educational practices among academic English language stakeholders for designing, developing and evaluating data-driven language learning resources?
2. What impact do the underlying business models and cultural practices of institutions and organizations have on open educational practices for remixing open access content in the design, development, implementation, and dissemination of resources for EAP in higher education?

In response to the language support collections developed for Study 2 for applications in non-formal online learning (MOOCs), the following research questions were devised as a basis to collect data from participants on their perceived experience of using the FLAX system:

1. Are automated domain-specific terminology learning support systems perceived as motivating to use (i.e., user-friendly and efficacious) in non-formal online learning where there is no formal language support provision?
2. Do the affordances of being able to browse and search data-mined course content that has been linked to auxiliary resources positively augment the learning and usability experience of MOOC platforms and learning management systems?

An identifiable lack of corpus-informed resources for teaching and learning specialized varieties of English was a motivating factor for conducting the experiment described in Study 3 in response to the following research questions:

1. To what extent can the digital commons of open and authentic content enrich data-driven learning across formal and informal language learning?

2. What effect does the application of DDL methods for querying open and authentic content have on the acquisition of specialized terminology, as opposed to accessing non-DDL-based online resources?

Research Methodology

Study 1 employs automated content analysis to mine the concept of open educational systems and practices from qualitative reflections spanning 2012-2019 with stakeholders from an on-going multi-site design-based research study with the FLAX project. Design considerations are presented for remixing domain-specific open access content for academic English language provision across formal and non-formal higher education contexts. Primary stakeholders in this ongoing research collaboration include the following: knowledge organizations – libraries and archives including the British Library and the Oxford Text Archive, universities in collaboration with Massive Open Online Course (MOOC) providers; an interdisciplinary team of researchers; and knowledge users in formal higher education – English for Academic Purposes (EAP) practitioners.

Study 2 presents a data-driven experiment in non-formal higher education by triangulating user query system log data with learner participant data from surveys (N=174) on the interface designs and usability of an automated open source digital library scheme, FLAX. Text and data mining approaches (TDM) common to NLP were applied to pedagogical English language corpora, derived from the content of two MOOCs, (Harvard University with edX, and the University of London with Coursera), and one networked course (Harvard Law School with the Berkman Klein Center for Internet and Society), which were then linked to external open resources (e.g., Wikipedia, the FLAX Learning Collocations system, WordNet), so that learners could employ the information discovery techniques (e.g., searching and browsing) that they have become accustomed to using through search engines (e.g., Google, Bing) for discovering and learning the domain-specific language features of their interests.

Study 3 presents a data-driven experiment in formal higher education from the legal English field to measure quantitatively the usefulness and effectiveness of employing the open Law Collections in FLAX in the teaching of legal English at the University of Murcia in Spain. Informants were divided into an experimental and a control group and were asked to write an essay on a given set of legal English topics, defined by the subject instructor as part of their final assessment. The experimental group only consulted the FLAX English Common Law MOOC collection as the single source of information to draft their essays, and the control group used any information source available from the Internet to draft their essays.

Summary of Findings

Themes arising from the qualitative dataset in Study 1 point to affordances as well as barriers with the adoption of open policies and practices for remixing open access content for data-driven language learning applications in higher education. These themes are discussed against the backdrop of different business models and cultural practices present within participating

knowledge organizations. The design research interventions and findings captured in Study 1 are further supported and evolved by the mixed methods of research inquiry employed in Studies 2 and 3.

Findings from Study 2 indicate a positive user experience with interfaces that include advanced affordances for course content browse, search and retrieval that transcend the MOOC platform and learning management system (LMS) standard. Further survey questions derived from an open education research bank from the Hewlett Foundation are reused in this study and presented against a larger dataset from the Hewlett Foundation (N=1921) on motivations for the uptake of open educational resources.

Findings from an analysis of the two learner corpora of essays in Study 3 indicate that members of the experimental group appear to have acquired the specialized terminology of the area better than those in the control group, as attested by the higher term average obtained by the texts in the FLAX-based corpus (56.5) as opposed to the non-FLAX-based text collection, at 13.73 points below.

Implications

A new paradigm for open data-driven language learning systems design in higher education: A basic premise underpinning the new research paradigm presented in this thesis, and demonstrated by the FLAX project, is that open data-driven language learning systems design as an approach is learner-centric and operates with the interface to the learner. Whether the learner is operating fully online in non-formal or informal learning mode or in a blended modality that is based both within and beyond the formal language classroom, this approach requires that the tools and interfaces, and indeed the corpora, be openly accessible and remixable for development or adaptation to meet this specific learner requirement. This method is different from existing DDL approaches which assume specialised knowledge or experience with DDL tools, interfaces and strategies, operating on mostly inaccessible corpora in terms of cost or design, or alternatively assuming training to, hopefully, compensate for this lack of knowledge and experience.

From a research and development (R&D) standpoint, the paradigm presented here also operates with the interface to knowledge organisations (universities, libraries, archives) and researchers who are engaging with open educational practices to push at the parameters of open policy for the non-commercial reuse and remix of authentic research and pedagogic content that is increasingly abundant in digital open access format for TDM purposes. This open access content is highly relevant to learning features of specialist varieties of English from across the academy but is otherwise off limits for development into proprietary learning materials by the commercial education publishing industry. Indeed, the open corpus development work presented in this thesis would not have been possible had it not been for the campaigners for copyright reform, the Internet activists, the open policy makers, the open-source software developers, and the advocates for open access, open data, and open education that have made these resources available for reuse and remix.

This paradigm leads down several paths, including research into understanding how users actually perceive, appropriate and use the approach based on the open tools and resources provided. This inquiry informs their design and development, in an R&D process that is presented here through the methodological lens of design-based research and design ethnography. This approach will be fundamentally different than if we assume that the user is actually a DDL or linguistics expert or that such an expert will be the learner's interface to the system, by preparing output for the learner to experience and learn from. This approach will also be necessarily different than if we assume the user is always a formally registered student at a university with access to EAP support that may or may not offer DDL or linguistics expertise for learning the language features of specific discourse communities from across the academy.

The assumption behind this new paradigm that the right tools and resources can allow the end-learner to drive the processes autonomously is fundamentally revolutionary. This premise goes to the original contribution to knowledge of this thesis, but it also challenges and directs researchers and practitioners in the field to consider and take up this new direction with open data-driven language learning systems design for applications that can be scaled in higher education to meet the increasing numbers of learners who are coming online.

The focus on domain-specific terminology learning support via data-driven approaches is of course also decidedly different from the current EAP paradigm which in mainstream practice has been steadily evolving away from its roots in English for Specific Purposes (ESP), domain specificity and DDL processes towards the generic skills and knowledge programs currently in vogue that are arguably being steered by generic EAP coursebook publications from the commercial education publishing industry.

Thus, this is also a new paradigm based on DDL approaches, driving domain-specific terminology learning support for EAP across formal, non-formal and informal learning modalities in higher education. It will transform, potentially, the focus of DDL systems design developments in language support and learning in general toward the non-specialist end-learner, but also hopefully help re-establish the centrality of language specificity to the field of EAP.

The new paradigm is necessarily rooted in greater inter- or multi-disciplinarity. Given the goal of facilitating, in particular, the increasing number of learners who are coming online, and users of large-scale MOOC platforms who are trying to function in domain-specific subject areas that are invariably offered in the English language, the approach requires collaboration and cooperation among platform providers, subject academics and instructors, educational technologists, software developers, educational researchers, linguists and EAP practitioners with expertise in corpus-based and DDL approaches, and policy makers in knowledge organizations, such as in libraries, universities, and archives.

References

Ädel, A. (2010). Using corpora to teach academic writing. In M. C. Campoy Cubillo, B. Bellés Fortuño & M. L. Gea-Valor (Eds.), *Corpus-based approaches in English language teaching* (pp. 39–55). London, UK: Continuum.

Alcorn, B., Christensen, G., & Kapur, D. (2015) Higher education and MOOCs in India and the global south. *Change: The Magazine of Higher Learning*, *47*(3), 42–49. DOI: 10.1080/00091383.2015.1040710

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

Amiel, T., & Reeves, T. C. (2008). Design-based research and educational technology: Rethinking technology and the research agenda. *Educational Technology & Society*, *11* (4), 29–40.

Ananiadou, S., & Mcnaught, J. (Eds.) (2006). *Text mining for biology and biomedicine*. Boston, MA/London, UK: Artech House.

Ananiadou, S., Rea, B., Okazaki, N., Procter, R., & Thomas, J. (2009). Supporting systematic reviews using text mining. *Social Science Computing Review*, *27*, 509–523. doi:10.1177/0894439309332293

Ananiadou, S., Thompson, P., Thomas, J., Mu, T., Oliver, S., Rickinson, M., Sasaki, Y., Weissenbacher, D., & McNaught, J. (2010). Supporting the education evidence portal with text mining. *Philosophical Transactions of the Royal Society*, *368*, 3,829–3,844. doi:10.1098/rsta.2010.0152

Angus, D., Rintel, S., &Wiles, J. (2013). Making sense of big text: A visual-first approach for analyzing text data using Leximancer and Discursis. *International Journal of Social Research Methodology*, *16*(3), 261–267.

Anderson, T. (2007). Design-based research: A new research paradigm for open and distance learning. Keynote address given at the Community of Inquiry Chais Research Centre Annual Education Technology Conference. The Open University of Israel, Ra'anana, Israel. Retrieved from http://www.slideshare.net/terrya/design-based-research-new-research-paradigm

Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher*, *41*(1), 16–25.

Anthony, L. (2014, July). A view to the future in corpus tools development. 11th Teaching and Language Corpora Conference (TALC 11) Keynote Address. Lancaster University, UK.

Anthony, L. (2018). *Introducing English for specific purposes*. London, UK: Routledge.

Argyris, C., & Schön, D. (1991). Participatory action research and action science compared. In W. F. Whyte (Ed.), *Participatory action research* (pp. 85–96). Newbury Park, NJ: Sage.

Arum, R., & Roska, J. (2011). *Academically adrift: limited learning on college campuses*. Chicago, IL: Chicago University Press.

Atenas, J., Havemann, L., & Priego, E. (2015). Open data as open educational resources: Towards transversal skills and global citizenship. *Open Praxis*, 7(4), 377–389.

Atenas, J. (2015). Model for democratisation of the contents hosted in MOOCs. *RUSC Universities and Knowledge Society Journal*, *12*(1), 3–14. doi http://dx.doi.org/10.7238/rusc.v12i1.2031

Bainbridge, J., Melitski, J., Zahradnik, A., Lauria, E., Jayaprakash, S., & Baron, J. (2015). Using learning analytics to predict at-risk students in online graduate public affairs and administration education. *Journal of Public Affairs Education*, *21*(2), 247–262.

Balfour, S. P. (2013). Assessing writing in MOOCs: Automated essay scoring and calibrated peer review™. *Research & Practice in Assessment*, *8*, 40–48.

Barab, S., & Squire, L. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences*, *13*(1), 1–14.

Barefoot, B. (2004). Higher education's revolving door: confronting the problem of student drop out in US colleges and universities. *Open Learning*, *19*(1), 9–19.

Baskerville, R. L., & Myers, M. D. (2015). Design ethnography in information systems. *Information Systems Journal*, *25*(1), 23–46.

Bayne, S., Knox, J., & Ross, J. (2015). Open education: The need for a critical approach. *Learning, Media and Technology*, *40*(3), 247–250. DOI: 10.1080/17439884.2015.1065272

Belcher, D. (2010). What ESP is and can be: An introduction. In D. Belcher (Ed.), *English for specific purposes in theory and practice* (pp. 1-20). Ann Arbor, MI: University of Michigan Press.

Bell, P. (2004). On the theoretical breadth of design-based research in education. *Educational Psychologist*, *39*(4), 243–253.

Benkler, Y. (2007). *The wealth of networks: How social production transforms markets and freedom*. New Haven, CT: Yale University Press.

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

Bereiter, C. (2002). Design research for sustained innovation. *Cognitive Studies, Bulletin of the Japanese Cognitive Science Society, 9*(3), 321-327.

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

Bernardini, S. (2004). Corpora in the classroom: An overview and some reflections on future developments. In J. McH. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 15–36). Amsterdam, The Netherlands: John Benjamins.

Bhatia, V. K. (1993). *Analysing genre: Language use in professional settings*. London, UK: Longman.

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at…: Lexical bundles in university teaching and textbooks. *Applied Linguistics, 25*(3), 371–405.

Biber, D. (2006). *University language: a corpus-based study of spoken and written registers*. Amsterdam, The Netherlands: John Benjamins.

Biber, D., & Kurjian, J. (2007). Towards a taxonomy of web registers and text types: A multi-dimensional analysis. *Language and Computers, 59*(1), 109–130.

Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purpose, 26*, 263–286.

Bichard, J., & Gheerawo, R. (2010). The ethnography in design. In A. J. Clarke (Ed.), *Design anthropology: Object culture in the 21st century*. New York, NY: SpringerWien.

Biesta, G. (2010). *Good education in an age of measurement: Ethics, politics, democracy*. Boulder, CO: Paradigm.

Bishop, H. (2004). The effect of typographical salience on the look up and comprehension of unknown formulaic sequences. In N. Schmidt (Ed.). *Formulaic sequences: Acquisition, processing and use* (pp. 227–244). Philadelphia, PA: John Benjamins.

Blomberg, J. (1993). Ethnographic field methods and their relation to design. In D. Schuler & A. Namioka (Eds.), *Participatory design: principles and practices*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bollier, D. (2007). The growth of the commons paradigm. In C. Hess & E. Ostrom (Eds.). *Understanding knowledge as a commons: From theory to practice* (pp. 27–40). Cambridge, MA: MIT Press.

Bordet, G. (2014). Influence of collocational variations on making the PhD abstract an effective "would-be insider" self-promotional tool. In M. Bondi & R. Lores Sanz (Eds.), *Abstracts in academic discourse: Variation and change* (pp. 131–160). Bern, Switzerland: Peter Lang.

Bordet, G. (2015). The role of "Lexical Paving" in building a text according to the requirements of a target genre. In P. Thompson & G. Diani (Eds.), *English for academic purposes: Approaches and implications* (pp. 43–66). Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

Borgman, C. (2015). *Big data, little data, no data: Scholarship in the networked world*. Cambridge, MA: MIT Press

Borthwick, K., & Gallagher-Brett, A. (2014). Inspiration, ideas, encouragement: teacher development and improved use of technology in language teaching through open educational practice. *Computer Assisted Language Learning*, *27*(2), 163–183. DOI: 10.1080/09588221.2013.818560

Bosman, J. (2012, March 13). After 244 years, Encyclopaedia Britannica stops the presses. *The New York Times*.

Boulton, A. (2009). Testing the limits of data-driven learning: language proficiency and training. *ReCALL*, *21*(1), 37–54.

Boulton, A. (2010a). Data-driven learning: Taking the computer out of the equation. *Language Learning*, *60*(3), 534–572.

Boulton, A. (2010b). Data-driven learning: On paper, in practice. In T. Harris & M. Moreno Jaén, (Eds.). *Corpus linguistics in language teaching* (pp. 17–52). Bern, Switzerland: Peter Lang.

Boulton, A. (2011). Data-driven learning: The perpetual enigma. In S. Roszkowski & B. Lewandowska-Tomaszczyk (Eds.), *Explorations across languages and corpora* (pp. 563–580). Frankfurt, Germany: Peter Lang.

Boulton, A. (2012a). Beyond concordancing: Multiple affordances of corpora in university language degrees. *Languages, Cultures and Virtual Communities*, *34*, 33–38.

Boulton, A. (2012b). Hands-on / hands-off: Alternative approaches to data-driven learning. In J. Thomas & A. Boulton (Eds.), *Input, process, and product: Developments in teaching and language corpora* (pp. 152–168). Brno, Czech Republic: Masaryk University Press.

Boulton, A., & Thomas, J. (2012). Corpus language input, corpus processes in learning, learner corpus product. In J. Thomas & A. Boulton (Eds.), *Input, process and product: Developments in teaching and language corpora* (pp. 7–34). Brno, Czech Republic: Masaryk University Press.

Boulton, A. (2013). Wanted: Large corpus, simple software. No timewasters. In A. Leńko-Szymańska. *Proceedings of TaLC10: 10th International Conference on Teaching and Language Corpora* (pp.1–6). Warsaw, Poland: University of Warsaw.

Boulton, A., & Pérez-Paredes, P. (2014). ReCALL special issue: Researching uses of corpora for language teaching and learning Editorial Researching uses of corpora for language teaching and learning. *ReCALL, 26*(2), 121-127. doi:10.1017/S0958344014000068

Boulton, A. (2015). Applying data-driven learning to the web. In A. Lenko-Szymanska & A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 267–296). Amsterdam, The Netherlands: John Benjamins.

Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning 67*(2), 348–393.

Bowen, G. (2006). Grounded theory and sensitizing concepts. *International Journal of Qualitative Methods*, *5*, 12–23.

Brabazon, T. (2006). The Google effect. *Libri*, *56*(3), 157–167.

Brabazon, T. (2013). *Digital dieting: From information obesity to intellectual fitness*. London, UK & New York, NY: Routledge Taylor & Francis Group.

British Library. (n.d.). EThOS Toolkit | Re-use by external services: EThOS as a data provider: Metadata. Retrieved from http://ethostoolkit.cranfield.ac.uk/tiki-index.php?page=Re-use+by+external+services

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

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences*, *2*(2), 141–178.

Breslow, L., Pritchard, D. E., DeBoer, J., Stump, G. S., Ho, A. D., & Seaton, D. T. (2013). Studying learning in the worldwide classroom: Research into edX's first MOOC. *Journal of Research & Practice in Assessment*, *8*, 13–25.

Budapest Open Access Initiative (BOAI). (2002). BOAI declaration. Retrieved from http://www.budapestopenaccessinitiative.org/read

Burkhardt, H. (2006). From design research to large-scale impact. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieeven (Eds.), *Educational design research* (pp. 185–228). London, UK: Routledge.

Campbell, C., Pitt, L., Parent, M., & Berthon, P. (2011). Understanding consumer conversations around ads in a web 2.0 world. *Journal of Advertising*, *40*(1), 87–102.

Campbell, L. (2013, June 5). What do FutureLearn's terms and conditions say about open content? [Blog post]. Retrieved from http://blogs.cetis.org.uk/lmc/2013/06/05/what-do-futurelearns-terms-and-conditions-say-about-open-content/

Cape Town Open Education Declaration. (2007). Cape Town open education declaration: Unlocking the promise of open educational resources. Retrieved from http://www.capetowndeclaration.org/read-the-declaration

Carroll, J. B. (1964). Words, meanings, & concepts. *Harvard Educational Review*, *34*(2), 178–202.

Castilho, S., Gaspari, F., Moorkens, J., & Way, A. (2017). Integrating machine translation into MOOCs. *Proceedings of the EDULEARN 2017 conference* (pp. 9,360–9,365). Barcelona, Spain.

Chalmers, I. (2003). Trying to do more good than harm in policy and practice: the role of rigorous, transparent, up-to-date evaluations. *The Annals of the American Academy of Political Science*, *589*(1), 22–40. doi:10.1177/ 0002716203254762

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

Chang, J.-Y. (2014). The use of general and specialized corpora as reference sources for academic English writing: A case study. *ReCALL*, *26*(2), 243–259.

Charles, M. (2012). Proper vocabulary and juicy collocations: EAP students evaluate do-it-yourself corpus-building. *English for Specific Purposes*, *31*(2), 93–102.

Charles, M. (2015). Same task, different corpus: The role of personal corpora in EAP classes. In A. Lenko-Szymanska, & A. Boulton (Eds.) *Multiple affordances of language corpora for data-driven learning* (pp. 131–153). Amsterdam, The Netherlands: John Benjamins.

Charles, M., & Pecorari, D. (2016). *Introducing English for academic purposes*. London, UK: Routledge.

Chen, H. H. (2011) Developing and evaluating a web-based collocation retrieval tool for EFL students and teachers. *Computer Assisted Language Learning*, *24*(1), 59–76.

Chemers, M. M., Hu, L., & Garcia, B. F. (2001). Academic self-efficacy and first year college student performance and adjustment. *Journal of Educational Psychology*, *93*(1), 55–64. doi:10.1037/0022-0663.93.1.55

Chesbrough, H., Vanhaverbeke, W., & West, J. (2008). *Open innovation: Researching a new paradigm*. Oxford, UK: Oxford University Press.

Chinnery, G. (2008). You've got some GALL: Google-assisted language learning. *Language Learning & Technology, 12*(1), 3–11.

Christensen, C.M. (1997). *The innovator's dilemma: When new technologies cause great firms to fail*. Boston, MA: Harvard Business School Press.

Christensen, C. M., Raynor, M., & McDonald, R. (December 2015): What is disruptive innovation? *Harvard Business Review 93*(12), 44–53.

Christensen, G., Steinmetz, A., Alcorn, B., Bennett, A., Woods, D., & Emanuel, E.J. (November 6, 2013). The MOOC phenomenon: Who takes Massive Open Online Courses and why? Retrieved from http://ssrn.com/abstract=2350964

Chuang, I., & Ho, A. D. (2016). HarvardX and MITx: Four years of open online courses -- Fall 2012–Summer 2016. Retrieved from SSRN: https://ssrn.com/abstract=2889436 or http://dx.doi.org/10.2139/ssrn.2889436

Clifford, J. (1990). Notes on (field)notes. In R. Sanjek (Ed.), *Fieldnotes: The makings of anthropology* (pp. 47–70). Ithaca, NY: Cornell University Press.

Cobb, P., Confrey, J., di Sessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, *32*(1), 9–13.

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

Cobb, T. (1999). Applying constructivism: A test for the learner-as-scientist. *Educational Technology Research & Development*, *47*(3), 15–31.

Cobb, T. (2006). Internet and literacy in the developing world: Delivering the teacher with the text. *Educational Technology Research & Development*, *54*(6), 627–645.

Cobb, T., & Boulton, A. (2015). Classroom applications of corpus analysis. In D. Biber & R. Reppen (Eds.), *Cambridge handbook of English corpus linguistics* (pp. 478–497). Cambridge, UK: Cambridge University Press.

Collins, A. (1992). Toward a Design Science of Education. In E. Scanlon & T. O'Shea (Eds.), *New directions in educational technology* (pp. 15–22). New York, NY: Springer-Verlag.

Colpaert, J. (2004). Transdisciplinarity. *Computer Assisted Language Learning*, *17*(5), 459–472.

Colpaert, J. (2015). Reflections on present and future: towards an ontological approach to LMOOCs. In E. Martín-Monje & E. Bárcena (Eds.) *Language MOOCs. Providing learning, transcending boundaries* (pp. 161–172). Berlin, Germany: De Gruyter Open. Retrieved from https://www.degruyter.com/downloadpdf/books/9783110422504/9783110422504.10/9783110422504.10.pdf

Colpaert, J. (2016). Big content in an educational engineering approach. *Journal of Technology and Chinese Language Teaching*, *7*(1), 1–14. Retrieved from http://tclt.us/journal/2016v7n1/colpaert.pdf

Colpaert, J. (2018): Transdisciplinarity revisited. *Computer Assisted Language Learning*, *31*(5-6), 483–489 DOI: 10.1080/09588221.2018.1437111

COnnectedREpositories (CORE): Aggregating the world's open access papers. (n.d.). Retrieved from https://core.ac.uk/

Cost of Knowledge. (n.d.). *Elsevier statement*. Retrieved from https://gowers.files.wordpress.com/2012/02/elsevierstatementfinal.pdf

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*(2), 213–238.

Coxhead, A., & Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Second Language Writing*, *16*(3), 129–147.

Crabtree, A., Rouncefield, M., & Tolmie, P. (2012). *Doing design ethnography*. London, UK: Springer.

Creative Commons. (2015). *State of the commons*. Retrieved from https://stateof.creativecommons.org/2015/cc-sotc-2015-xx12.html

Daniel, J. (2013). Rankings and online learning: a disruptive combination for higher education? In P. T. M. Marope, P. J. Wells, & E. Hazelkorn (Eds), *Rankings and accountability in higher education: Uses and misuses* (pp. 95–112). Paris, France: UNESCO

Daskalovska, N. (2015). Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning*, *28*(2), 130–144. doi: 10.1080/09588221.2013.803982

De Groot, F. O. (2017). Tracing the potential of out-of-class digitally mediated language learning practice back to the classroom: A nexus of practice perspective. In M. Carrier, R. M. Damerow, and K. M. Bailey (Eds.), *Digital language learning and teaching* (pp 25–37). New York, NY: Routledge.

Dede, C. (2004). If design-based research is the answer, what is the question? A commentary on Collins, Joseph, and Bielaczyc; diSessa and Cobb; and Fishman, Marx, Blumenthal, Krajcik, and Soloway in the JLS special issue on design-based research. *Journal of the Learning Sciences*, *13*(1), 105–114. doi: 10.1207/s15327809jls1301_5

Design-Based Research Collective (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 32*(1), 5–8.

Digital commons (economics). (2016, March 13). In *Wikipedia, the free encyclopaedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=Digital_commons_(economics)&oldid=709892273

Ding, A., & Bruce, I. (2017). *The English for academic purposes practitioner: Operating on the edge of academia*. Cham, Switzerland: Palgrave Macmillan.

Downes, S. (2004). Nine rules for good technology. In *The learning marketplace: Meaning, metadata, and content syndication in the learning object economy* (pp. 11–15). Retrieved from http://www.downes.ca/files/book3.htm

Downes, S. (2007). Models for sustainable open educational resources. *Interdisciplinary Journal of Knowledge and Learning Objects*, *3*. Retrieved from http://ijklo.org/Volume3/IJKLOv3p029- 044Downes.pdf

Drouin, P. (2003). Term extraction using non-technical corpora as a point of leverage. *Terminology*, *9*(1), 99–117.

Engeström, Y. (2009). The future of activity theory: A rough draft. In A. Sannino, H. Daniels, & K. D. Gutierrez (Eds), *Learning and expanding with activity theory* (pp. 303–329). Cambridge, UK: Cambridge University Press.

Eaglestone, B., Ford, N., Brown, G.J., & Moore, A. (2007). Information systems and creativity: an empirical study. *Journal of Documentation*, *63*(4), 443–464.

Edwards, L., Hilty, R., Hugenholtz, B., Klimpel, P., Kreutzer, T., Lynn, J., Manara, C., Perez, N., & Vivant, M. (2012). Intellectual property and innovation: A framework for 21st century growth and jobs. In I. Hargreaves & P. Hofheinz (Eds.), *Lisbon Council Policy Brief* (Vol. VI, No. 2) (pp. 1-55). ISSN 2031 0943 Retrieved from http://www.lisboncouncil.net/publication/publication/84-intellectual-property-and-innovation-a-framework-for-21st-century-growth-and-jobs-.html

Elliott, R., Fischer, C. T., & Rennie, D. L. (1999). Evolving guidelines for publication of qualitative research studies in psychology and related fields. *Evolving Guidelines for Publication of Qualitative Research Studies in Psychology and Related Fields*, *38*(3), 215–229.

Emery, M., & Devane, T. (2007) Participative design workshop. In P. Holman, T. Devane, S. Cady (Eds.), *The change handbook: The definitive resource on today's best methods for engaging whole systems* (pp. 419– 435). San Francisco, CA: Berrett-Koehler.

Epstein, K. (2012). Academic Spring sees widening boycott of Elsevier. *BMJ News*, *344*(7846), 4–5.

Ericsson, K., & Simon, H. (1987). Verbal reports on thinking. In C. Faerch & G. Kasper (Eds.), *Introspection in second language research* (pp. 24–54). Clevedon, UK: Multilingual Matters.

ETIC. (1975). *English for academic study: Problems and perspectives*. London, UK: British Council.

Farrow, R., Pitt, R., de los Arcos, B., Perryman, L.-A., Weller, M., & McAndrew, P. (2015). Impact of OER use on teaching and learning: data from OER Research Hub (2013–2014). *British Journal of Educational Technology*, *46*(5), 972–976.

Finch Group. (2012). Accessibility, sustainability, excellence: how to expand access to research publications. *Report of the Working Group on Expanding Access to Published Research Findings*. Retrieved from http://www.researchinfonet.org/publish/finch/

Fisher, W. W. (2014a). CopyrightX: Lecture 12.3, Remedies: Criminal Penalties. Retrieved from https://www.youtube.com/watch?v=x6W8gk9fOFk&feature=youtube_gdata_player

Fisher, W. W. (2014b). HLS1X: CopyrightX course report. HarvardX working paper series No. 5. Available at SSRN: http://ssrn.com/abstract=2382332 or http://dx.doi.org/10.2139/ssrn.2382332

Fitzgerald, A. (2013a). *Openness in English for academic purposes*. Open Educational Resources case study with Durham University: Pedagogical development from OER practice. Commissioned by the Higher Education Academy (HEA) and the Joint Information Systems Committee (JISC), United Kingdom. Retrieved from https://www.heacademy.ac.uk/knowledge-hub/openness-english-academic-purposes

Fitzgerald, A. (2013b). *TOETOE International: FLAX weaving with Oxford Open Educational Resources*. Open Educational Resources International case study with the University of Oxford. Commissioned by the Higher Education Academy (HEA) and the Joint Information Systems Committee (JISC), United Kingdom. Retrieved from https://www.heacademy.ac.uk/knowledge-hub/toetoe-international-flax-weaving-oxford-open-education-resources

Fitzgerald, A. (2013c, February 14). Love is a stranger in an open car to tempt you in and drive you far away … toward OEP [Blog post] *TOETOE Technology for Open English Toying with Open E-resources (ˈtɔɪtɔɪ)*. Retrieved from http://www.alannahfitzgerald.org/love-is-a-stranger-in-an-open-car-who-tempts-you-in-and-drives-you-far-away/

Fitzgerald, A., Wu, S., & Barge, M. (2014). Investigating an open methodology for designing domain-specific language collections. In S. Jager, L. Bradley, E. J. Meima, & S. Thouësny (Eds.), *CALL design: Principles and practice; Proceedings of the 2014 EUROCALL Conference, Groningen, The Netherlands* (pp. 88–95). Dublin, Ireland: Research-publishing.net.

Fitzgerald, A., Wu, S., & Marín, M. J. (2015). FLAX: Flexible and open corpus-based language collections development. In K. Borthwick, E. Corradini, & A. Dickens (Eds.), *10 years of the LLAS elearning symposium: Case studies in good practice* (pp. 215–227). Dublin, Ireland: Research-publishing.net.

Fitzgerald, A. (2015, September 14). Rebalancing English language education: Access, materials writing and copyright. [Blog post] *ELTjam*. Retrieved from http://eltjam.com/rebalancing-english-language-education-access-materials-writing-and-copyright/

Fitzgerald, A., Marín, M. J., Wu. S., & Witten, I. H. (2017). Evaluating the efficacy of the digital commons for scaling data-driven learning. In M. Carrier, R. Damerow, K. Bailey (Eds.), *Digital language learning and teaching: Research, theory and practice* (pp. 38–51). New York, NY: Routledge & TIRF.

Fitzgerald, A. (2017). MOOC linguistic support. Data files available at https://osf.io/juakn/

Fitzgerald, A., Wu, S., König, J., Witten, I. H., & Shaw, S. (Expected 2020). Designing and evaluating an automated open data-driven language learning support system for MOOCs.

Fitzgerald, A., Wu, S. & Witten, I.H. (Expected 2020). Reflections on remixing open access content for data-driven language learning systems design in higher education.

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

Flowerdew, J., & Peacock, M. (2001). Issues in EAP: A preliminary perspective. In J. Flowerdew & M. Peacock (Eds.), *Research perspectives on English for academic purposes* (pp. 8-24). Cambridge, UK: Cambridge University Press.

Flowerdew L. (2009). Applying corpus linguistics to pedagogy: A critical evaluation. *International Journal of Corpus Linguistics*, *14*(3), 393–417.

Ferguson, G., Pérez-Llantada, C., & Plo, R. (2011). English as an international language of scientific publication: A study of attitudes. *World Englishes*, *30*(1), 41–59.

Franken, M. (2014). The nature and scope of student search strategies in using a web derived corpus for writing. *The Language Learning Journal*, *42*(1), 85–102.

Frankenberg-Garcia, A. (2005). A peek into what today's language learners as researchers actually do. *International Journal of Lexicography*, *18*, 335–55.

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

Gatt, C., & Ingold, T. (2013). From description to correspondence: anthropology in real time. In W. Gunn, T. Otto, & R. C. Smith (Eds), *Design anthropology: Theory and practice* (pp. 139–158). London, UK: Bloomsbury Academic.

Geertz, C. (1973). *The interpretation of cultures: selected essays*. New York, NY: Basic Books.

Geiller, I. (2014). How EFL students can use Google to correct their "untreatable" written errors. *The Eurocall Review*, *22*(2), 26–45.

Geluso, J., & Yamaguchi, A. (2014). Discovering formulaic language through data-driven learning: Student attitudes and efficacy. *ReCALL*, *26*(2), 225–242. doi: 10.1017/S0958344014000044

Genzuk, M. (2003). A synthesis of ethnographic research. *Occasional Papers Series*. Center for Multilingual, Multicultural Research. Los Angeles, CA: Rossier School of Education, University of Southern California.

Gibson, W. (1999, November 30). The science in science fiction. *Talk of the Nation*, NPR. Retrieved from http://www.npr.org/templates/story/story.php?storyId=1067220

Gillani, N., & Eynon, R. (2014). Communication patterns in massively open online courses. *Internet and Higher Education*, *23*, 18–26.

Gillett, A. (2018, April 13). Is EAP ESP? [Blog post] *Uefap*. Retrieved from http://www.uefap.net/blog/?p=933

Godwin-Jones, R. (2012). Challenging hegemonies in online learning. *Language Learning & Technology*, *16*(2), 4–13. Retrieved from http://llt.msu.edu/issues/june2012/emerging.pdf

Gonzalez, G. H., Tahsin, T., Goodale, B. C., Greene, A. C., & Greene, C. S. (2016). Recent advances and emerging applications in text and data mining for biomedical discovery. *Brief Bioinform*, *17*(1), 33–42.

Graddol. D. (2006). *English Next – why English as a global language may mean the end of 'English as a Foreign Language'*. London, UK: The British Council.

Graham, C. R. (2006). Blended learning systems: Definition, current trends, and future directions. In C. J. Bonk & C. R. Graham (Eds.), *The handbook of blended learning: Global perspectives, local designs* (pp. 3–21). San Francisco, CA: Pfeiffer.

Graham, C. R. (2012). Emerging practice and research in blended learning. In M. G. Moore (Ed.), *Handbook of distance education* (pp. 333–350). New York, NY: Routledge.

Green, C. (2015, June 2). edX makes it easy for authors to share under Creative Commons [Blog post]. Retrieved from https://creativecommons.org/2015/06/02/edx-makes-it-easy-for-authors-to-share-under-creative-commons/

Groom, J., & Lamb, B. (2014). Reclaiming innovation. *EDUCAUSE Review*, *49*(3). Retrieved from https://www.educause.edu/visuals/shared/er/extras/2014/ReclaimingInnovation/default.html

Groot, P. J. M. (2000). Computer assisted second language vocabulary acquisition. *Language Learning & Technology*, *4*(1), 60–81.

Gunn, W., & Donovan, J. (2013). *Design and anthropology*. Burlington, VT: Ashgate.

Gunn, W., Otto, T., & Smith, R.C. (2013). *Design anthropology: Theory and practice*. London, UK: Bloomsbury Academic.

Guo, P., & Reinecke, K. (2014). Demographic differences in how students navigate through MOOCs. *Proceedings of the first ACM conference on Learning @ Scale* (pp. 21–30). New York, NY: ACM.

Hackin, T. (2001). Abstracting from abstracts. In M. Hewings (Ed.), *Academic writing in context: Implications and applications* (pp. 93–103). Birmingham, UK: Birmingham University Press.

Hadley, G. (2015). *English for Academic Purposes in neoliberal universities: A critical grounded theory* (Vol. 22). Cham, Switzerland: Springer International Publishing. Retrieved from http://link.springer.com/10.1007/978-3-319-10449-2

Hafner, C. A., & Candlin, C. N. (2007). Corpus tools as an affordance to learning in professional legal education. *English for Academic Purposes*, *6*(4), 303–318.

Hake, R. (1998). Interactive engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*(1), 64–74.

Hakkarainen, P. (2009). Designing and implementing a PBL course on educational digital video production: Lessons learned from a design- based research. *Educational Technology, Research and Development*, 57(2), 211–228. doi: 10.1007/s11423- 007- 9039- 4

Hargreaves, I. (2011). *Digital opportunity – A review of intellectual property and growth*. London, UK: HM Government. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/32563/ipreview-finalreport.pdf

Hearst, M. (1999). Untangling text data mining. In R. Dale & K. Church (Eds.), *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics* (pp. 3–10). Morristown, NJ: ACL. doi:10.3115/1034678.1034679

Hearst, M. A. (2009). *Search user interfaces*. Cambridge, UK: Cambridge University Press.

Herrington, J., McKenney, S., Reeves, C. T., & Oliver, R. (2007) Design-based research and doctoral students: Guidelines for preparing a dissertation proposal. Retrieved from http://ro.ecu.edu.au/cgi/viewcontent.cgi?article=2611&context=ecuworks

Hill, J. (1999). Collocational competence. *ETP (English Teaching Professional), 11*, 1–6.

Hilton III, J. L., & Wiley, D. (2011). Open access textbooks and financial sustainability: A case study on flat world knowledge. *The International Review of Research in Open and Distributed Learning, 12*(5), 18–26.

Hilton III, J. (2016). Open educational resources and college textbook choices: A review of research on efficacy and perceptions. *Educational Technology Research & Development*, *64*(4), 573–590. https://doi.org/10.1007/s11423-016-9434-9

Ho, A. D., Chuang, I., Reich, J., Coleman, C. A., Whitehill, J., Northcutt, C. G., William, J. J., Hansen, J. D., Lopez, G., & Petersen, R. (2015). HarvardX and MITx: Two years of open online courses Fall 2012-Summer 2014 (SSRN Scholarly Paper No. ID 2586847). Rochester, NY: Social Science Research Network. Retrieved from: http://papers.ssrn.com/abstract=2586847

Howatt, A. P. R. (2004). *A history of English language teaching* (2nd ed.) Oxford, UK: Oxford University Press.

Hughes, J. A., Randall, D., & Shapiro, D. (1992). Faltering from ethnography to design. In *Proceedings of ACM 1992 Conference on Computer-Supported Cooperative Work: Sharing Perspectives* (pp. 115–123). New York, NY: ACM Press.

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge, UK: Cambridge University Press.

Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. Harlow, UK: Longman.

Hyland, K. (2002). Specificity revisited: How far should we go now? *English for Specific Purposes*, *21*, 385–395.

Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. London, UK: Continuum.

Hyland, K. (2006). *English for Academic Purposes: An advanced resource book*. London, UK: Routledge.

Hyland, K. (2008a). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics, 18*(1), 41–62.

Hyland, K. (2008b). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes, 27*(1), 4–21.

Hyland, K. (2018). Sympathy for the devil: A defense of EAP. *Language Teaching, 51*(3), 383–399.

Hyland, K. (2019). Participation in publishing: The demoralising discourse of disadvantage. In P. Habibie & K. Hyland (Eds.), *Novice writers and scholarly publication: Authors, mentors, gatekeepers* (pp. 1–33). Cham, Switzerland: Palgrave Macmillan.

IT Services, University of Oxford. (n.d.). [Oxford Text Archive] [*BAWE* Terms and Conditions Text]. Retrieved from http://ota.ox.ac.uk/scripts/download.php?otaid=2539

Johns, T. (1994). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. In T. Odline (Ed.), *Perspectives on pedagogical grammar* (pp. 293–313). Cambridge, UK: Cambridge University Press.

Johns, T. (1993). Data-driven learning: An update. *TELL&CALL, 2*, 4–10.

Johns, T. (1997). Contexts: The background, development and trialling of a concordance-based CALL program. In A. Wichmann, S. Figelston, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 100–115). London, UK: Longman.

Johns, T. (2002). Data-driven learning: The perpetual challenge. In B. Kettemann & G. Marko (Eds.), *Proceedings of Teaching and Learning by Doing Corpus Analysis* (pp.107–117). The Fourth International Conference on Teaching and Language Corpora, Graz 19–24 July 2000. Amsterdam, The Netherlands: Rodopi.

Kalman, Y. (2014). A race to the bottom: MOOCs and higher education business models. *Open Learning: The Journal of Open, Distance and eLearning, 29*(1), 5–14.

Kelty, C. M. (2008). *Two bits: The cultural significance of free software*. Durham, NC: Duke University Press.

Kilbourn, K. (2013). Tools of movement and engagement: design anthropology's style of knowing. In W. Gunn, T. Otto, R. C. Smith (Eds.), *Design anthropology: Theory and practice* (pp. 68–82). London, UK: Bloomsbury Academic.

Kilgariff, A., & Grefenstette, G. (2003). Introduction to the social issue on the web as corpus. *Computational Linguistics, 29*(3), 333–347.

Kincheloe, J. L. (2005). On to the next level: Continuing the conceptualization of the bricolage. *Qualitative Inquiry, 11*(3), 323–350.

King, M. (2007). General principles of user-oriented evaluation. In L. Dybkjær, H. Hemsen & W. Minker (Eds). *Evaluation of text and speech systems* (pp. 125–161). Berlin/Heidelberg, Germany/New York, NY: Springer.

Kizilcec, R., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: Analysing learner subpopulations in Massive Open Online Courses. *Proceedings of the third international conference on learning analytics and knowledge* (pp. 170–179). New York, NY: ACM.

Knoth, P., & Zdrahal, Z. (2012). CORE: three access levels to underpin open access. *D-Lib Magazine, 18*(11/12), article number 4.

Knox, J. (2013). The limitations of access alone: Moving towards open processes in education technology. *Open Praxis, 5*(1), 21–29.

Koseoglu, S., & Bozkurt, A. (2018). An exploratory literature review on open educational practices, *Distance Education*, *39*(4), 441–461.

Kovanović, V., Joksimović, S., Gašević, D., Siemens, G., & Hatala, M. (2015), What public media reveals about MOOCs: A systematic analysis of news reports. *British Journal of Educational Technology*, *46*(3), 510–527. doi:10.1111/bjet.12277

Kuper, A., Lingard, L., & Levinson, W. (2008). Critically appraising qualitative research. *BMJ, 337*(7671), 687–689. https://doi.org/10.1136/bmj.a1035.

Laakso, M., Welling, P., Bukvova H., Nyman, L., Björk, B-C., & Hedlund, T. (2011). The development of open access journal publishing from 1993 to 2009. *PLoS One, 6*(6), e20961. 10.1371/journal.pone.0020961

LeCompte, M., & Schensul, J. (1999). *Analyzing and interpreting ethnographic data*. Walnut Creek, CA: AltaMira Press.

Lederman, D. (2017, April 28). Clay Christensen, doubling down. *Inside Higher Ed*. Retrieved from https://www.insidehighered.com/digital-learning/article/2017/04/28/clay-christensen-sticks-predictions-massive-college-closures

Lepore, J. (2014, June 23). The disruption machine: What the gospel of innovation gets wrong. *The New Yorker.* Retrieved from https://www.newyorker.com/magazine/2014/06/23/the-disruption-machine

Lessig, L. (2004). *Free culture: How big media uses technology and the law to lock down culture and control creativity*. New York, NY: Penguin.

Lessig, L. (2008). *Remix: Making art and commerce thrive in the hybrid economy*. New York, NY: Penguin Press.

Levin, B. (2011). Mobilising research knowledge in education. *London Review of Education*, *9*(1), 15–26. doi:10.1080/14748460.2011.550431.

Levy, M. (1997). *CALL: Context and conceptualisation*. Oxford: Oxford University Press.

Li, L., Franken, M., & Wu, S. (2017). Bundle-driven metadiscourse analysis: Sentence initial bundles in Chinese and New Zealand postgraduates' thesis writing. In C. Hatipoglu, E. Akbas & Y. Bayyurt (Eds.), *Metadiscourse in written genres: Uncovering textual and interactional aspects of texts*. Amsterdam, The Netherlands: Peter Lang Publishing.

Liao, S., & Lei, L. (2017). What we talk about when we talk about corpus: A bibliometric analysis of corpus-related research in linguistics (2000-2015). *Glottometrics*, *38*, 1–20.

Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. London, UK: Sage Publications Inc.

Littlejohn, A., Hood, N., Milligan, C., & Mustain, P. (2016). Learning in MOOCs: Motivations and self-regulated learning in MOOCs. *Internet and Higher Education*, *29*, 40–48.

Littlejohn, A., & Hood, N. (2017). How educators build knowledge and expand their practice: The case of open education resources. *British Journal of Educational Technology*, *48*(2), 499–510. doi: https://doi.org/10.1111/bjet.12438

Liyanagunawardena, T., Adams, A., & Williams, S. (2013). MOOCs: A systematic study of the published literature 2008–2012. *The International Review of Research in Open and Distance Learning*, *14*(3), 202–227.

Lock, S. (1988). Structured abstracts. *BMJ: British Medical Journal*, *297*(66420, 156.

Lockley, P. (2015, December 9). Re: Open policies for MOOCs [Discussion-list]. JISCMail - OER-DISCUSS Archives. Retrieved from https://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind1512&L=OER-DISCUSS&F=&S=&P=16754\

Lugmayr, A., Stockleben, B., Zou, Y., Anzenhofer, S., & Jalonen, M. (2014). Applying "design thinking" in the context of media management education. *Multimedia Tools and Applications*, *71*(1), 119–157.

Lyotard, J-F. (1984). *The Postmodern Condition: A Report on Knowledge.* Minneapolis, MN: University of Minnesota Press.

Malhotra, A., Younesi, E., Gurulingappa, H., & Hofmann-Apitius, M. (2013). 'HypothesisFinder:' a strategy for the detection of speculative statements in scientific text. *PLoS Computational Biololgy*, *9*(7), e1003117.

Marín, M. J., & Rea, C. (2014). Assessing four automatic term recognition methods: Are they domain-dependent? *English for Specific Purposes World*, *42*(15), 1–27.

Marín, M. J. (2014). Evaluation of five single-word term recognition methods on a legal corpus. *Corpora*, *9*(1), 83–107.

Marin, M.J., Ortis Llopaz, M., & Fitzgerald, A. (2017). A Data-driven learning experiment in the legal English classroom using the FLAX platform. *Miscelánea: A Journal of English and American Studies*, *55*, 37–64.

Martin, M. (1984). Advanced vocabulary teaching: The problem of synonyms. *Modern Language Journal*, *69*(2), 130–137.

McAlpine, J., & Myles, J. (2003). Capturing phraseology in an online dictionary for advanced users of English as a second language: a response to user needs. *System*, *31*(1), 71–84.

McEnery, T., & Wilson, A. (1997). Teaching and language corpora. *ReCALL*, *9*(1), 5–14.

McEnery, T., & Wilson, A. (2001). *Corpus linguistics.* Edinburgh, UK: Edinburgh University Press.

McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: an advanced resource book*. London, UK: Routledge.

McGill, L., Falconer, I., Dempster, J.A., Littlejohn, A., & Beetham, H. (2013). OER journeys: HEFCE OER review final report. JISC. Retrieved from https://oersynth.pbworks.com/w/page/60338879/HEFCE-OER-Review-Final-Report

McGill, L., Beetham, H., Falconer, I., & Littlejohn, A. (2010) JISC/HEA OER programme: Pilot phase synthesis and evaluation report. Retrieved from https://oersynth.pbworks.com/w/page/29688444/Pilot-Phase-Synthesis-and-Evaluation-Report

McKenney, S., & Reeves, T. (2012). *Conducting Educational Design Research*. London, UK and New York, NY: Routledge.

McLuhan, M. (2005). Fordham University: First lecture (1967). In S. McLuhan and D. Staines (Eds.), *Understanding me: Lectures and interviews*. Cambridge, UK: MIT Press

McMillan Cottom, T. (2015). Intersectionality and critical engagement with the Internet. In U. Safiya, N. Tynes & B. Tynes (Eds.), *The intersectional Internet: Race, sex, class, and culture online.* Amsterdam, Netherlands: Peter Lang Publishing.

McSherry, C. (2015, October 16). Big win for fair use in Google books lawsuit. *Electronic Frontier Foundation*. Retrieved from https://www.eff.org/deeplinks/2015/10/big-win-fair-use-google-books-lawsuit

Mellinkoff, D. (1963). *The language of the law*. Boston, MA: Little, Brown & Co.

Merlucci, A. (1996). *Challenging codes: Collective action in the information age*. Cambridge, UK: Cambridge University Press.

Meyers, C., & Jones, T. (1993). *Promoting active learning: Strategies for the college classroom.* San Francisco, CA: Jossey-Bass.

Mezynski, K. (1983). Issues concerning the acquisition of knowledge: Effects of vocabulary training on reading comprehension. *Review of Educational Research*, *53*(2), 253–279.

Milligan, C., Littlejohn, A., & Margaryon, A. (2013). Patterns of engagement in connectivist MOOCs. *Merlot Journal of Online Learning and Teaching*, *9*(2), 149–159.

Milne, D., & Witten, I. H. (2013). An open-source toolkit for mining Wikipedia. *Artificial Intelligence*, *194*, 222–239.

Moxley, J. (2018, March). Letter from the founder. *Writing Commons*. Retrieved from

https://writingcommons.org/component/acymailing/listid-225-master-wc-list/mailid-105-writing-commons-march-2018

Natarajan, J., Berrar, D., Dubitzky, W., Hack, C., Zhang, Y., DeSesa, C., Brocklyn, J. R. V., & Bremer, E. G. (2006). Text mining of full-text journal articles combined with gene expression analysis reveals a relationship between sphingosine-1-phosphate and invasiveness of a glioblastoma cell line. *BMC Bioinformatics*, *7*(1), 373.

Naughton, John. (2011). *The elusive technological future*. Keynote address given at the Association for Learning Technology Annual Conference. Leeds, UK. Retrieved from https://www.youtube.com/watch?v=yUXh-GPa5dI

Naur, P. (1983). Program development studies based on diaries. In T. Green, S. Payne, G. Veer (Eds.), *Psychology of computer use* (pp. 159–170). London, UK: Academic Press.

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., & Gardner S. (2012). *Genres across the disciplines: student writing in Higher Education*. Cambridge, UK: Cambridge University Press.

Nuijten, M.B., Hartgerink, C. H. J., van Assen, M. A. L. M, Epskamp, S., & Wicherts, J. M. (2015). The prevalence of statistical reporting errors in psychology (1985–2013). *Behavior Research Methods*, *48*(4), 1205–1226. ISSN 1554-351X, 1554-3528.

O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for reporting qualitative research: A synthesis of recommendations. *Academic Medicine, 89*(9), 1245–1251. https://doi.org/10.1097/ACM.0000000000000388.

O'Keeffe, A., McCarthy, M., & Carter R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge, UK: Cambridge University Press.

Okerson, A. (1991). With feathers: Effects of copyright and ownership on scholarly publishing. *College & Research Libraries, 52*(5), 425–38.

Onwuegbuzie, A. J., & Leech, N. L. (2005). On becoming a pragmatic researcher: The importance of combining quantitative and qualitative research methodologies. *International Journal of Social Research and Methodology, 8*(5), 375–387.

O'Reilly, T. (2005, September 30). What is Web 2.0? End of the software release cycle. *O'Reilly Media, Inc*. Retrieved from https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html?page=4

Orlikowski, W.J. (1991). Integrated information environment or matrix of control? The contradictory implications of information technology. *Accounting, Management and Information Technologies, 1*(1), 9–42.

O'Sullivan, I., & Chambers, A. (2006). Learners' writing skills in French: Corpus consultation and learner evaluation. *Journal of Second Language Writing, 15*(1), 49–68.

Otto, T., & Smith, R. C. (2013). Design anthropology: a distinct style of knowing. In W. Gunn, T. Otto, R.C. Smith (Eds.), *Design anthropology: Theory and practice* (pp. 1–29). London, UK: Bloomsbury Academic.

Pappano, L. (2012, November 2). The year of the MOOC. *The New York Times*. Retrieved from http://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html

Park, K. (2012). Learner-corpus interaction: A locus of microgenesis in corpus-assisted L2 writing. *Applied Linguistics, 33*(4), 361–385.

Parr, C. (2013, May 10). Not staying the course, *Times Higher Education.* Retrieved from https://www.insidehighered.com/news/2013/05/10/new-study-low-mooc-completion-rates.

Peachey, N. (2005). Concordancers in ELT. *British Council Teaching English*. Retrieved from http://www.teachingenglish.org.uk/think/articles/concordancers-elt

Peter, S., & Deimann, M. (2013). On the role of openness in education: A historical reconstruction. *Open Praxis, 5*(1), 7–14. doi: http://dx.doi.org/10.5944/openpraxis.5.1.23

Pérez-Paredes, P., Ordoñana Guillamón, C., & Aguado Jiménez, P. (2018): Language teachers' perceptions on the use of OER language processing technologies in MALL. *Computer Assisted Language Learning*, *31*(5-6), 522-545. DOI: 10.1080/09588221.2017.1418754

Rapoport, R. (1970). Three dilemmas of action research. *Human Relations*, *23*(6), 499–513.

Raymond, E. (1997, May 27). Release early, release often. *The cathedral and the bazaar: Musings on Linux and open source by an accidental revolutionary*. Retrieved from http://www.catb.org/~esr/writings/cathedral-bazaar/cathedral-bazaar/ar01s04.html

Reason, P., & Bradbury, H. (2007). *Handbook of action research*. London, UK: Sage.

Renkle, A., Atkinson, R., Maier, U. & Staley, R. (2002). From example study to problem solving: Smooth transitions help learning. *Journal of Experimental Education*, *70*(4), 293–315.

Reppen, R. (2010). *Using corpora in the language classroom.* Cambridge, UK: Cambridge University Press.

Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, *4*(2), 155–169. https://doi.org/10.1007/BF01405730

Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning*, *45*(2), 283–331.

Rodgers, O., Chambers, A., & Le Baron-Earle, F. (2011). Corpora in the LSP classroom: A learner-centred corpus of French for biotechnologists. *International Journal of Corpus Linguistics*,16(3), 391–411.

Römer, U. (2010). Using general and specialized corpora in English language teaching: Past, present and future. In M.C. Campoy-Cubillo, B. Belles-Fortuño, & L. Gea-Valor (Eds.), *Corpus-based Approaches to English Language Teaching* (pp. 18–38). London, UK: Continuum.

Römer, U. (2011). Corpus research applications in second language teaching. *Annual Review of Applied Linguistics 31*, 205–225.

Rott, S. 2004. A comparison of output interventions and un-enhanced reading conditions on vocabulary acquisition and text comprehension. *The Canadian Modern Language Review*, *61*(2), 169–202.

Salvador, T., Bell, G., & Anderson, K. (1999). Design ethnography. *Design Management Journal (Former Series)*, *10*(4), 35–41.

Sanjek, R. (2004). Going public: responsibilities and strategies in the aftermath of ethnography. *Human Organization*, *63*(4), 444–456.

Santiago-Delefosse, M., Gavin, A., Bruchez, C., Roux, P., & Stephen, S. L. (2016). Quality of qualitative research in the health sciences: Analysis of the common criteria present in 58 assessment guidelines by expert users. *Social Science & Medicine, 148*, 142–151. https://doi.org/10.1016/j.socscimed.2015.11.007.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge, UK: Cambridge University Press.

Schmitt, N., & Schmitt, D. (1995). Vocabulary notebooks: theoretical underpinnings and practical suggestions. *ELT Journal, 49*(2), 133–143.

Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt, and M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp.199–227). Cambridge, UK: Cambridge University Press.

Scott, M. (2008). *WordSmith Tools version 5*. Liverpool, UK: Lexical Analysis Software.

Seimens, G. (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology & Distance Learning*, *2*(1), 3–10.

Seimens, G. (2011, October 13). The race to platform education [Blog post]. *elearnspace.* Retrieved from http://www.elearnspace.org/blog/2011/10/13/the-race-to-platform-education/

Seimens, G. (2012, July 25). MOOCs are really a platform [Blog post]. *elearnspace*. Retrieved from http://www.elearnspace.org/blog/2012/07/25/moocs-are-really-a-platform/

Selwyn, N. (2014). Data entry: towards the critical study of digital data and education. *Learning, Media and Technology*, *40*(1), 64–82. https://doi.org/10.1080/17439884.2014.921628

Selwyn, N. (2015). Minding our language: why education and technology is full of bullshit … and what might be done about it. *Learning, Media and Technology*, *41*(3), 437–443. http://doi.org/10.1080/17439884.2015.1012523

Shah, D. (2016, December 29). Monetization over massiveness: Breaking down MOOCs by the numbers in 2016. *EdSurge*. Retrieved from https://www.edsurge.com/news/2016-12-29-monetization-over-massiveness-breaking-down-moocs-by-the-numbers-in-2016

Shah, D. (2018a, December 11). By the numbers: MOOCs in 2018. Retrieved from https://www.class-central.com/report/mooc-stats-2018/

Shah, D. (2018b, January 22). A product at every price: A review of MOOC stats and trends in 2017. *Class Central*. Retrieved from https://www.class-central.com/report/moocs-stats-and-trends-2017/

Shah, D. (2018c, June 3). The second wave of MOOC hype is here and it's online degrees. Retrieved from https://www.class-central.com/report/second-wave-of-mooc-hype/

Shah D. (2018d, January 17). MOOC trends in 2017: Content paywalls. *Class Central*. Retrieved from https://www.class-central.com/report/mooc-trends-content-paywalls/

Shapiro, G. (2006, October 26). A rebellion erupts over journals of academia. *New York The Sun*. Retrieved from https://www.nysun.com/arts/rebellion-erupts-over-journals-of-academia/42317/

Shapiro, H. B., Lee, C. H., Wyman Roth, N. E., Li, K., Çetinkaya-Rundel, M., & Canelas, D. A. (2017). Understanding the massive open online course (MOOC) student experience: An examination of attitudes, motivations, and barriers. *Computers & Education*, *110*, 35–50.

Shaw, G. B. (1913). Getting married: A disquisitory play. London, UK: Constable & Company Ltd.

Shei, C. C. (2008). Discovering the hidden treasure on the Internet: using Google to uncover the veil of phraseology. *Computer Assisted Language Learning*, *21*(1), 67–85.

Shneiderman, B. & Plaisant, C. (2004). *Designing the user interface: strategies for effective human–computer interaction*. Boston, MA: Addison-Wesley Longman.

Simon, H.A. (1996). *The sciences of the artificial*. Cambridge, Mass: MIT Press.

Sinclair, J. McH. (2004a). *How to use corpora in language teaching*. Amsterdam, The Netherlands: John Benjamins.

Sinclair, J. McH. (2004b). *Trust the text: Language, corpus and discourse*. London, UK: Routledge.

Smith, A. E. (2000a). Machine learning of well-defined thesaurus concepts. In A.-H. Tan & P. S. Yu (Eds.), *Proceedings of the International Workshop on Text and Web Mining (PRICAI 2000)* (pp. 72–79). Melbourne: *PRICAI.*

Smith, A. E. (2000b). Machine mapping of document collections: The Leximancer system. In *Proceedings of the Fifth Australasian Document Computing Symposium*. Sunshine Coast, Australia: DSTC.

Smith, A. E. (2003). Automatic extraction of semantic networks from text using Leximancer. In *HLT-NAACL 2003 Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics: Companion volume* (pp. Demo23- Demo24). Edmonton, Canada: ACL.

Smith, A. E., & Humphreys, M. S. (2006). Evaluation of unsupervised semantic mapping of natural language with Leximancer concept mapping. *Behavior Research Methods, 38*(2), 262–279.

Sosoni, V. (2017). Casting some light on experts' experience with translation crowdsourcing.

The Journal of Specialised Translation, 28, 362–384.

SPARC. (2018, March). Congress funds $5 million open textbook grant program in 2018 spending bill. Retrieved from https://sparcopen.org/news/2018/open-textbooks-fy18/

Stahl, S.A., & Fairbanks, M.M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research, 56*(1), 72–110.

Stallman, R.M. (2002). *Free software, free society: Selected essays of Richard M. Stallman*. J. Gay (Ed.). Boston, MA: The Free Software Foundation.

Stich, A.E., & Reeves, T.D. (2017). Massive open online courses and underserved students in the United States. *Internet and Higher Education, 32*, 58–71.

Strevens, P. (1988). ESP after twenty years: a reappraisal. In M. Tickoo (Ed.), *ESP: state of the art* (pp. 1–13). Singapore: SEAMEO Regional Language Centre.

Stubbs, M. (1996). *Text and corpus analysis*. Oxford, UK: Basil Blackwell.

Stubbs, M., & Barth, I. (2003) Using recurrent phrases as text-type discriminators. A quantitative method and some findings. *Functions of Language, 10*(1), 61–104.

Swales, J. M. (1985). *Episodes in ESP.* Oxford, UK: Pergamon Press.

Swales, J. (1990). *Genre analysis: English in academic and research settings.* Cambridge, UK: Cambridge University Press.

Swales, J. & Feak, C. (2009). *Abstracts and the writing of abstracts.* The Michigan series in English for academic and professional purposes. Ann Arbor, MI: University of Michigan Press.

Swan, A. (2012). *Policy guidelines for the development and promotion of open access*. Paris, France: UNESCO. Retrieved from http://unesdoc.unesco.org/images/0021/002158/215863e.pdf

Swartz, A. (2008). *Guerilla open access manifesto*. Eremo, Italy: The Internet Archive. Retrieved from https://archive.org/stream/GuerillaOpenAccessManifesto/Goamjuly2008_djvu.txt

Tennant, J. P., Waldner, F., Jacques, D. C., Masuzzo, P., Collister, L. B., & Hartgerink, C. H. J. (2016). The academic, economic and societal impacts of open access: An evidence-based review [version 2; referees: 4 approved, 1 approved with reservations]. *F1000Research*, 5:632. doi: 10.12688/f1000research.8460.2

Thomas, J. (2017). *Discovering English with Sketch Engine* (2nd Edition). Versatile.

Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam, The Netherlands: John Benjamins.

Tomlinson, B., & Masuhara, H. (2010). Published research on materials development for language learning. In B. Tomlinson & H. Masuhara (Eds.), *Research for materials development in language learning*. London, UK: Continuum.

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care*, *19*(6), 349–357. https://doi.org/. http://dx.doi.org/10.1093/intqhc/mzm042.

Tribble, C. (2015). Teaching and language corpora: Perspectives from a personal journey. In A. Lenko-Szymanska, & A. Boulton, A. (Eds.) *Multiple Affordances of Language Corpora for Data-driven Learning* (pp. 37–64). Amsterdam, The Netherlands: John Benjamins,

Turner, J. (2004). Language as academic purpose. *Journal of English for Academic Purposes*, *3*(2), 95–109.

UNESCO. (2008). Education for All by 2015. Will we make it? *Education for all global education monitoring report 2008.* United Nations Education Scientific Cultural Organisation. Retrieved from www.efareport.unesco.org

UNESCO. (2015). Education for All 2000-2015: Achievements and challenges. *Education for all global education monitoring report 2015.* United Nations Education Scientific Cultural Organisation. Retrieved from http://unesdoc.unesco.org/images/0023/002322/232205e.pdf

UNESCO. (2017/2018). Accountability in education: Meeting our commitments. *Education for all global education monitoring report 2017/2018.* United Nations Education Scientific Cultural Organisation. Retrieved from http://unesdoc.unesco.org/images/0025/002593/259338e.pdf

University of California Academic Senate. (2019, February 28). University of California Academic Council statement on the university's negotiations with Elsevier Publishing. Retrieved from https://senate.universityofcalifornia.edu/_files/reports/academic-council-statement-elsevier-feb28.pdf

University of California Office of the President. (2019, February 28). UC terminates subscriptions with world's largest scientific publisher in push for open access to publicly funded research. Retrieved from https://www.universityofcalifornia.edu/press-room/uc-terminates-subscriptions-worlds-largest-scientific-publisher-push-open-access-publicly

Uvalić-Trumbić, S., & Daniel, J. (2011). Let a thousand flowers bloom! *UNESCO global forum on rankings and accountability in higher education: Uses and misuses*. Retrieved from http://www.col.org/resources/speeches/2011presentation/Pages/2011-05-16.aspx

Varley, S. (2009) I'll just look that up in the concordancer: Integrating corpus consultation into the language learning environment. *Computer Assisted Language Learning*, *22*(2), 133–152.

Vollmer, T. (2012, November 1). Keeping MOOCs open [Blog post]. Retrieved from https://creativecommons.org/2012/11/01/keeping-moocs-open/

Vyatkina, N. (2016). Data-driven learning of collocations: Learning performance, proficiency, and perceptions. *Language Learning & Technology*, *20*(3), 159–179.

Wakkary, R. (2005). Framing complexity, design and experience: a reflective analysis. *Digital Creativity*, *16*(2), 65–78.

Walker, D. (2006). Towards productive design studies. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieeven (Eds.), *Educational design research* (pp. 9–18). London, UK: Routledge.

Wang, T. (2013, May 13). Big data needs thick data. Retrieved from http://ethnographymatters.net/blog/2013/05/13/big-data-needs-thick-data/

Watters, A. (2013, November 7). The education apocalypse. 2013 Open Education Conference Keynote Address. In *Hack Education* #opened13 [Blog post]. Retrieved from http://hackeducation.com/2013/11/07/the-education-apocalypse

Watters, A. (2014, November 16). From "open" to justice. 2014 OpenCon Keynote Address. In *Hack Education* #OpenCon14 [Blog post]. Retrieved from http://hackeducation.com/2014/11/16/from-open-to-justice

Watters, A. (2015, February 19). The history of the future of education [Blog post]. In *Hack Education*. Retrieved from http://hackeducation.com/2015/02/19/the-history-of-the-future-of-education

Watters, A. (2016). *The curse of the monsters of educational technology*. Tech Gypsies: Amazon Digital Services LLC.

Weber, R.P. (1990). *Basic content analysis*. (Second edition). Newbury Park, CA: Sage Publications.

Weick, K. (1995). Organizational redesign as improvisation. In G. P. Huber & W. H. Glick (Eds.), *Organizational change and redesign: Ideas and insights for improving performance*, (pp. 346-379). London, UK: Wiley-Blackwell.

Weller, M. (2011). *The digital scholar: How technology is changing scholarly practice*. London, UK: Bloomsbury Academic. doi: http://dx.doi.org/10.5040/9781849666275 http://www.bloomsburyacademic.com/view/DigitalScholar_9781849666275/book-ba-9781849666275.xml

Weller, M. (2014). *The battle for open: How openness won and why it doesn't feel like victory*. London, UK: Ubiquity Press. doi: http://dx.doi.org/10.5334/bam

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

West, R. (2018). Developing an open textbook for learning and instructional design technology. *Tech Trends, 63*(2), 226–235. https://doi.org/10.1007/s11528-018-0263-z

Whitaker, M. P. (1996). Ethnography as learning: A Wittgensteinian approach to writing ethnographic accounts. *Anthropological Quarterly*, 69(1), 1–13.

Widdowson, H. G. (1983). *Learning purpose and language use.* Oxford, UK: Oxford University Press.

Widdowson, H. (2000). The limitations of linguistics applied. *Applied Linguistics, 21*(1), 3–25.

Wiley, D. (n.d.). Defining the "open" in open content. Retrieved from http://opencontent.org/definition/

Wiley D. (2011, July 27). Openwashing – the new greenwashing [Blog post]. *Iterating toward openness*. Retrieved from http://opencontent.org/blog/archives/1934

Willinsky, J. (2002). Copyright contradictions in scholarly publishing. *First Monday*, *7*(11).

Willis, J. (1998). Concordances in the classroom without a computer. In Tomlinson, B. (Ed.). *Materials Development in Language Teaching* (pp. 44–66). Cambridge, UK: Cambridge University Press,

Wise, A. F., Cui, Y., & Vytasek, J. (2017). Mining for gold: Identifying content-related MOOC discussion threads across domains through linguistic modelling. *Internet and Higher Education*, *32*, 11–28.

Witten, I. H., Paynter, G., Frank, E, Gutwin, C., & Neville-Manning, C.G. (1999). KEA: Practical automatic keyphrase extraction. In *Proceedings of Digital Libraries* (pp. 254–255).

Witten, I. H., Bainbridge, D., & Nichols, D. M. (2009). *How to build a digital library* (Second Edition). Burlington, MA: Morgan Kaufmann.

Witten, I. H., Wu, S., & Yu.X. (2011). Linking digital libraries to courses...with particular application to language learning. In *Proceedings of the 4th International Conference on Computer Supported Education* (pp. 5–14). Noordwijkerhout, The Netherlands: SciTePress.

Witten, I. H., Wu, S., Li, L., & Whisler, J. (2013). *The book of FLAX: A new approach to computer assisted language learning*. (Second Edition). University of Waikato, New Zealand. Retrieved from http://flax.nzdl.org/BOOK_OF_FLAX/BookofFLAX%202up.pdf

Witten, I. H., Eibe, F., Hall, M., & Pal, C.J. (2016). *Data mining: Practical machine learning tools and techniques* (Fourth Edition). Burlington, MA: Morgan Kaufmann.

Witten, I. H. (September 2017). Mining MOOCs to assist second language learners. Seminar given at the Knowledge Media Institute, Berrill Building, The Open University, September 2017. Milton Keynes, UK. Retrieved from http://stadium.open.ac.uk/stadia/preview.php?s=29&whichevent=2905

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.

Wu, S., Franken, M., & Witten, I. H. (2012). Collocation games from a language corpus. In H. Reinders (Ed.), *Digital games in language learning and teaching* (pp. 209–229). Basingstoke, UK: Palgrave Macmillan.

Wu, S., Fitzgerald, A., & Witten, I. H. (2014). Second language learning in the context of MOOCs. In S. Zvacek, M. T. Restivo, J. Uhomoibhi, & M. Helfert (Eds.), *Proceedings of the 6th International Conference on Computer Supported Education* (pp. 354–359). Barcelona: SciTePress.

Wu, S., Li, L., Witten, I. H., & Yu, A. (2016). Constructing a collocation learning system from the Wikipedia corpus. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)*, *6*(3), 18–35. doi:10.4018/IJCALLT.2016070102.

Wu, S., & Witten, I. H. (2016). Transcending concordance: augmenting academic text for L2 writing. *International Journal of Computer-Assisted Language Learning and Teaching*, *6*(2), 1–18. doi:10.4018/IJCALLT.2016040101.

Wu, S., Fitzgerald, A., Witten, I. H. & Yu, A. (2018). Automatically augmenting academic text for language learning: PhD abstract corpora with the British Library. In B. Zou, M. Thomas (Eds.), *Integrating Technology into Contemporary Language Learning and Teaching*, (pp. 512–537). Hershey, PA: IGI Global.

Wu, S., Fitzgerald, A., Yu. A., & Witten, I. H. (2019). Developing and evaluating a learner-friendly collocation system with user query data. *International Journal of Computer-Assisted Language Learning and Teaching*, *9*(2), 53–78.

Yeh, Y., Li, Y.-H., & Liou, H.-C. (2007). Online synonym materials and concordancing for EFL college writing. *Computer Assisted Language Learning, 20*(2), 131–152. doi: 10.1080/09588220701331451

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

Yu, A., Wu, S., Witten, I. H., & König, J. L. (2016). Learning Collocations with FLAX Apps. In L. E. Dyson, W. Ng & J. Fergusson (Eds.), *Proceedings of the 15th World Conference on Mobile and Contextual Learning (mLearn 2016) – Sustaining Quality Research and Practice in Mobile Learning*, (pp. 291–294). Sydney: Australia.

Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: effects of frequency and contextual richness. Canadian Modern Language Review, *15*(4), 541–572.

Zajacova, A., Lynch, S., & Espenshade, T. (2005). Self-efficacy, stress, and academic success in college. *Research in Higher Education*, *46*, 677–706. doi:10.1007/s11162-004-4139-z

Zimmerman, B. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, *25*(1), 3–17.