English for Specific Purposes 51 (2018) 69–83

Contents lists available at ScienceDirect

English for Specific Purposes

journal homepage: http://ees.elsevier.com/esp/default.asp

From the Editors

The nature of vocabulary in academic speech of hard and soft-sciences

Thi Ngoc Yen Dang*

Vietnam National University, Hanoi, Viet Nam

ARTICLE INFO

Article history:

Keywords:
Hard science
Soft science
Academic spoken discourse
Vocabulary
Word list
Corpora

ABSTRACT

Little is known about the similarities and differences between the vocabulary in hard-sciences (e.g., Maths, Engineering, Medicine) and soft-sciences (e.g., Business, Law, History), especially in spoken discourse. To address this gap, a Soft Science Spoken Word List (SSWL) was developed for second language learners of soft-sciences at English-medium universities. The list consists of the 1,964 most frequent and wide-ranging word-families in a 6.5 million word corpus of soft-science speech, which represents 12 subjects across two equally-sized sub-corpora. The list may allow learners to recognize 94%–97% of the words in academic speech of soft-sciences. A comparison of the SSWL with Dang's (2018) Hard Science Spoken Word List revealed that although the most frequent 3,000 words are important for comprehending academic speech of both soft- and hard-sciences, the value of these words in soft-sciences is greater than in hard-sciences. Pedagogical implications related to this nature of vocabulary in hard- and soft-science speech are provided.

1. Introduction

English has been widely used as the medium of instruction in academic courses at tertiary levels—in face-to-face, distance learning, and online contexts—in both English speaking and non-English speaking countries. Second language (L2) learners in these courses have to comprehend not only reading materials such as textbooks and research articles but also lectures, seminars, labs, and tutorials (Biber, 2006). Because vocabulary knowledge and comprehension are closely related (Laufer & Ravenhorst-Kalovski, 2010; Schmitt, Jiang, & Grabe, 2011; van Zeeland & Schmitt, 2013), it is essential for learners to master the words that they are likely to encounter often in a wide range of academic written and spoken texts. A large number of wordlists have been developed to assist L2 learners' comprehension of academic writing (e.g., Coxhead, 2000; Coxhead & Hirsh, 2007; Gardner & Davies, 2014; Liu & Han, 2015; Martínez, Beck, & Panza, 2009; Watson-Todd, 2017; Wang, Liang, & Ge, 2008). Yet little has been done to help these learners comprehend academic speech. In fact, understanding academic spoken English is a great challenge for L2 learners in different contexts (Flowerdew & Miller, 1992; Mulligan & Kirkpatrick, 2000). Given this fact, it is crucial to create wordlists that capture the most frequent and wide-ranging words in academic speech. Together with written wordlists, these spoken wordlists are valuable resources for English for Academic Purposes (EAP) programs to support L2 learners' comprehension of academic English.

* The Faculty of English, University of Languages and International Studies, Vietnam National University, Level 5, Building B2, 2 Pham Van Dong, Cau Giay, Hanoi, Viet Nam.
E-mail address: ngocyen1011@gmail.com.

https://doi.org/10.1016/j.esp.2018.03.004
0889-4906/

Depending on the target subject areas of their learners, EAP programs can be divided into two types: English for Specific Academic Purposes (ESAP) and English for General Academic Purposes (EGAP).[1] Learners in ESAP programs are fairly homogeneous in terms of target subject areas. That is, they all plan to study hard-science subjects (e.g., Mathematics, Physics, Engineering or Medicine) or soft-science subjects (e.g., Linguistics, Law, Business, or Education). The hard/soft division refers to the existence of a paradigm, or 'a body of theory which is subscribed to by all members of the field' (Biglan, 1973b, p. 201). Earlier research on students' learning strategies and scholars' behavior and opinions about various aspects of academic disciplines (e.g., teaching, learning, research styles) (e.g., Becher, 1989; Biglan, 1973a, 1973b; Neumann, Parry, & Becher, 2002) has indicated that the hard/soft division is the strongest dimension[2] that distinguishes academic subjects in higher education. Hard-sciences (e.g., Mathematics, Engineering) are likely to have a single paradigm that allows scholars working in these areas to reach a wide consensus on research methods and key concepts. In contrast, soft-sciences (e.g., Law, Philosophy) are likely to lack a single paradigm, and scholars working in these areas seem to argue over methods and key concepts. While the hard/soft division cannot fully reflect the complexity and variation in inquiry processes and knowledge structures in various disciplines (Becher & Trowler, 2001; Nesi, 2002, pp. 351–358), this division is a useful shorthand when attempting to explain the complexity and diversity of academic discourse.

Learners in EGAP programs, however, are more heterogeneous in terms of their target disciplines. In other words, there is a mixture of hard-science and soft-science students in these programs. EGAP programs can also be the programs where (a) learners have not yet identified their target subject areas, (b) learners plan to study interdisciplinary subject areas, (c) or teachers lack background knowledge of learners' specific subject areas.

According to Hyland (2016), general and specific EAP approaches should be considered as a continuum rather than a dichotomy. Depending on the teaching and learning context of a particular EAP program, either a general academic wordlist or a discipline-specific wordlist is more suitable than the other (Dang, Coxhead, & Webb, 2017). A general academic wordlist is more relevant to EGAP programs. The diversity in learners' academic subject areas may make it challenging for teachers in these programs to satisfy the specific needs of every learner. Meanwhile, a discipline-specific wordlist is more suitable for ESAP programs. As specialized vocabulary tends to occur more often in specialized texts (Chung & Nation, 2004), discipline-specific wordlists focus learners' attention on items that occur very often in their specific areas and provide a shortcut to reduce the amount of learning (Nation, 2013). Learners are more motivated to learn items from discipline-specific wordlists because they can see clearly the link between what they learn in their ESAP course and their subject courses (Coxhead & Hirsh, 2007; Hyland, 2016). Additionally, the similarities between learners' academic disciplines may make it easier for teachers to focus on specialized vocabulary in a particular discipline.

Several general academic spoken wordlists have been developed for EGAP programs such as Nesi's (2002) Spoken Academic Word List, Simpson-Vlach and Ellis's (2010) Academic Formulas List, and Dang et al.'s (2017) Academic Spoken Word List. As a result, we have a fairly good understanding about the shared spoken vocabulary across hard- and soft-sciences. In contrast, only one spoken discipline-specific wordlist has been developed for ESAP programs, and it focuses on spoken vocabulary in hard-sciences: Dang's (2018) Hard Science Spoken Word List (HSWL). No attempts have been made to identify the most frequent and wide-ranging words in soft-science speech. The lack of such a list makes it challenging to compare the vocabulary in academic speech of the hard- and soft-sciences.

The present study was conducted with two aims. The first aim was to develop a Soft Science Spoken Word List (SSWL) for ESAP programs which consist of solely soft-science students. The second aim was to compare this list with Dang's (2018) HSWL to see the similarities and differences between the most frequent and wide-ranging words in hard disciplines and soft disciplines. The research thus provides soft-science students in ESAP programs with a useful instrument to achieve better comprehension of academic speech, and sheds light on the nature of vocabulary in hard- and soft-science speech.

1.1. Background

Studies that investigated the nature of vocabulary in spoken texts of hard- and soft-sciences either examined the vocabulary load of the two disciplines or focused on the shared vocabulary between these disciplines. Vocabulary load studies (Coxhead, Dang, & Mukai, 2017; Dang & Webb, 2014) determined the number of words required to understand academic speech. They looked at the lexical coverage of different 1,000-word frequency levels of general vocabulary in academic speech and estimated the number of words needed to reach 95% and 98% coverage of these texts. Lexical coverage is the percentage of words covered by items from a particular wordlist in a text (Nation & Waring, 1997). The 95% and 98% figures have been widely used as the coverage cut-off points to indicate high and stable degrees of comprehension (van Zeeland & Schmitt, 2013).

Dang and Webb (2014) examined the vocabulary load of lectures and seminars of the hard- and soft-sciences represented in the British Academic Spoken English Corpus (BASE). They found that a vocabulary size of 3,000–4,000 word families is needed to reach 95% coverage of soft-science speech while a vocabulary size of 5,000–7,000 word families is necessary to

[1] These are the definitions of EGAP and ESAP in the present study. There is some variation in the understanding of these terms in the field of EAP, however.

[2] The other dimensions are pure/applied and life/non-life, which refer to the concern of the areas with (a) application to practical problems and (b) life systems, respectively.