Advancing disciplinary literacy through English for academic purposes: Discipline-specific wordlists, collocations and word families for eight secondary subjects

Clarence Green[*], James Lambert

Department of English Language and Literature, National Institute of Education, Nanyang Technological University, NIE3-03-118, 1 Nanyang Walk, 637616, Singapore

A R T I C L E   I N F O

A B S T R A C T

The knowledge of academic vocabulary is crucial for educational success, and recently there has been a push amongst teachers and researchers to assist students at the pre-tertiary level to develop their disciplinary literacy and understanding of how academic English varies across disciplines. EAP research has developed advanced methods for producing corpus-informed vocabulary resources, but these have yet to be fully leveraged to promote disciplinary literacy within the secondary school context. For example, the focus of most previous wordlists has been on general academic vocabulary or the discipline-specific vocabulary needed in tertiary education. The current research contributes a series of discipline-specific wordlists for secondary school education, the *Secondary School Vocabulary Lists* (SVL), covering eight core subjects: Biology, Chemistry, Economics, English, Geology, History, Mathematics, and Physics. Further, the SVL goes beyond wordlists alone in developing accompanying word family and word association (i.e. collocation) lists for the disciplinary lexis. The SVL thus provides secondary education teachers with an unprecedented set of resources covering key vocabulary for the eight core disciplines informed by innovative EAP corpus methods.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction: EAP affordances to advance disciplinary literacy in secondary education

With an increasing number of school students continuing to tertiary education, developing students' academic literacy during secondary school is essential. As Wingate (2011) notes, "in today's mass higher education system, many students are not fully prepared for the demands of academic study" (p. 66). Thus, research-based resources with the potential to assist teachers prepare secondary students for tertiary education are essential. Vocabulary is a vital component of educational success in both first and second language contexts (Webb & Nation, 2017), and in English for Academic Purposes (EAP) pedagogical materials to facilitate instructed vocabulary acquisition have been the focus of much research; in particular, advanced methods have developed for producing corpus-informed wordlists (Gardner & Davies, 2014; Lei & Liu, 2016). However, such wordlists have largely been developed in the context of tertiary level ESL education. The current advances in EAP for developing such educational resources have yet to be fully leveraged in other educational contexts, but there are

---

* Corresponding author.
    E-mail address: clarence.green@nie.edu.sg (C. Green).

recent trends in this direction as exemplified by the *Middle School Vocabulary Lists* (Greene & Coxhead, 2015). This paper aims to further advance the trend by reporting on a large corpus project that releases to the research and teaching community the *Secondary School Vocabulary Lists* (SVL), a series of lemma-based discipline-specific academic vocabulary lists. These lists cover important vocabulary from eight core subjects that students need to master during secondary education: Biology, Chemistry, Economics, English, Geology, History, Mathematics, and Physics. Furthermore, the current study goes beyond previous wordlist development by contributing accompanying pedagogical lists containing word associations (i.e. collocations) and word families.

The SVL is designed for secondary school as a resource to assist teachers in the development of disciplinary literacy. Disciplinary literacy is a pedagogical approach and research area of increasing significance in secondary education internationally (Airey, Lauridsen, Räsänen, Salö, & Schwach, 2017). It emphasizes the connection between language and the disciplines, e.g. rather than English teachers focussing on a general academic vocabulary and subject area teachers on content, this approach calls for more explicit teaching of disciplinary language by all teachers. Shanahan and Shanahan (2017) suggest that students benefit the most from the disciplinary literacy approach during middle and secondary school as language becomes increasingly specialised. A resource that profiles the important language of secondary disciplines by adapting the methods of EAP research could therefore be very useful for such pedagogy. It would provide a resource from which, for example, teachers might select study words, knowing that the vocabulary chosen is important to the disciplines of secondary education (Ogle, Blachowicz, Fisher, & Lang, 2016). As Römer (2011, p. 209, p. 209) notes, even if a content area teacher does not focus on certain vocabulary until they reach a particular topic, language teachers can still introduce such discipline-specific vocabulary beforehand with the confidence that students will be better prepared for content classes. In the broader context, corpus-informed wordlists developed using the advanced methods of EAP within ESL research are increasingly relevant to schooling in predominantly English speaking countries. In the past, secondary classrooms have perhaps been seen as largely native speaker environments, but this is certainly no longer the case and secondary teachers are fully aware they need pedagogies and materials that support the different language backgrounds and proficiencies of contemporary classrooms. Further, the SVL can assist teachers in the growing number of international schools and EFL classes at the secondary level world-wide.

2. The pedagogical issues addressed by corpus-informed wordlists

The value of pedagogical material informed by objective methodological procedures developed in corpus linguistics is widely recognized (Brezina & Gablasova, 2017). The *General Service List* (GSL) of West (1953) has had no less than two recent updates (Brezina & Gablasova, 2013; Browne, Culligan, & Phillips, 2013), the *Academic Wordlist* (Coxhead, 2000) has recently been supplemented by the *Academic Spoken Wordlist* (Dang, Coxhead, & Webb, 2017), and there is an increasing trend toward the development of discipline-specific wordlists (Lei & Liu, 2016). The insight underlying such wordlists is that frequency, combined with metrics such as range and dispersion, profiles for teachers and students the relative usefulness of words (Brezina & Gablasova, 2017). Nation (2006, p. 79), for example, calculated that for any general text, comprehension requires about 98% understanding of its vocabulary, and that remarkably a teacher can provide this coverage by targeting the most frequent 8000–9000 word families of English. The general idea has a long tradition in educational research. Thorndike and Lorge (1944) compiled a corpus of primary and secondary school texts to produce a series of wordbooks to "enable a teacher to know not only the general importance of each word so far as frequency of occurrence measures that" (p. 1).

Vocabulary provides a foundation from which grammar, phonology, and morphology emerge, and in a subject area it provides access to conceptual knowledge (Coxhead, 2018). Vocabulary selection for pedagogical purposes is therefore crucial. Nation (2016) suggests a well-planned curriculum might progress from teaching general high frequency words, to general academic words, to discipline-specific and technical words, guided at each stage by corpus-informed resources that assist teachers in selecting vocabulary. Ogle et al. (2016), discussing how secondary teachers might intuitively select vocabulary, offer the example *pinnae* and recommend that although a word in secondary science textbooks that students would not understand, it should be treated as low-utility by teachers planning a vocabulary curriculum given its low frequency and topic-specificity. A corpus-informed wordlist could clearly facilitate such decisions. There are limitations to teacher intuition; as Schmitt (2010, p. 67) notes, given the idiolectal nature of language, teachers vary on which vocabulary they think important. This is also true of individual textbook writers (Harwood, 2014).

3. Corpus-informed wordlists: academic and discipline-specific vocabulary

To succeed in school, students need to develop an academic vocabulary (Shanahan & Shanahan, 2017). Perhaps the seminal wordlist for academic vocabulary is Coxhead's (2000) *Academic Wordlist* (AWL), designed to assist ESL speakers prepare for university. The AWL's impact on tertiary-level language education has been significant (Schmitt, Schmitt, & Mann, 2011) and there has been a growing awareness of corpus-based wordlists in the K-12 context. DiCerbo, Anstrom, Baker, and Rivera (2014) in a review of pedagogical approaches to academic vocabulary in US schools recommend the AWL as containing much vocabulary relevant to secondary school, even though the AWL is based on a corpus of university texts, largely sourced in New Zealand. Methodologically, Coxhead (2000) developed the AWL using the definition of word families in Bauer and Nation (1993), with a word family being a headword and its inflectional and derivational morphological forms, e.g. *react, reaction, reactor.* Coxhead (2000, p. 221) was particularly concerned with establishing corpus