

From Corpus to Classroom: Teaching Semi-technical Business English Vocabulary

Lidan Chen¹

¹ School of English for International Business, Guangdong University of Foreign Studies, Guangzhou, China

Correspondence: Lidan Chen, School of English for International Business, Guangdong University of Foreign Studies, Guangzhou, 510420, China.

Received: September 20, 2023 Accepted: November 12, 2023 Online Published: November 24, 2023

doi:10.5539/ijel.v13n6p55 URL: <https://doi.org/10.5539/ijel.v13n6p55>

Abstract

Semi-technical vocabulary has been considered a challenging and neglected area of English for Specific Purposes (ESP) instruction. This paper employs AntConc to extract a keyword list of a self-compiled business English textbook corpus (BETC). Through manual identification from corpus keywords, we focus on semi-technical vocabulary, addressing the fundamental question of “how to identify”. We also draw upon pedagogical materials from the academic sub-corpora within the Corpus of Contemporary American English (COCA) to design a corpus-based language pedagogy (CBLP) lesson. This lesson serves as a model for instructing the multifaceted meanings and diverse patterns of the semi-technical word “default” across various disciplines and contexts, addressing the question of “how to teach”. Our research leverages the rich resources provided by pedagogical corpora, offering in-depth analysis, including collocation, colligation, semantic preference, and semantic prosody, as effective teaching aids. In doing so, it promotes interdisciplinary, comparative, and exploratory teaching and learning of semi-technical business English vocabulary. By bridging the gap between corpus analysis and classroom instruction, it provides innovative strategies for educators in the field of business English, and by implication, in various ESP disciplines.

Keywords: Business English Instruction, Semi-Technical Vocabulary, CBLP, AntConc

1. Introduction

The realm of English for Specific Purposes (ESP) encompasses a vast landscape of specialized language use, tailored to the unique demands of various disciplines and professional contexts. Within this diverse area, semi-technical vocabulary emerges as a highly important and particularly challenging aspect of vocabulary instruction. Its importance was underscored in Inman’s research on a 100 000-word scientific and technical corpus. She found that it accounts for a substantial 70%, surpassing function words at 9% and technical words at 21% (Inman, 1978, as cited in Farrell, 1990). While semi-technical vocabulary isn’t “fully technical” or “fully specialized”, it plays an integral part in professional communication, where clarity and fluency are paramount. Consequently, educators should acknowledge the vital role these words play in empowering students to excel in their chosen field of study and career development.

As defined by Farrell (1990), semi-technical vocabulary consists of lexical items used in general language but with a higher frequency of usage and a strong dependence on the context in specialized discourse. Highly technical vocabulary, which is explicitly taught by subject matter instructors, doesn’t usually pose significant challenges for non-native learners. In contrast, students often struggle with the comprehension of semi-technical vocabulary, as these words frequently adopt extended meanings within technical contexts (Li & Pemberton, 1994; Trimble, 1995; Gardner, 2007; Hyland & Tse, 2009; Li & Miller, 2023). Unlike technical words (sometimes named terminology or technical terms), semi-technical words evolve from the common lexicon, often conveying multiple meanings and displaying different patterns. This tendency for semantic divergence and dynamism of semi-technical vocabulary endows it with a notable degree of abstraction, diversity, and complexity, making it a challenge for foreign language learners (Dun, 2002; Chen & Gao, 2010).

Due to its dynamic and intricate nature, semi-technical vocabulary often poses challenges for traditional teaching approaches that rely solely on a single teaching resource, such as a textbook. These approaches frequently fall short of equipping learners with the nuanced comprehension and application of these words across various disciplines. In response to these challenges, the present research adopts a corpus-based language pedagogy

(CBLP) approach to bridge the gap from corpus analysis to language instruction. To achieve this, we draw on pedagogical materials extracted from a self-compiled business English textbook corpus (BETC) and enrich the resources by incorporating academic sub-corpora sourced from the Corpus of Contemporary American English (COCA). Through these combined resources, we attempt to explore the complexities of semi-technical business vocabulary, addressing two key questions: “how to identify” (identification and analysis) and “how to teach” (teaching approach and model). The extraction of corpus keywords in BETC by use of AntConc, which is supplemented by manual identification, provides a preliminary answer to the “what” question. Employing a corpus-based language pedagogy approach, we navigate the intricate web of meanings and versatile patterns of the sample word across diverse domains and contextual settings. The multi-step teaching model integrates corpus concordance analysis to provide the necessary teaching scaffolding. In so doing, we address the equally significant question of “how to teach”.

2. Literature Review

2.1 Definition and Identification of Semi-Technical Vocabulary

While academia presents different definitions of semi-technical vocabulary, vocabulary categorization in ESP consistently centers on the degree of technicality, dividing vocabulary into technical, semi-technical/sub-technical, and non-technical categories (Cowan, 1974; Baker, 1988; Nation, 2001, 2013). The controversy in defining semi-technical vocabulary has two primary references. The first pertains to context-independent words which have high-frequency usage across different disciplines (Cowan, 1974; Yang, 1986; Baker, 1988; Nation, 2001; Zhen & Wang, 2014). The second includes common words with special meanings and words that possess one or more general meanings but assume extended meanings in technical contexts (Trimble, 1985; Flowerdew, 1993; Menon & Mukundan, 2010; Le & Miller, 2023). In this paper, we adopt a working definition for words that are commonly used in both ESP and general English but exhibit differences in meaning and higher frequency in specialized discourse. For example, the word “administer” typically conveys the meaning of “to manage and organize the affairs of a company, an organization, or a country” (Parkinson & Noble, 2000) in Business English. However, in the domain of Medical English, it signifies “to give a drug to a patient as a treatment” (Collins, 2018). This example illustrates how semi-technical vocabulary can take on different meanings in various specialized disciplines, emphasizing the importance for learners to comprehend these nuances when communicating in different contexts. Furthermore, the dynamic nature of meaning brings about complexities in the usage patterns of these words in discipline-specific discourses, adding to the challenge of identifying and understanding semi-technical vocabulary.

Efforts to identify semi-technical vocabulary have evolved significantly within the domain of technical or specialized vocabulary (Yang, 1986; Flowerdew, 1993; Coxhead, 2000; Fraser, 2007, 2012; Nation, 2013; Hsu, 2013; Watson-Rodd, 2017; Le & Miller, 2023). Chung and Nation (2003) summarized two general approaches to identifying technical vocabulary. The first involves using the intuition of a subject expert which was achieved through the use of a rating scale, a technical dictionary, or contextual clues. The second approach is to use a corpus-comparison approach, comparing word frequencies in a technical text with those in a different corpus. This approach, known as the keyword technique in corpus linguistics, focuses on words that exhibit higher-than-average frequencies in specialized corpora than in general corpora (Coxhead, 2000). Comparing these two approaches, Chung and Nation (2003) concluded that the computer-based approach offers ease of application and high practicality. Some highly relevant studies on semi-technical words have either relied on the knowledge and experience of the researcher and an independent coder’s selection from the wordlist (Menon & Mukundan, 2010) or focused on manually selected frequent or salient semi-technical items in the corpus keyword list (Baker, 1988; Maher, 2016; Le & Miller, 2023). Additionally, some research centered on the development of a Business Word List, which will be elaborated in Section 2.2. Irrespective of the identification approach chosen, all relevant studies emphasize the importance of collocation and patterns (Baker, 1988; Menon & Mukundan, 2010; He & Zhang, 2011; Zhen & Wang, 2014; Maher, 2016, among others). The assumption underlying these studies is that these words exhibit text- and subject-specific usage, often occurring in close association with other domain-specific terms. By examining collocational patterns, researchers gain insights into the contextual use of these words.

2.2 Analysis and Teaching of Semi-Technical Business Vocabulary

Business English learners should possess a strong command of specialized vocabulary to operate effectively in their professional roles, as a profession’s knowledge and its specialized vocabulary are closely linked (Coxhead, 2018). Hence, it is important to review relevant research on the analysis and teaching of technical business vocabulary, from which semi-technical business vocabulary is extracted.

Researchers have explored business English corpora to develop business wordlists or keyword lists, from which semi-technical business words are discovered. For instance, Farmer (1967, as cited in Nelson, 2000) extracted a wordlist from a 10956-word business letter corpus. Nelson (2000) categorized business keywords into people in business, business activities, business actions, business descriptions, and business events and entities. Camiciottolio (2007) identified 174 technical words in the Business Studies Lectures Corpus, including semi-technical words like “market,” “product,” and “price”, which are neither strictly technical nor exclusive to a specific discipline. Concurrently, Konstantakis (2007) developed a business wordlist by drawing on the Business English Textbook Corpus, a part of Nelson’s (2000) corpora. More recent studies have continued to contribute to the study of technical or specialized business vocabulary. Tangpijaikul (2014) retrieved a 134-word list of technical terms in the 890 000-word business and economic news corpus in Thailand, which were jointly agreed upon by the researcher and two experienced business managers. Browne and Culligan (2016), using a 64.5-million-word business corpus, developed a Business Service List containing approximately 17 000 high-frequency general business English words, spanning from everyday vocabulary to a more specialized lexicon. Across these studies, a consistent discovery is that semi-technical business vocabulary often comprises polysemous words exhibiting both general and technical meanings that are related (Jackson, 1988, as cited in Farrell, 1990).

While these corpus-based studies of specialized business vocabulary offer significant implications for analyzing and teaching semi-technical business vocabulary, it’s noteworthy that the analysis of semi-technical business vocabulary has not received the attention it deserves, particularly in teaching them in the ESP classroom. Given the importance of semi-technical vocabulary and the learning difficulty due to its semantic dynamism and multiple forms, the analysis and teaching of semi-technical business vocabulary should be highly valued, highlighting a significant research gap.

Traditional vocabulary teaching methods often fall short of providing the depth and context necessary for learners to grasp the nuances in the semantic dynamism and multifaceted nature of semi-technical words effectively (Nation, 2001). The emergence of corpus linguistics has transformed language research and pedagogy in that corpora can provide vast collections of authentic language data with rich contextual information, enabling educators to gain insights into real-world language use (Biber, 1993). Corpus linguistics has greatly advanced our understanding of language analysis, “shifting from intuition- to corpus-based inductive observations of language accounts” (Ma, Chiu, Lin, & Mendoza, 2023, p. 20). As mentioned earlier, most corpus-based studies of semi-technical vocabulary, whether in business English or other ESP areas, emphasize collocation and pattern analysis to facilitate lexical learning in a practical and context-based manner. This strongly justifies the application of corpora in teaching semi-technical business English vocabulary.

The integration of corpus into language teaching and learning, formally known as Data-Driven Learning (DDL), was first introduced by Tim Johns (Johns, 1991; Sinclair, 1991) and subsequently endorsed by many scholars (Leech, 1997; Gilquin & Granger, 2010; Chang, 2014; Boulton & Cobb, 2017; Crosthwaite, 2017, among others). DDL is a student-centered and teacher-guided approach that employs corpus techniques like concordance to help students observe and summarize the target word’s lexico-grammatical patterns, providing deeper insights into its contextual meaning. A recent development in this approach is corpus-based language pedagogy (CBLP), which combines corpus linguistics with classroom pedagogy. This innovative pedagogy integrates corpora and corpus technology into classroom language teaching based on four design principles to facilitate language instruction (Ma & Mei, 2021), which will be used to develop a teaching case in Section 4.

The literature review in both specialized vocabulary studies and CBLP underscores the crucial role of corpora in the identification and instruction of semi-technical business English vocabulary. Motivated by this, the researcher will detail how corpora, coupled with corpus techniques, can be used for pedagogical purposes in the context of teaching semi-technical business English vocabulary to address the two research questions:

- 1) How could semi-technical vocabulary be identified in business English pedagogy?
- 2) How could this vocabulary be effectively instructed so that learners can master their dynamic meanings and multifaceted forms?

To answer these two questions, we will present a systematic approach to identifying and analyzing semi-technical business English vocabulary in the following section.

3. Research Method

This section outlines the corpus tool, corpora, and data-processing procedures to investigate the first research question “how could semi-technical business English vocabulary be identified for teaching”. This approach

adopted is in alignment with the growing shift from intuition-based to corpus-based language analysis and pedagogy, allowing for practical and context-based corpus analysis. The findings from this section will provide valuable teaching scaffolding in the subsequent teaching case.

3.1 Instrument

The corpus concordancer AntConc 3.5.9 was utilized to process the self-compiled pedagogical corpus BETC. To extract a keyword list, the general corpus CROWN (appropriate 1 000 000 running words) (Xu & Liang, 2010) served as the reference corpus. Following definitions by researchers such as Farrel (1990), the identification of semi-technical business English vocabulary was conducted using the keyness-ordered keyword list, in collaboration between the researcher and a subject teacher with over 10 years of teaching experience. In cases of uncertainty or disagreement during the identification process, references like the Oxford Business English Dictionary and Collins Cobuild Learners' Dictionary were consulted.

Apart from extracting keyword lists, AntConc provided concordance lines and collocates of the target word across various pedagogical corpora in this study, including BETC, COCA Academic Legal Sub-corpus, and COCA Academic Science and Technology Sub-corpus. These diverse corpus resources enable students to compare lexical and grammatical structures and multiple meanings of semi-technical vocabulary in various disciplinary contexts.

3.2 The Corpora

Three specialized corpora are employed for pedagogical purposes in this study. BETC (with 26 419 running words) includes all reading texts of a business English textbook (Cai, 2009), specifically chosen for a full semester of instruction, excluding instructional content. These texts cover topics such as gender issues in business, cross-cultural business communication, marketing, social security, leadership, money and banking, investment, and business ethics. The rationale behind extracting semi-technical business vocabulary from academic genres, such as textbooks and the COCA academic section, is twofold. Firstly, business English textbooks play a significant role in mainland China university education in that it remains a primary source of students' exposure to the English language and business knowledge, particularly in the university that uses English as the Medium of Instruction (EMI). Thus, the analysis of textbook corpora could substantially aid instructors and learners in identifying key language items and their specific patterns (Sinclair, 1991; Hunston & Francis, 2000). Second, the textbook corpus, along with two COCA academic sub-corpora, serves as accessible and relevant pedagogical resources for undergraduate students. Consequently, the researcher compiled a pedagogical corpus from the business English textbook used by second-year Business English undergraduates. The remaining two corpora were drawn from the COCA academic section (with 120 988 361 running words), which covers the full range of academic disciplines and is sourced from more than 200 different peer-reviewed journals. Given that the selected semi-technical word "default" to be demonstrated in the subsequent teaching case has multiple usages within the two sub-corpora of law and science/technology from the COCA academic section, these two specific sub-corpora were accessed to extract the concordance lines of "default".

3.3 Data Processing

The data processing procedures started with the extraction of the keyword list from BETC using AntConc. The identification of semi-technical business vocabulary and extraction of relevant concordance lines from three corpora followed a manual approach. The steps undertaken are delineated below.

Step 1: Extraction of BETC keyword list and identification of semi-technical vocabulary

This research adopted the computer-based approach to identifying specialized vocabulary, guided by the work of Chung and Nation (2003, 2004) for its easy application and high practicality, as previously reviewed. The process commenced by loading BETC, the target corpus, into AntConc. A general corpus CROWN was then loaded as a reference corpus to generate a keyword list organized by keyness. Keywords in corpus linguistics are words that demonstrate significantly higher or lower frequency in a given corpus compared to a reference corpus which is typically larger. Corpus keywords can reveal the aboutness (subject knowledge) of the corpus (Hunston, 2002; Bondi, 2010). As shown in Figure 1, most of the high-ranked keywords displayed strong specialization and a close semantic connection with business knowledge. They are highly relevant to the business topics covered within the corpus. For instance, keywords like "inflation", "rate", "bank(s)", "collateral", and "lending" are specific to "money and banking", a financial topic of the business discipline. Similarly, keywords like "investment", "bond(s)", "stock(s)", "security", "investing", "risk", and "return" are associated with "investment", another financial topic in the business discipline. In this sense, the extracted positive keywords, which occur in BETC with significantly higher frequency, effectively indicate the subject matter of the corpus.

Rank	Freq	Keyness	Effect	Keyword
1	76	+ 445	0.0057	inflation
2	62	+ 271.52	0.0047	leadership
3	38	+ 253.47	0.0029	cola
4	35	+ 236.76	0.0026	coca
5	30	+ 212.01	0.0023	dhl
6	85	+ 202.81	0.0063	social
7	44	+ 195.28	0.0033	managers
8	64	+ 163.91	0.0048	security
9	33	+ 149.19	0.0025	welfare
10	40	+ 142.03	0.003	leader
11	45	+ 132.08	0.0034	rates
12	17	+ 125.11	0.0013	walmart
13	22	+ 125.02	0.0017	supermarkets
14	61	+ 116.43	0.0045	company
15	37	+ 115.32	0.0028	investment
16	25	+ 114.64	0.0019	bonds
17	31	+ 107.86	0.0023	vision
18	19	+ 100.96	0.0014	default
19	22	+ 94.75	0.0017	bond
20	19	+ 92.38	0.0014	investing
21	26	+ 85.79	0.002	stock
22	26	+ 82.78	0.002	loans
23	33	+ 79.95	0.0025	leaders
24	41	+ 79.7	0.0031	rate
25	34	+ 79.68	0.0026	return
26	39	+ 79.07	0.0029	interest

Figure 1. BETC Keyword list (arranged by keyness)

An intriguing observation in the keyword list is the presence of specialized words such as “security”, “bond”, “default”, “return” and “interest”, among others. These words, whose multiple meanings are demonstrated in Appendix A Table A1–5, exhibit a dual nature, functioning as common words in general English and as semi-technical words with different or extended meanings across various disciplines. Their polysemous quality across different disciplines positions them as prime candidates for semi-technical in our research, identified through both manual selection by the researcher and a subject teacher and a computer-based method. Additionally, the selection process takes into account students’ experiences as a supplementary criterion. Drawing from over 15 years of experience in teaching business English, observations frequently reveal students grappling with these polysemous words in the business context. Therefore, students’ comprehension challenges augment this selection process in filtering semi-technical vocabulary from the extracted keyword list.

Unlike other candidates, whose multiple meanings and usages are explicitly elaborated in dictionaries, the word “default” has not received comprehensive coverage in commonly used learner’s dictionaries or business English dictionaries. This absence in reference materials is a primary reason why this word was chosen as the focal point for the designed teaching case in Section 4.

Step 2: Extraction of concordance lines from BETC and affordances of teaching scaffolding

The concordance line typically displays each occurrence of the target word (also referred to as the “node word”) in the center of the line. These lines can be sorted in alphabetical order based on the L1 (preceding) or R1 (following) word of the target word so that the collocates and lexico-grammatical patterns become more apparent in the context (Nesi, 2013). As illustrated in Figure 2, 23 concordance lines of “default” are extracted and arranged in R1 alphabetical order for Extended Unit of Meaning (EUM) analysis. Initially introduced by Sinclair in 1996 and later refined into a model comprising five key components, EUM offers a comprehensive framework to explore the usage-based meaning of the node word. These components include the core element (the node word itself), collocation (co-occurring vocabulary), colligation (grammatical patterns), semantic preference (semantic attributes), and semantic prosody (attitudinal and evaluative meanings) (Sinclair, 2008; Wei, 2012; He, 2013, 2015; Pu, 2020). Operating within a hierarchy that progresses from vocabulary to grammar and ultimately to semantics and pragmatics, this model revolves around the node word, facilitating a systematic examination of its co-selection patterns. For instance, a EUM analysis of “default” concordance lines in Figure 2 reveals its frequent collocates such as “on” and “rate”, which form verbal phrases like “default on” and nominal phrases like “default rate”. Examining its recurring lexico-grammatical patterns, we find that “default” functions as a verb, often in patterns like “Noun [borrower /securities/loan] default” or “Noun 1 [government/buyer/homeowner] default on Noun 2 [debt/obligations/mortgages/loans]” and as a noun in expressions like “Noun [risk/probability] of default”, “default rate” and “Verb [be/go] in/into default”.

In most of these collocates and patterns, “default” maintains a core meaning associated with the failure to repay money, as evidenced in the third concordance line in Figure 2. However, two exceptions are found in Line 5 and Line 11. The former relates to computer science, where default signifies a specific set of instructions that a computer follows unless instructed otherwise. This meaning is likely transferred from computer science to business, given its contextual relevance in Line 3 when discussing Facebook’s internet homepage. In the latter exception which is less businesslike and more culture-relevant to cross-cultural business communication, “cultural default mode” is used to denote a state in which individuals revert to their familiar cultural norms, values, and behaviors when faced with a different culture. In both exceptional lines, “default” is used in a different pattern: default + Noun [mode/homepage].

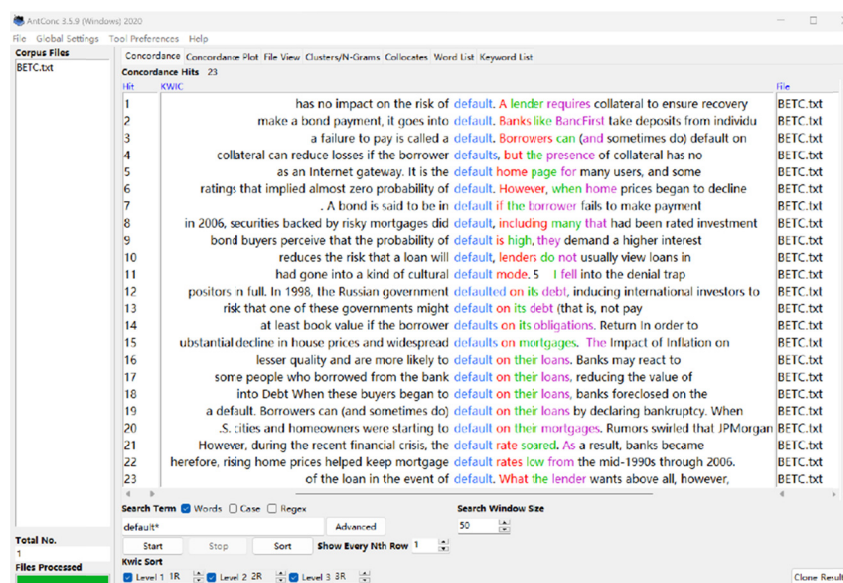



Figure 2. Concordance lines of “default” in BETC






The concordance analysis of “default” in the context of business English aligns with Chung & Nation’s emphasis on the importance of elucidating to students how a word’s new or technical usage is connected to its basic meaning and how it might differ in terms of “collocations and grammatical patterns” (Chung & Nation, 2003, p. 113). This analysis, embracing these aspects, also incorporates both semantic preference and semantic prosody. The semantic preference of “default” can be encapsulated as its usage to denote vulnerability to repayment, often stemming from poor financial standing. As for the semantic prosody, “default” is predominately used, except in the two exceptional concordance lines, to convey a derogatory sense of difficulty and failure in fulfilling financial obligations. In conclusion, this EUM-based corpus analysis of default’s concordance lines will serve as fundamental materials in the subsequent teaching case.






Step 3: Extraction of concordance lines from COCA and affordances of teaching scaffolding

To broaden students’ understanding of “default” in various contexts, our research extends the corpus analysis beyond the business domain to law and computer science, facilitating a cross-disciplinary approach for comparative analysis. Understanding the varied implications of semi-technical business English vocabulary in diverse contexts aids in comprehending their extended and metaphorical meanings (Li, Yu, & Lin, 2015). The frequencies of “default” in two sub-corpora in the COCA academic section, Law/PolSci (with 12 285 693 running words) and Sci/Tech (with 17 454 120 running words), are demonstrated in Figure 3, strongly proving that this word is frequently used in these two disciplines in academic genres. We then extracted concordance lines of “default” in these two sub-corpora and transformed them into text files, referred to as COCAL and COCAST. These files have been meticulously cleaned and prepared to serve as accessible resources for students’ hands-on search in classroom teaching which will be detailed in the next section.



Corpus of Contemporary American English

SEARCH

FREQUENCY

CONTEXT

ACCOUNT

SEE CONTEXT: CLICK ON WORD (ALL SECTIONS) OR NUMBER (SPECIFIED SECTION)

SEE # TEXTS [\[HELP...\]](#)

SEC 1 (ACAD:Law/PolSci): 12,285,693 WORDS

WORD/PHRASE	TOKENS 1	TOKENS 2	PM 1	PM 2	RATIO
1 DEFAULT	455	203	37.0	11.6	3.2

SEC 2 (ACAD:Sci/Tech): 17,454,120 WORDS

WORD/PHRASE	TOKENS 2	TOKENS 1	PM 2	PM 1	RATIO
1 DEFAULT	203	455	11.6	37.0	0.3

Figure 3. Query result of “default” in two COCA academic sub-corpora

In light of the considerably higher frequency of “default” in both COCA sub-corpora compared to BETC, we opted for the collocate function of AntConc for collocation analysis, as shown in Figure 4. By setting the frequency threshold for L1 and R1 collocates at 10, we identified the most frequent collocates within COCAL. Two notable collocations are “default rule(s)/principle” and “default on”. Another two collocates, “benefit” and “by”, can be observed in the phrases “benefit default” and “by default”. This information, represented in Figure 4, provides clickable access to the corresponding concordance lines in which “default” and the collocate co-occur, as demonstrated in Figure 5. Reading these co-occurring concordance lines can facilitate students’ discovery of default’s patterns when it co-occurs with these collocates. Replicated processes were applied to COCAST for analyzing “default” in the computer science sub-corpus, although the details have been omitted due to spatial constraints and the repetitive nature of the procedures.

Rank	Freq	Freq(L)	Freq(R)	Stat	Collocate
1	68	0	68	4.41517	principles
2	51	0	51	4.37318	rule
3	10	10	0	4.36845	penalty
4	48	48	0	4.33381	benefit
5	82	0	82	4.09422	rules
6	31	28	3	3.03110	by
7	13	2	11	2.40119	or
8	54	49	5	2.19655	a
9	14	14	0	2.17206	s
10	10	0	10	1.86595	on
11	14	8	6	1.51998	for
12	10	9	1	1.44562	are
13	87	82	5	1.29471	the
14	43	40	3	1.28297	of
15	22	13	9	1.20218	in
16	16	1	15	0.85013	and
17	12	6	6	-0.33430	to

Figure 4. Collocate list of “default” in COCAL

Line	Text
1	can acquire its property interest upon default by means of the judicial process and
2	erest. It is typically nonpossessory until default by the debtor and it can be
3	often, for some period of time, "win" by default: a failure to reach an accommodation
4	re to adopt enforceable standards will, by default, allow countries that exploit their workers
5	hat is, with Batista's candidate winning by default, Ambassador Smith tried to blunt the
6	they wind up dealing with the WTO by default, and as I mentioned, they will
7	y filing pendency actions that they won by default, and obtained temporary restraining orders to
8	live security operations, the NACC may by default assume that responsibility. As a bridge
9	requirements of 26 U.S.C. 501(c) (3) are by default assumed to be operating as religious
10	in nature, territorial sovereignty would by default be viewed as extending beyond the
11	industrial policy is "Yes". It is formed by default by the sum total of the
12	is more common that they are decided by default (721 defendants or 9%) or by summary disposi
13	is meant by other-regarding is defined by default fiduciary duties of loyalty"). # 118. See supra
14	a hearing (not including cases decided by default) (i.e., the same as Reduced
15	do math. # Rhode: Well, it does win by default, I think, for a fair number
16	end, 24 also backs up those messages by default on its iCloud servers in a
17	simply falling into the dystopian future by default -- or by complacently trusting in creative

Figure 5. Concordance lines including “default” and “by” in collocation

4. A CBLP Case of Default

This section uses the above corpus query results and processed data to craft a teaching case following the four principles of effective CBLP lessons proposed by Ma et al. (2021). The primary emphasis is on the teaching procedures that address the question “how to teach” with a focus on collocation and pattern analysis of the semi-technical business English word “default”. The framework provided by Ma et al. (2021) serves as the guiding structure for the design of corpus-based teaching activities.

- 1) Testing student knowledge (detect lexical errors)
- 2) Hands-on corpus searches by students (observing and analyzing the language)
- 3) Inductive discovery by students (summarizing the language use pattern)
- 4) Output exercise (practice using the language)

(Ma & Mei, 2021, p. 186)

Subsequently, a 4-step CBLP lesson is outlined to showcase how “default” is taught in a business English classroom.

4.1 Lesson Background

Business English undergraduates feel confused on the meaning and usage of “default” in their reading materials, exemplified in the text “Inflation, Banking, and Economic Growth”, part of the unit of “Money and Banking” in the business English textbook. While these students are familiar with its meaning in computer science, where “default” refers to preset configurations of a computer, their understanding within the business domain and potentially across other disciplines remains limited. Illuminating these disparities and interconnections between the general and technical connotations of the term falls within the English instructor’s purview (Farrell, 1990). This teaching case within the CBLP framework, set in the reading classroom, offers significant value for business English learners to acquire cross-disciplinary vocabulary knowledge.

4.2 Step 1: Guided Questions for Knowledge-Testing & Consciousness-Raising

The instructor uses a series of questions, derived from the quoted text lines, to evaluate students’ grasp of “default” within the specific business context. These questions intend to assess their comprehension of its meaning in this context, its diverse parts of speech, as well as their knowledge of associated lexical and grammatical structures. Simultaneously, these questions aim to enhance students’ awareness of the semantic dynamics and the form-meaning correspondence in semi-technical vocabulary.

Quoted text lines: “Higher inflation can decrease the real rate of return on assets. Lower real rates of return discourage saving but encourage borrowing. At this point, new borrowers entering the market are likely to be of lesser quality and are more likely to default on their loans” (Cai, 2009).

Question 1: How can “default on loans” be paraphrased in the context?

Question 2: Apart from its use as a verb, what other parts of speech does it exhibit?

Question 3: In what other phrases or structures can “default” be used?

Question 4: In which specific business situations or topics might this term be used?

Question 5: Does this word in a business context hold any particular attitude or emotional connotation?

Question 6: Beyond its use in the business context, are you familiar with its meanings and forms in other disciplines or subjects?

4.3 Step 2: Guided Hands-On Corpus Search by Students

The instructor begins by demonstrating the use of AntConc for retrieving concordances of “default” within BETC, as illustrated in Figure 2 in the preceding Section 3.3. Additionally, students are instructed to use the software’s collocate function to explore its high-frequency collocates (L1 and R1). Informed by the questions and answers in Step 1, both teacher and students collaborate to explore various facets of the search term “default” within BETC. This includes its meaning, part of speech, collocates, grammatical structures, semantic preferences, and semantic prosody.

Following the instructor’s demonstration, students autonomously employ AntConc to conduct searches for the word “default” within COCAL and COCAST provided by the instructor. With the instructor’s guidance, they retrieve collocate lists (reference as Figure 4 in Section 3.3 as in COCAL), and corresponding concordance lines depicting the co-occurrence of “default” with its high-frequency collocates (as seen in Figure 5 in Section 3, as

in COCAL). COCAST equivalents are also extracted for comparative analysis. These resources serve as a foundation for the subsequent phase of inductive discovery.

4.4 Step 3: Guided Inductive Discovery by Students

In Step 3, students engage in a guided inductive process. Empowered with the corpus resources extracted from two COCA sub-corpora, and insights from the instructor's BETC corpus results presentation, students undertake a comprehensive comparative analysis. By reading collocate lists and conducting an EUM analysis of concordance lines, students consolidate their findings. They derive an intricate comprehension of the word "default", unraveling its nuanced meanings and complex language structures across three distinct academic disciplines: business, law, and computer science. This process extends the collocation and colligation analysis to reveal the subtler facets of semantic preferences and semantic prosody, thereby portraying a more profound profile of the term's complexities in these distinct academic domains.

This step not only aids students in grasping the intricate uses of "default" but also equips them with the tools and high-order thinking skill necessary for effective language learning. It enriches their linguistic and cognitive abilities within academic discourses. Furthermore, through hands-on corpus searches and generalization, students are expected to cultivate self-directed learning and inductive skills by working as "language detectives" (Johns, 1997, p. 101).

4.5 Step 4: Vocabulary Knowledge Consolidation Through Productive Exercises

To consolidate students' productive knowledge of the word "default", the instructor devises targeted vocabulary exercises for after-class practice on corpus query findings. These exercises are structured to accommodate second-year business English undergraduates in mainland China, accounting for their intermediate English proficiency and adequate understanding of business concepts. Two exercises, "Sentence Rearrangement" and "Gap Filling", are tailored to meet their needs, which are outlined in Appendix B. It's essential to note that, depending on the students' proficiency levels, other productive vocabulary exercise formats such as "True/False Statements", "Word Matching", "Word Sorting/Categorization" and "Cloze" may be more relevant, and can be accordingly integrated.

5. Pedagogical Implications

This study offers multiple pedagogical implications for business English vocabulary instruction and the broader spectrum of ESP instruction. It highlights the semantic dynamism and diverse forms of semi-technical business English vocabulary through corpus analysis, presenting an innovative pedagogical approach that bridges corpus analysis and ESP instruction. Its pedagogical implications are articulated as follows.

The incorporation of corpus linguistics into language instruction, facilitated by corpus analysis and the CBLP teaching model, promises to provide valuable insights for both ESP instructors and learners. This integration fosters a deeper understanding of semi-technical vocabulary based on rich disciplinary corpus resources. The structured teaching procedures offer a systematic framework for specialized vocabulary instruction. This model accommodates both the receptive and productive knowledge of language and can be used to provide students with a comprehensive vocabulary foundation, particularly focusing on words characterized by inter-disciplinary complexities.

The pedagogical model advances interdisciplinary comparative learning and inductive discovery, fostering self-directed learning among students. This approach nurtures a holistic understanding of the technicality of ESP discourses, facilitating communication in diverse professional contexts. It also equips students with the skills to explore, understand, and use semi-technical vocabulary by publicly available and self-constructed corpora, along with corpus tools.

Moreover, the study underscores the inherent value of corpus analysis techniques, including keyword analysis, concordance analysis and collocate analysis, with a particular focus on EUM-framed concordance analysis. This involves studying collocations, colligations, semantic preferences, and semantic prosody. These corpus linguistics techniques serve as valuable tools for researchers and educators engaging in register analysis of ESP discourses.

6. Conclusion

This study explores semi-technical business English vocabulary, a challenging and often overlooked aspect of ESP analysis and teaching. It effectively addresses questions of "how to identify" and "how to teach". The employed corpus analytical techniques offer valuable insights for ESP researchers and educators engaged in interdisciplinary studies. From the identification and analysis of semi-technical business vocabulary to classroom

pedagogy design, this research fully integrates corpus resources and techniques which adds vitality and innovation to ESP instruction.

By bridging the gap between corpus analysis and language instruction, our research holds the potential to enhance business English teaching and equip learners with the essential corpus tools and skills necessary to navigate the complexities of business English discourses. The insights drawn from the CBLP teaching case offer invaluable guidance for instructors seeking innovative strategies to impart semi-technical vocabulary within business English, and by implication, within any branch of ESP.

However, it is essential to acknowledge that the effectiveness and efficiency of this innovative pedagogical approach require further support from future empirical studies, particularly based on students' feedback.

Funding

This study was funded by the Thirteenth Five-Year Plan Discipline Co-construction Project by the Guangdong Planning Office of Philosophy and Social Science (Project number GD20XWY10).

References

- Baker, M. (1988). Sub-technical vocabulary and the ESP teacher: An analysis of some rhetorical items in medical journal articles. *Reading in a Foreign Language*, 4(2), 91–105.
- Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing*, 8(4), 243–257. <https://doi.org/10.1093/lc/8.4.243>
- Bondi, M., & Scott, M. (2010). *Keyness in texts*. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/scl.41>
- Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta - analysis. *Language Learning*, 67(2), 348–393. <https://doi.org/10.1111/lang.12224>
- Browne, C., & Culligan, B. (2016). *A new business service list*. Retrieved from <http://www.newgeneralservicelist.org/bsl-business-service-list/>.
- Cai, Y. (2009). *Business English: A comprehensive course*. Beijing: Higher Education Press.
- Camiciottoli, B. C. (2007). *The language of business studies lectures*. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/pbns.157>
- Chen, Q., & Gao, Y. (2010). Semi-technical vocabulary in academic English writings. *Foreign Language Education*, 6, 42–46.
- Chung, T. M., & Nation, P. (2003). Technical vocabulary in specialized texts. *Reading in a Foreign Language*, 2, 103–116.
- Chung, T. M., & Nation, P. (2004). Identifying technical vocabulary. *System*, 32(2), 251–263. <https://doi.org/10.1016/j.system.2003.11.008>
- Collins, H. (2018). *Collins COBUILD advanced learner's dictionary*. Harper Collins Publishers.
- Cowan, J. R. (1974). Lexical and syntactic research for the design of EFL reading materials. *TESOL Quarterly*, 4, 389–399. <https://doi.org/10.2307/3585470>
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213–238. <https://doi.org/10.2307/3587951>
- Coxhead, A. (2018). *Vocabulary and English for specific purposes*. NY: Routledge. <https://doi.org/10.4324/9781315146478>
- Crosthwaite, P. (2017). Retesting the limits of data-driven learning: Feedback and error correction. *Computer Assisted Language Learning*, 30(6), 447–473. <https://doi.org/10.1080/09588221.2017.1312462>
- Dun, G. G. (2002). Characteristics and translation of words in English for business and economics. *Shandong Foreign Language Teaching*, 3, 37–40, 57.
- Farrell, P. (1990). *A lexical analysis of the English of electronics and a study of semi-technical vocabulary*. Dublin: Trinity College.
- Flowerdew, J. (1993). Concordancing as a tool in course design. *System*, 21(2), 231–244. [https://doi.org/10.1016/0346-251X\(93\)90044-H](https://doi.org/10.1016/0346-251X(93)90044-H)
- Fraser, S. (2007). Providing ESP learners with the vocabulary they need: Corpora and the creation of specialized

- word lists. *Hiroshima Studies in Language and Language Education*, 10, 127–143.
- Fraser, S. (2012). Factors affecting the learnability of technical vocabulary: Finding from a specialized corpus. *Hiroshima Studies in Language and Language Education*, 15, 123–142.
- Gardner, D. (2007). Validating the construct of word in applied corpus-based vocabulary research: A critical survey. *Applied Linguistics*, 28(2), 241–265. <https://doi.org/10.1093/applin/amm010>
- Gilquin, G., & Granger, S. (2010). *How can data-driven learning be used in language teaching?* Routledge. <https://doi.org/10.4324/9780203856949-26>
- He, A. P. (2013). From the philosophy of “meaning as use” to the exploration of “unit of meaning” in corpus linguistics. *Foreign Languages and Their Teaching*, 3, 44–48. <https://doi.org/10.13458/j.cnki.flatt.003915>
- He, A. P. (2015). On vocabulary development in streamlined EFL course books: A perspective of dynamic system theory. *Foreign Language Teaching and Research*, 6, 898–908.
- He, Y., & Zhang, J. D. (2011). Semantic research on semi-technical terms in scientific and technical English. *Journal of Donghua University*, 3, 193–199.
- Hsu, W. (2013). Bridging the vocabulary gap for EFL medical undergraduates: The establishment of a medical word list. *Language Teaching Research*, 17(4), 454–484. <https://doi.org/10.1177/1362168813494121>
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139524773>
- Hunston, S., & Francis, G. (2000). *Pattern grammar: A corpus-driven approach to the lexical grammar of English*. Amsterdam: John Benjamins. <https://doi.org/10.1075/sc1.4>
- Hyland, K., & Tse, P. (2009). Academic lexis and disciplinary practice: Corpus evidence for specificity. *International Journal of English Studies*, 9(2), 111–129.
- Johns, T. (1991). Should you be persuaded: Two examples of data-driven learning. *English Language Research Journal*, 4, 1–16.
- Konstantakis, N. (2007). Creating a business word list for teaching English. *Estudios de lingüística inglesa aplicada*, 7, 79–102.
- Le, C. N. N., & Julia, M. (2023). A core meaning-based analysis of English semi-technical vocabulary in the medical field. *English for Specific Purposes*, 70, 252–266. <https://doi.org/10.1016/j.esp.2023.01.006>
- Leech, G. (1997). Teaching and language corpora: A convergence. In A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles (Eds.), *Teaching and language corpora* (pp. 1–23). Longman. <https://doi.org/doi:10.4324/9781315842677-1>
- Li, C. F., Yu, J., & Lin, J. (2015). Business English vocabulary: description, identification and pedagogy. *Foreign Language World*, 3, 83–88, 96.
- Li, S. L. E., & Pemberton, R. (1994). An investigation of student knowledge of academic and sub-technical vocabulary. In I. Flowerdew & A. K. K. Tong (Eds.), *Emerging Text* (pp. 183–196). Hong Kong: The Hong Kong University of Science & Technology.
- Ma, Q., Chui, M. M., Lin, S., & Mendoza. (2023). Teachers’ perceived corpus literacy and their intention to integrate corpora into classroom teaching: A survey study. *ReCALL*, 35(1), 19–39. <https://doi.org/10.1017/S0958344022000180>
- Ma, Q., & Mei, F. (2021). Review of corpus tools for vocabulary teaching and learning. *Journal of China Computer-Assisted Language Learning*, 1(1), 177–190. <https://doi.org/10.1515/jccall-2021-2008>
- Maher, P. (2016). The use of semi-technical vocabulary to understand the epistemology of a disciplinary field. *Journal of English for Academic Purposes*, 22, 92–108. <https://doi.org/10.1016/j.jeap.2016.01.010>
- Menon, S., & Jayakaran, M. (2010). Analysing collocational patterns of semi-technical words in science textbooks. *Pertanika J. Soc. Sci. & Hum.*, 18(2), 241–258.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>
- Nesi, H. (2013). ESP and corpus studies. In B. Paltridge & S. Starfields (Eds.), *The handbook of English for specific purposes* (pp. 407–426). Wiley-Blackwell. <https://doi.org/10.1002/9781118339855.ch21>
- Parkinson, D., & Joseph, N. (2005). *Oxford business English dictionary for learners of English*. Oxford: Oxford

University Press.

- Pu, J. Z. (2020). Reinterpretation of the model of extended units of meaning. *Foreign Languages Research*, 2, 1–8.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Sinclair, J. (2008). The phrase, the whole phrase, and nothing but the phrase. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 407–410). Amsterdam & Philadelphia: John Benjamins Publishing Company.
- Tangpijaikul, M. (2014). Preparing business vocabulary for the ESP classroom. *RELC Journal*, 45(1), 51–65. <https://doi.org/10.1177/0033688214522641>
- Trimble, L. (1985). *English for science and technology: A discourse approach*. New York: Cambridge University Press.
- Watson-Todd, R. (2017). An opaque engineering word list: Which words should a teacher focus on? *English for Specific Purposes*, 45, 31–39. <https://doi.org/10.1016/j.esp.2016.08.003>
- Wei, N. X. (2012). Theory of co-selection and corpus-driven phraseology. *Journal of PLA University of Foreign Languages*, 1, 1–6, 74.
- Xu, J. J., & Liang, M. C. (2011). *CROWN corpus*. National Research Center for Foreign Language Education, Beijing Foreign Studies University.
- Yang, H. Z. (1986). A new technique for identifying scientific/technical terms and describing science texts. *Literary and Linguistic Computing*, 1(2), 93–103. <https://doi.org/10.1093/lc/1.2.93>
- Zhen, F. C., & Wang, H. (2014). The rhetorical functions of sub-technical vocabulary: An analysis of evidence, research and result in linguistics journal articles. *Foreign Languages and Their Teaching*, 5, 54–60.

Appendix A

General and specialized meanings of some semi-technical business words

Table A1. General and specialized meanings of security[noun]

Semi-technical vocabulary	General meaning	Meaning in Business
security	Measures taken to protect a place or ensure that only people with permission enter it or leave it	1. [plural]A financial asset such as a share or bond [Finance] 2. A valuable item that you agree to give to somebody if you are unable to pay back the money that you have borrowed from them[finance]

Table A2. General and specialized meanings of bond [noun]

Semi-technical vocabulary	General meaning	Meaning in Business	Meaning in Law
bond	1. A strong feeling of friendship, love or shared beliefs and experiences that unites them 2. A close connection that people have with each other 3. The way two things stick to each other or are joined	1. An agreement by a government or an organization to pay back the money an investor has lent plus a fixed amount of interest on a particular date[finance] 2. A word used for certain kinds of insurance policies that protect companies from loss[insurance]	Legal written agreement or promise

Table A3. General and specialized meanings of default [noun]

Semi-technical vocabulary	General meaning	Meaning in Law	Meaning in Computer Science
default	A default situation is what exists or happens unless someone or something changes it	Failure to do something that is required by an agreement or by law, especially paying a debt	What happens or appears if you do not make any other choice or change

Table A4. General and specialized meanings of return [noun]

Semi-technical vocabulary	General meaning	Meaning in Business	Meaning in Computer Science
return	An arrival back at the place where you have been before	1. The amount of profit or income that you get from a particular investment[finance] 2. An official report or statement that gives particular information about something to an official body[accounting] 3. Goods that a customer has bought or ordered and then returned/the acting of returning a product[commerce]	The button that you press on a computer when you reach the end of an instruction or to begin a new line

Table A5. General and specialized meanings of interest [noun]

Semi-technical vocabulary	General meaning	Meaning in Business	Meaning in Law
interest	1. If you have an interest in something, you want to learn or hear more about it 2. If something is in the interests of a particular person or group, it will benefit them in some way	1. The extra money that you have to pay when you borrow money 2. The extra money that you receive when you invest money 3. A share in a business or company and a legal right to a share of its profits 4. A connection with something that affects your attitude to it, especially because you may benefit from it in some way[finance]	A legal right to land or property

Appendix B

CBLP teaching case Step 4 Vocabulary Knowledge Consolidation Through Productive Exercises

Sentence Rearrangement Task: Rearrange the given words in scrambled sentences to form grammatically correct sentences.

Scrambled Sentences:

1. default, under control, the risk, is, to ensure, collateral, requires, a lender, of
2. default, low, home prices, helped, mortgage, rising, keep, rates
3. default, widely, has, rule, risk aversion, accepted, become a
4. default, in the absence of, a specific choice, the system, setting, will, the, select

Rearranged Sentences for reference:

1. A lender requires collateral to ensure the risk of default is under control.
2. Rising home prices helped keep mortgage default rates low.
3. Risk aversion has become a widely accepted default rule.
4. In the absence of a specific choice, the system will select the default setting.

Gap Filling Task: Fill in the gaps with appropriate words to complete the sentences.

1. If the company fails to make a bond payment, it goes ____ default.
2. The remedy is revoking the common stock, and most actions are decided ____ default.
3. Banks may react to lesser quality borrowers by rationing credit, as they are more likely to default _ their loans.
4. When bond buyers perceive that the _____ of default is high, they demand a higher interest.

Reference answers: 1. into, 2. by, 3. on, 4. probability/risk

Copyrights

Copyright for this article is retained by the author, with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).