

Ha, A.Y.H. & Hyland, K. (2017). What is technicality? A Technicality Analysis Model for EAP vocabulary. Journal of English for Academic Purposes. 28: 35-49.

## **What is technicality? A Technicality Analysis Model for EAP vocabulary**

### **Abstract**

The identification of technical words for teaching discipline-specific EAP courses remains a problem for materials designers and teachers alike. This study proposes a method that identifies technicality and measures the degree of technicality of a word. The Technicality Analysis Model (TAM) suggests five levels of technicality: *least technical*, *slightly technical*, *moderately technical*, *very technical* and *most technical*. In identifying technicality we take four factors into account: 1) both general and specialised senses of a word; (2) the banding of a word in reference word lists; (3) the polysemy of a word; (4) the literal meaning of a word. The set of categorisation criteria is stringent in the sense that even *least technical words* may have specialised senses in a specific discipline but those senses may be almost the same as the general sense. All words in more technical categories have specialised senses. We trialled the TAM with 837 financial-sector-specific words generated from a 6.7-million-word corpus of financial texts. Results show that with the categorisation criteria in the technicality analysis, every financial-sector-specific word could be categorised into one of the technical word categories. Future research may use the TAM to develop a repertoire of discipline-specific vocabulary for EAP teaching and learning.

### **Keywords**

Technicality Analysis Model; specialised vocabulary; technical vocabulary; technicality; terminology

### **Highlights**

- A Technicality Analysis Model is proposed to measure the technicality of a word.
- Criteria include dictionary and literal meanings, frequency, and polysemy.
- The model was trialled with financial-sector-specific words.
- 837 words were categorised along a continuum of five degrees of technicality.
- Future research may use the model to build discipline-specific word lists for EAP.

# What is technicality? A Technicality Analysis Model for EAP vocabulary

## 1. Introduction

Technical vocabulary is widely recognised as playing a fundamental role in specialised texts and is central to the teaching and learning of disciplinary specific English in EAP courses (e.g. Chung and Nation, 2003; Woodward-Kron, 2008; Kwary, 2011). The acquisition of technical vocabulary is, however, often a major problem for students. Large-scale surveys by Hyland (1997), Evans and Green (2007) and Evans and Morrison (2011) found that specialist vocabulary was among the biggest challenges facing undergraduate students in Hong Kong while Wu and Hammond (2011) obtained similar results with non-native English speakers in a Canadian university. Yet while there is considerable research regarding the nature and behaviour of high frequency academic words (e.g. Coxhead, 2000; Gardner & Davies, 2014) a reliable method for establishing which words are actually technical remains.

Technical vocabulary is important in EAP classes as it helps learners develop their subject knowledge. EAP teachers, however, often lack the specific field knowledge to develop suitable teaching materials about technical vocabulary and often feel vulnerable in this area (e.g. Spack, 1988). The ability to identify technical vocabulary is a pre-requisite for the teaching of discipline-specific subject matter (Chung and Nation, 2004), but EAP teachers lack clear guidelines to understand the nature of vocabulary (e.g. Kwary, 2011; Mukundan and Yu, 2012; Woodward-Kron, 2008). Woodward-Kron (2008, p. 235), for example, observes that

..there appears to be little or no discussion of the role of the specialist language of students' chosen discipline for learning nor any guidelines for teachers to understand the nature of the specialist language of different academic disciplines... many EAP teachers and language and learning advisors have backgrounds in education and applied linguistics, yet in their work with students they regularly encounter unfamiliar disciplines and need to engage with the specialist language of those disciplines in their language advising role.

An understanding of specialist technical vocabulary is therefore indispensable to those designing EAP courses and materials. A means of assisting EAP teachers to assess how technical a word is in order to create discipline-specific word lists for their classes would therefore be invaluable. This would enable them to select words of appropriate technicality levels from the reference discipline-specific word lists according to students' needs and abilities.

This study is an attempt to address the definitional issue of *technicality* which underpins pedagogic decisions by proposing a Technicality Analysis Model (TAM) that considers different dimensions of a

word. Four factors concerning senses, frequency, polysemy, and literal meaning are amalgamated into a set of criteria which capture and categorise the technicality of a word. This shows that technicality is not a binary term; we cannot say a word is either technical or not. Instead our model allows researchers to categorise words along a continuum with five degrees of technicality. In ascending order of technicality, technical words are *least technical* (TAM 1), *slightly technical* (TAM 2), *moderately technical* (TAM 3), *very technical* (TAM 4) and *most technical* (TAM 5).

The study analysed word types instead of word families. Most previous lexical studies used the word family as the unit of analysis (e.g. Sutarsyah, Nation & Kennedy, 1994; Chung & Nation, 2004; Hsu, 2011) while recent studies began to give more attention to word types (e.g. Durrant, 2014; Ward, 2009). A word-type-based threshold has an edge over a word-family-based threshold in cases where one of the word family members occurs frequently enough to be included in the analysis while other word family members are not. A word-type-based threshold can capture such a word whereas the word might be filtered out using a word-family-based threshold.

In what follows we intend to address the following questions:

1. How can technicality be operationalised when studying vocabulary?
2. How can particular words be categorised in terms of degrees of technicality?
3. What words and combinations are specific to finance and how technical are they?

We first review the main descriptions of technical vocabulary before introducing our model and go on to discuss how this was trialled with a large, specially developed corpus of financial-sector-specific vocabulary. We conclude with a discussion of the significance and implications of the model and suggestions for future research.

## **2. Characterisations of Technical Vocabulary**

Previous studies made attempts to create words lists for specific discipline areas and their findings reveal that a technical text contains a substantial amount of technical words (Nation, 2016). A number of studies have attempted to define *technical vocabulary* or *technical words* (e.g. Martin, 1976; Paquot, 2010; Schmitt, 2010; Valipouri & Nassaji, 2013) yet no consensus has been reached on a common set of features. Most researchers relate technicality to specialisation and a particular community of users. For example, Valipouri and Nassaji (2013, p. 249) state that technical vocabulary “includes words closely related to a specific sub-field and not frequent in other fields”. Paquot (2010, p. 13) also believes that technical terms are “domain-specific” and “require scientific knowledge to understand” while Mudraya (2006, p. 238-239) identifies technical terms as those which have “no exact synonyms and have a very narrow range of interpretations within a particular field”. Thus the notion of infrequency and having a narrow, specialised range of use are commonly invoked as indicators of technicality (Mudraya, 2006; Nation & Newton, 1997; Valipouri & Nassaji, 2013).

Nation (2013) recognises the overlap between technical words and other categories of words and regards technical words as specialised.

But while some technical vocabulary may be restricted to a given domain of users, the picture is complicated by the fact that many items may be relatively high frequency forms in general use which can carry specialised meanings within a particular field (Schmitt, 2010). As Dudley-Evans and St. John (1998, p. 83) observe: “The specialised uses of general vocabulary in specific disciplines we would regard as an aspect of technical vocabulary”. Mudraya (2006, p. 238-239) also refers to “strictly technical” words, hinting at the possibility of a cline of technicality and taking semantic properties into account by considering whether a word has an exact synonym and is resistant to semantic change or not. Thus words in general use might also have specialised uses, leading to the additional complication that a technical vocabulary might contain homonyms and polysemes and that a word might not necessarily be unique to a field. Thus Ward (2007, p. 23) suggests “two strands of meaning in the word ‘technicality’: specialisation and difficulty”. Nation (2016, p.7) sees two types of specialist vocabularies – “those are commonly known by people who are not specialists in the field” and “those that are typically only known by specialists”. We would not, however, exclude words that also have a general sense so that specialisation may be a feature of one meaning of a word but accessibility that of another meaning of the same word.

Moving beyond these general characterisations, various methods have been used to identify and measure the technicality of words. The intuitions of disciplinary experts, either directly through interviews or by consulting specialist dictionaries, have been used although these are often laborious, risk excluding terms through subjective judgements and are not always effective for identifying non-specialist terms (Chung & Nation, 2004; Kwary, 2011). Pearson (1998) investigated the metalanguage patterns describing terms in specialised corpora as a means of identifying technical terms, although this seems very dependent on individual writers’ views. Another approach is corpus-comparison, whereby word frequencies in a technical text are compared with those in a reference corpus and words that occur more frequently in the technical corpus are seen as technical (e.g. Sutarsyah, Nation & Kennedy, 1994). This method is unlikely to be comprehensive, however, as it is often specialised collocations which take on technical meanings and neither multi-word units or everyday words with technical meanings will show up as technical in such comparisons. Moreover, like keyword analyses, comparisons fail to distinguish the word class, so that the high-frequency word *stock*, for example, will have a technical sense as a noun but not as a verb (Kwary, 2011).

To overcome limitations of the existing methods of identifying technical terms, Kwary (2011) proposed a two-step method which subjects the keywords identified by corpus comparisons to systematic classifications by disciplinary experts, a rather cumbersome approach that still risks missing key items. Perhaps the most comprehensive rating scale developed thus far is Chung and

Nation's (2003) four-step model, with words having no relationship to the field eliminated in step 1 and those remaining in step 4 having specific meanings to the discipline and not likely to be known to outsiders. Most problematic were words in step 3 which had meanings, for example, closely related to anatomy, such as *chest*, *neck*, *heart* and *muscles*, but which also "occur in general use with little change in meaning" (p. 105). Thus polysemes continue to create difficulties for categorising items which we attempt to resolve in our model by categorising words in terms of relative degrees of technicality and using more refined criteria.

Given the significance of discipline-specific vocabulary in the everyday work of academic disciplines and the teaching and learning of their discourses, we here propose a systematic and principled method to assess the technicality of words along a cline of five degrees of technicality. We trialled the method with 837 financial-sector-specific words identified from a self-built financial corpus of 6.7 million words. We introduce the model and describe the corpus in following sections.

### **3. The Technicality Analysis Model**

#### **(i) Key terms and instruments**

The Technicality Analysis Model is essentially a set of criteria that measures the degree of technicality of a word and was trialled with financial-sector-specific words. Every financial-sector-specific word was analysed and categorised along a continuum using the technicality analysis. The five degrees along the continuum are *least technical* (TAM1), *slightly technical* (TAM2), *moderately technical* (TAM3), *very technical* (TAM4), and *most technical* (TAM5), in order of ascending technicality.

The degree of technicality of every financial-sector-specific word was assessed using multiple instruments including the *Cambridge Advanced Learners Dictionary & Thesaurus*, *Cambridge Business English Dictionary*, the Concord function in *WordSmith Tools*, the random concordancing function in *WordBanks*, and other reference sources such as the *Merriam-Webster Online Dictionary* and *Investopedia Dictionary* whenever necessary. Technical dictionaries are compiled by subject experts and have been used in technicality studies to access the assessments of disciplinary specialists (e.g. Chung & Nation, 2003). We also used the New General Service List (NGSL) (Browne, Culligan, & Phillips, 2013) and the set of twenty-five 1000 word family lists from the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA). The NGSL, which was built upon the original General Service List of English Words (West, 1953), provides a register of general high frequency vocabulary while the BNC/COCA sub-lists are well-established word frequency lists which can provide credible banding information of each word in the study. The 1st and 2nd

BNC/COCA sub-lists are comparable to the non-specialist NGSL while the 3rd sub-list onwards use BNC/COCA rankings (Nation, 2012) to provide our study with word frequency information.

We also need to define a number of key terms used in the analysis. A *general sense* is defined as a sense shown in a general dictionary, which can be retrieved from the Cambridge Advanced Learners Dictionary & Thesaurus or the Merriam-Webster Online Dictionary. A *specialised sense* is defined as a sense shown in a specialised dictionary or glossary, which can be retrieved from the Cambridge Business English Dictionary or the Investopedia Dictionary. A word with only one entry in a general/specialised dictionary or with the same entry in a general dictionary and a specialised dictionary is regarded as *monosemous*. A word with two or more entries of different senses is considered *polysemous*. A word with eight or more entries of different senses in a general dictionary and/or a specialised dictionary is seen as *semantically depleted*. The sense of a semantically depleted needs to be derived from its context and is not available independently of it (Moon, 2010).

## (ii) Key Principles

The analysis attempts to unpack the construct of technicality by first distinguishing words according to their types of sense – general or specialised – then benchmarking them against established reference word lists as well as comparing the specialised sense(s) with the general sense(s). The issue of polysemy is also considered in the set of categorisation criteria in which *monosemous words* gain the upper hand in attaining a higher degree of technicality. The set of categorisation criteria is stringent in the sense that even *least technical* words may have specialised senses in finance but those senses are the same or almost the same as the general sense whilst all words in categories *slightly technical* or higher have specialised senses in finance.

The technicality analysis considers four factors:

1. both general and specialised senses of a word,
2. the banding of a word in reference word lists,
3. the polysemy of a word,
4. the literal meaning of a word.

The rationale for using these four factors was determined by these three key principles.

### **A word is more technical...**

#### *a. the more its specialised sense departs from its general sense.*

If a word's specialised and general senses are equivalent, the specialised sense is no longer considered "specialised" and so assigned a minimal degree of technicality. The more semantically the word's specialised sense departs from its general sense, the more technical the word's specialised sense is. Even if a word has very high occurrences in a large and representative corpus, it does not

rule out the possibility of its being technical because it may exhibit a specialised sense that is distinct from its general sense.

***b. the less frequently the word occurs***

Given two words of the same specialised sense, the lower-frequency word is considered more technical than the higher-frequency word on the grounds that fewer encounters in daily life lowers a word's chance of being understood. The degree of technicality increases as the word's chance of being understood decreases. Frequency information can be obtained by benchmarking a word against established word lists including the BNC/COCA sub-lists and the NGSL discussed above. The general principle is that with the same specialised sense, a word from a lower frequency band should be more technical than that from a higher frequency band.

***c. the word is monosemous***

The degree of technicality of a word is increased if the specialised sense is the only sense that a word has. Given the same specialised sense, a monoseme that cannot be understood literally or decoded from its word parts (e.g. *accretable*) is considered to be more technical than a monoseme that can be understood literally or decoded from its word parts (e.g. *policyholder*). Thus the *most technical* words are always monosemous and cannot be understood literally or decoded from their word parts.

**iii) Steps in the technicality analysis**

Using the instruments and principles sketched above, the technicality of a word can be assessed through successive steps. These are set out below and presented as a flow chart in Table 3.

1) Check if the word has a specialised sense

This can be done by consulting dictionaries. If the word does not have a specialised sense, the word will be categorised as TAM1 *least technical* and the technicality analysis ends. If the word has a specialised sense, the technicality analysis moves on to the second step.

2) Deal with polysemes

Compare the specialised sense with the general sense (if any). If the senses are the same or nearly the same, the word is categorised as TAM 1 *least technical* and the technicality analysis ends. If the senses differ, the technicality analysis moves on to the third step. This second step is skipped if the word does not have a general sense.

3) Identify the frequency band

Match the word against the NGSL and then the BNC/COCA sub-lists. If the word is within the NGSL or the 9<sup>th</sup> level BNC/COCA sub-list, its specialised sense is compared with its general sense to decide the semantic distance between them. This has two possible outcomes: Either its specialised

sense (1) contains more details or (2) is only remotely or not related to any of its related general sense(s). Depending on the banding of the word, it will then be categorised as TAM2 *slightly technical*, TAM3 *moderately technical*, or TAM4 *very technical*. The lexical coverage in the FC helps determine the cut-off point at the 9<sup>th</sup>/10<sup>th</sup> level BNC/COCA sub-lists. The lexical coverage in the FC of each BNC/COCA sub-list from 1<sup>st</sup> level to the 9<sup>th</sup> level is greater than 0.1% whilst each subsequent sub-list constitutes less than 0.1% of the FC.

i. A word is TAM2 *slightly technical* if it is in either the NGSL or the first and second BNC/COCA and has a specialised sense that contains more details than the related general sense (e.g. *capital* and *gearing*).

ii. A word is TAM3 *moderately technical* if (1) it is in either the NGSL or the first and second BNC/COCA sub-lists and is monosemous or its specialised sense is remotely or not related to any of its general senses; or (2) the word is in the third to ninth BNC/COCA and has a specialised sense that contains more details than the related general sense (e.g. *exposure* and *reinsurance*).

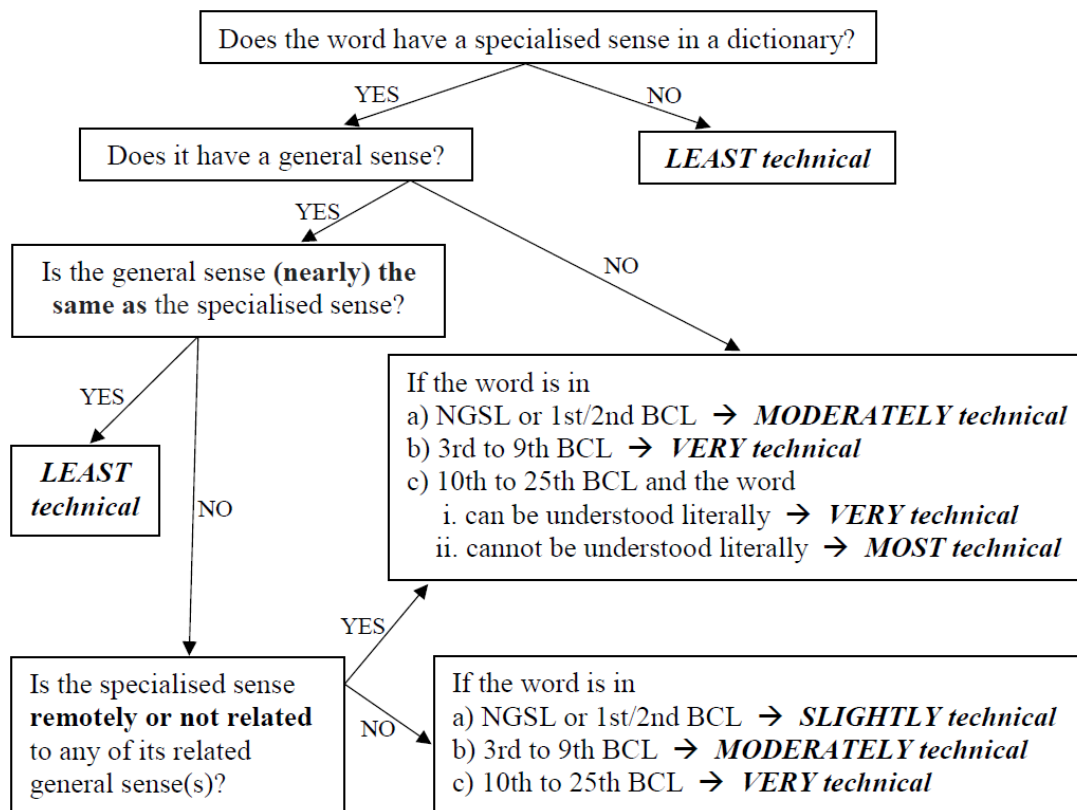
iii. A word is TAM4 *very technical* if (1) it is in the third to ninth BNC/COCA sub-lists and is monosemous or its specialised sense is remotely or not related to any of its general senses.

iv. A word is TAM 5 *most technical* if (1) it is in the 10th to 25th BNC/COCA, has a specialised sense, and cannot be understood literally

Table 3 A flow chart of the Technicality Analysis Model

---





#### 4. The Financial Corpus

The TAM was trialled with financial vocabulary using a large and representative financial corpus constructed for this purpose. Underpinning this decision to use discipline-specific vocabulary for the trial is the fact that specific vocabulary choices are core aspects of the particular practices, genres, and communicative conventions of academic or professional communities (Hyland, 2002). Hyland and Tse (2007), for example, show that lexical items cluster in individual disciplines, suggesting that the use of these items is discipline-specific. In other words, it is more likely that words with specific meanings to a community appear in highly specific texts. These words with specialised meanings are therefore ideal inputs for trialling our method. In addition, we selected Finance as the source of vocabulary as many readers will be familiar with financial terms in their everyday lives in a way which does not apply so much to technical terms in, say, engineering or chemistry.

The Financial Corpus (FC) consists of two central written and spoken genres of economics and finance: annual reports and earnings calls transcripts. Annual reports are company publications for reporting business and financial performances and contain a number of sections of different styles and linguistic features. Earnings calls, in contrast, are a spoken genre comprising both scripted presentations and spontaneous question-and-answer sessions which occur when a company releases its quarterly or annual results. The FC consists of 33 annual reports from 33 companies and 347 earnings call transcripts from 138 companies, totalling 6,753,212 words which are distributed almost evenly across the four financial sectors and between writing and speech (see Table 4). The corpus is a

representative sample of these two key genres compiled by collecting texts from the from the largest financial institutions ranked by market capitalisation on the Dow Jones Sector Titans Index (SL) (Dow Jones Indexes, 2012). This list ranks the top 60 listed companies by capitalisation globally in four financial sectors, namely Banks, Financial Services, Insurance, and Real Estate. The corpora were compiled by identifying the largest listed companies of each financial sector and collecting up to eight earnings call transcripts of each of these companies and their annual reports from their websites according to availability.

Table 4 Corpora word counts

	Financial				Totals
	Banks	Services	Insurance	Real Estate	
Annual Reports	881,428	838,954	848,062	805,784	3,374,228
Earnings Calls	866,423	841,215	844,982	826,364	3,378,984
Totals	1,747,851	1,680,169	1,693,044	1,632,148	6,753,212

Working with the assumption that highly technical words are likely to be more frequent in highly specific texts, we conducted keywords analyses to identify words that are specific to financial sectors using the academic sub-corpus in the BNC Baby as the reference corpus. Keywords analyses between financial sectors were then carried out to compile a financial-sector-specific vocabulary from the finance-specific vocabulary. A list of 837 words, accounting for 1,617,221 words (23.94%) in both corpora, were identified as specific to one or two financial sectors (see Table 5). These words comprised the data which we used to trial the Technicality Analysis Model. The model was trialled by the two authors and several graduate students working independently on a sample of 100 randomly selected items. Following the steps outlined below and the flowchart in Table 3, we arrived at the same TAM values with 95% agreement.

Table 5 Composition of the financial-sector-specific words

	Number of one-sector-specific words	Word Counts	Lexical coverage in all corpora (%)
Banks	181	283,854	4.20%
Financial Services	177	274,378	4.06%
Insurance	150	288,076	4.27%
Real Estate	204	273,727	4.05%
Specific to Two Sectors	125	497,186	7.36%
	837	1,617,221	23.94%

## 5. Conducting a Technicality Analysis

In this section, we explain the categorisation criteria of each of the five groups of technical words. We also explain, with examples, a number of possible scenarios which can emerge from a technicality analysis and show how decisions are made according to the categorisation criteria.

### a. Categorisation criteria for TAM 1 (least technical) words

Several instances in the TAM1 (*least technical*) vocabulary have been selected to illustrate how the categorisation works (see Table 6).

Table 6 Exemplifications of TAM 1 (Least technical)

Financial-sector-specific-word	General sense	Specialised sense	Analysis result
comparable (adj.)	Yes	No	TAM 1
headwinds (n.)	Yes	No	TAM 1
fee (n.)	Yes	Yes	same senses → TAM1
buybacks (n.)	Yes	Yes	almost same senses → TAM1

While listed in the NGSL and Cambridge Advanced Learners Dictionary & Thesaurus, the adjective *comparable* does not have an entry in the Cambridge Business English Dictionary and hence no specialised sense. It was easy to categorise the word as *least technical*. A similar instance without an entry in the Cambridge Business English Dictionary was the noun *headwinds*. Although the noun is neither in the NGSL or the first and second BNC/COCA sub-lists, it does not have a specialised sense and was therefore categorised as TAM1 *least technical*.

The noun *fee* means “an amount of money paid for a particular piece of work or for a particular right or service” in both the Cambridge Advanced Learners Dictionary & Thesaurus (n.d.) and the Cambridge Business English Dictionary (n.d.) and so is placed in the *least technical* category. Another noun *buybacks* means “the act of buying something from the same person you sold it to, especially an offer by a company to buy shares of its own stock from shareholders” in the Cambridge Business English Dictionary (n.d.). The meaning in the Cambridge Advanced Learners Dictionary & Thesaurus (n.d.) is “an arrangement in which a business or person sells something, especially shares in companies, and then buys them again according to a fixed agreement”, which is very similar in substance to that in the Cambridge Business English Dictionary. The specialised sense is in no way different from or more detailed than the general sense and thus *buybacks* was categorised as *least technical*. In short, TAM1 *least technical* words can be understood without any specialised knowledge.

**b. Categorisation criteria for TAM 2 (slightly technical) words**

The nouns *capital* and *gearing* exemplify the TAM 2 (*slightly technical*) words (see Table 7).

Table 7 Exemplifications of TAM 2 (Slightly technical)

Financial sector specific-word	General sense	Specialised sense	Word frequency	Analysis result
capital (n.)	Yes	Yes	NGSL or 1st/2nd BNC/COCA sub-list	The specialised senses contain more details. → TAM 2
gearing (n.)	Yes (marked “SPECIALIZED”)	Yes	NGSL or 1st/2nd BNC/COCA sub-list	Same specialised senses → TAM 2

From the seven entries for the noun *capital* in the *Cambridge Advanced Learners Dictionary & Thesaurus*, the relevant sense is “money and possessions, especially a large amount of money used for producing more wealth or for starting a new business” whereas the *Cambridge Business English Dictionary* presents three specialised senses in the context of finance: (1) money that is used for investment or for starting a business, (2) the total amount of money and property that an individual or company owns, and (3) money that is lent or borrowed and will have to be paid back. This example clearly shows that the entry in the Business Dictionary contains more details than the related general sense in the Learners Dictionary and so *capital* was categorised as TAM2 *slightly technical*.

Another way that a word from the NGSL or the first and second BNC/COCA sub-lists can be considered *slightly technical* is when the general sense in the Cambridge Advanced Learners Dictionary & Thesaurus is marked *SPECIALIZED* so that a general sense can be regarded as a specialised sense. An example is *gearing* which refers to “the amount a company has borrowed compared to its share capital” in both the Cambridge Advanced Learners Dictionary & Thesaurus (n.d.) and the Cambridge Business English Dictionary (n.d.). However, the entry in the former is marked *SPECIALIZED*. As a result, this word in the second BNC/COCA sub-lists was deemed to have a truly specialised sense and be TAM2 *slightly technical*.

**c. Categorisation criteria for TAM 3 (moderately technical) words**

Four words were selected to exemplify the circumstances under which words are categorised as TAM 3 (*moderately technical*) (see Table 8).

Table 8 Exemplifications of TAM 3 (Moderately technical)

Financial sector specific-word	General sense	Specialised sense	Word frequency	Analysis result
impairment (n.)	Yes	Yes	3rd to 9th BNC/COCA sub-lists	The specialised senses contain more details. → TAM 3
clearing (n.)	Yes	Yes	NGSL or 1st/2nd BNC/COCA sub-list	The specialised sense is REMOTELY related to the general one. → TAM 3
exposure (n.)	Yes	Yes	NGSL or 1st/2nd BNC/COCA sub-list	The specialised sense is NOT related to the general one. → TAM 3
metrics (n.)	No	Yes	NGSL or 1st/2nd BNC/COCA sub-list	Monoseme → TAM 3

The noun *impairment* means “a situation in which the value of an asset is recorded as being greater than the amount of money that it could be sold for” in the Cambridge Business English Dictionary (n.d.). The Cambridge Advanced Learners Dictionary & Thesaurus (n.d.) has a general sense “the act of spoiling something or making it weaker so that it is less effective” and a marked specialised sense “medical deterioration in the functioning of a body part, organ, or system which can be temporary or permanent, and which can result from injury or disease”. The specialised sense in finance contains more details than the general one in the sense that the former describes a specific situation in the context of accounting whereas the latter is an abstract description that attempts to generalise different possible circumstances. In this regard, the specialised sense is more than remotely related to the general one and the word type, which is in the fourth BNC/COCA sub-list, was therefore categorised as TAM3 *moderately technical*.

Some words fall into several possible word classes each with a distinct sense and in such cases it is necessary to ascertain the word class of the majority of tokens in the corpus. The word *clearing* in the NGSL, for instance, can be a verb or a noun but a concordance of instances in the corpora found only 10 out of 429 tokens were verbs. The two specialised senses in the Cambridge Business English Dictionary (n.d.) are “the process by which cheques and other payments are exchanged between customers of different banks” and “the process by which shares and money are exchanged at the end of a day of trading on a financial market”. The only entry for the noun form was in the general dictionary and not related to the specialised sense. Examining the other nine senses in the verb form,

the sense “to (cause a cheque to) go from one bank to another through a central organization, so that money can be paid to the person it is owed to” (*Cambridge Advanced Learners Dictionary & Thesaurus*, n.d.) is closely related to the first sense in the Cambridge Business English Dictionary.

Deciding between these competing meanings involved scrutinising concordances of *clearing* in the corpora. The relevant specialised sense was determined to be the second definition as the right collocates include *house* (44 times), *organizations* (41 times), and *services* (24 times). We thus categorised *clearing* as *moderately technical* with the justification that the specialised sense is only remotely related to the general sense about cheques. To gain further support for the decision, a total of 300 concordances of *clearing* were randomly generated from *WordBanks* and analysed. The results showed that only 4% of the concordances, all found in finance, had the specialised sense, suggesting a narrow range of the word. This extra evidence supports the maxim that the more technical a word is, the fewer contexts we encounter it.

Another example of a *moderately technical* word is *exposure*, which is in the second BNC/COCA sub-list. It has two specialised senses in finance: “the risk of losing money, for example through a loan or investment, or the amount of money that might be lost” and “the act of investing in something” (*Cambridge Business English Dictionary* n.d.). None of the six senses in the Cambridge Advanced Learners Dictionary & Thesaurus (n.d.) is relevant to these two specialised senses.

Inspecting the concordances, all the tokens in the corpora have the specialised meanings and thus *exposure* was categorised as *moderately technical* for the specialised senses are not related to any of its general senses. One more instance is a monosemous plural noun *metrics* which means “a set of numbers that give information about a particular process or activity” (*Cambridge Business English Dictionary*, n.d.). It does not have an entry in the Cambridge Advanced Learners Dictionary and is in neither in the NGSL nor the first or second BNC/COCA sub-list and was therefore categorised as TAM3 *moderately technical*.

**d. Categorisation criteria for TAM 4 (very technical) words**

Moving to the degree of TAM4 *very technical*, we once again elaborate possible word categorisation scenarios (see Table 9).

Table 9 Exemplifications of TAM 4 (Very technical)

Financial sector specific-word	General sense	Specialised sense	Word frequency	Analysis result

collateralized (adj./v.)	No	Yes	3rd to 9th BNC/COCA sub- lists	The word has only specialised senses. → TAM 4
facilities (n.)	Yes	Yes	3rd to 9th BNC/COCA sub- lists	The specialised sense is NOT related to the general ones. → TAM 4
amortized (adj./v.)	Yes	Yes	10th to 25th BNC/COCA sub- lists or beyond	Polysemous → TAM 4
guaranty (n.)	No	Yes	10th to 25th BNC/COCA sub- lists or beyond	Monoseme but can be understood literally → TAM 4

*Collateralized* is a word in the third to the ninth BNC/COCA sub-lists. This monosemous word appears as either an adjective or a verb in the corpora and means “to give property as collateral for a loan, bond, etc.” (*Cambridge Business English Dictionary*, n.d.). The Cambridge Advanced Learners Dictionary does not have an entry for this word, occurring in the eighth BNC/COCA sub-list, and so *collateralized* was categorised as TAM4 *very technical*. Another similar instance is *facilities* from the third BNC/COCA sub-list. Its specialised sense is “an arrangement that lets someone borrow money from a bank or other financial institution for an agreed period of time or up to a particular amount” (*Cambridge Business English Dictionary*, n.d.) and this is the sense that many tokens in the corpora have. The specialised sense is not related to any of the two general senses about building and ability in the Cambridge Advanced Learners Dictionary and thus *facilities* was categorised as TAM4 *very technical*.

The word type *amortized* is an instance of a *very technical* word from the 10<sup>th</sup> to 25<sup>th</sup> BNC/COCA sub-lists and in the corpora it occurs as an adjective or verb. It has a specialised sense in the context of accounting which is “to spread the value or cost of an asset in accounts over a number of years” and another sense “to reduce a debt by paying small regular amounts” (*Cambridge Business English Dictionary*, n.d.) but this sense in exactly the same wordings also appears in the Cambridge Advanced Learners Dictionary. The judgement to be made for word types at or beyond the 10<sup>th</sup> BNC/COCA sub-list is to determine if it is polysemous. The specialised sense and the general sense of the word type *amortized* are different and shall be deemed polysemous. Hence, *amortized* was categorised as TAM4 *very technical*.

One final example in the *very technical* category is *guaranty*. This ostensibly looks like a member of the family *guarantee* in the second BNC/COCA sub-list but it is not in any reference word list and is a monoseme which means “a legal agreement in which a person or organization promises to pay back a loan if the person or organization that originally borrowed the money cannot” (*Cambridge Business English Dictionary*, n.d.). Although the word *guaranty* is monosemous, given its grammatical properties, it will probably be seen as a variant of the noun *guarantee* and can be understood literally. The meaning of a *most technical* word cannot be understood literally or guessed easily so *guaranty* does not qualify as a *most technical* word and is only TAM4 *very technical*.

**e. Categorisation criteria for TAM 5 (most technical) words**

The last category on the cline of technicality is *most technical*. A *most technical* word is characterised by the fact that its only sense is specialised and cannot be understood literally or guessed easily. A total of nine word types fall into this category in finance as shown in Table 10.

Table 10 List of most technical words

	Word	No. of tokens	BNC/COCA banding
1	ACCRETABLE	128	#N/A
2	ACCRETIVE	124	#N/A
3	BANCASSURANCE	112	#N/A
4	ESCROW	146	BNC/COCA 14k
5	LIEN	352	BNC/COCA 10k
6	LIENS	132	BNC/COCA 10k
7	MORTGAGE-BACKED	435	#N/A
8	SUBPRIME	177	#N/A
9	TRIPLE-NET	174	#N/A

Examining the most frequent items, we find that *accretable* cannot be found in reference dictionaries while *accretive* has an entry in the Cambridge Business English Dictionary (n.d.) as “making the amount, level, or value of something gradually increase”. Studying the concordances, we might expect the specialised sense of *accretable* to be very similar to that of *accretive*, but there are considerable differences in usage. The majority of tokens for *accretable* modifies the collocate *yield* and 104 out of its 128 tokens occur in writing while *accretive* does not have a dominant collocate and 116 out of its 124 tokens occur in speech.

The third word in Table 10, *bancassurance*, means “the combination of banking and insurance services that is offered by many banks” (*Cambridge Business English Dictionary*, n.d.) is a blend of *bank* and *insurance* in lexicalisation. However, as a word borrowed from French, *bancassurance*



cannot be considered literally understood and so this monosemous off-list word was categorised as *most technical*. The monosemes *escrow*, *lien*, and *liens* are undoubtedly *most technical* words. The adjective *subprime* is “used to describe a mortgage that has a high risk of not being paid back” (*Cambridge Business English Dictionary*, n.d.). The neoclassic prefix *sub-* means below (*Cambridge Advanced Learners Dictionary & Thesaurus*, n.d.) and the base word *prime* can be seen as exocentric in the sense that *prime* means the prime rate in the mortgage business (*Cambridge Business English Dictionary*, n.d.). In this regard, *subprime* can possibly be understood as related to a mortgage that yield a below-than-expected return, which is different from the real specialised sense. Thus the adjective *subprime* is monosemous and cannot be understood literally which means it is classified as *most technical*.

Concerning the two hyphenated words in Table 10, *mortgage-backed* and *triple-net*, the definition for the former can be found in the Cambridge Business English Dictionary but the latter we need to resort to the *Investopedia Dictionary*. The word *mortgage-backed* is “used to describe an investment, especially a bond, in which the money that is used to pay back mortgages is used to pay interest on the investment” (*Cambridge Business English Dictionary*, n.d.). The adjective *triple-net* is usually used to modify the noun *lease* in the corpora. A *triple-net lease* refers to “[a] lease agreement that designates the lessee (the tenant) as being solely responsible for all of the costs relating to the asset being leased in addition to the rent fee applied under the lease” (*Investopedia Dictionary*, n.d.). Dissecting the specialised senses of the two hyphenated words, guessing the sense for *triple-net* should be impossible while the word *mortgage-backed* does not make sense to anyone without subject knowledge as it seems illogical that a mortgage, which is debt, can be used to “back” something. Even the description of the specialised sense in the Cambridge Business English Dictionary may hardly be understood. In the light of these considerations, the monosemes *mortgage-backed* and *triple-net* were categorised as TAM5 *most technical*.

#### **f. Categorisation criteria for hyphenated technical words**

The study does not separate words with hyphens which means that hyphenated words, for example, *asset-backed*, *long-term*, and *pre-tax* are counted as three words. The NGSL has five hyphenated words and the AWL (Academic Word List) 25. The BNC/COCA sub-lists do not have any, although there is a list of transparent compounds without hyphens. In the hope of exploring more facets of financial vocabulary, no criteria exclude hyphenated words in the study. Thus the TAM can be applied to hyphenated words with similar rationales.

Let’s take *pre-tax* as an example. We found its unhyphenated variant *pretax* in the first BNC/COCA sub-lists and so we classified it as *least technical*. Unlike *pre-tax*, the word *after-tax* is not in any reference word lists and the specialised sense “used to describe an amount of money that is left after

tax has been taken away” (*Cambridge Business English Dictionary*, n.d.) is simply a wordy version of *after tax*, which are both in the NGSL. The specialised sense of *after-tax* is therefore deemed the same as the general sense and *after-tax* is classified as *least technical*. Other hyphenated words such as *available-for-sale*, *cross-border*, and *high-quality* were categorised as *least technical* on the same grounds. The adjectival compound *credit-related* was considered *slightly technical* because the non-head *credit* is *slightly technical* and the adjectivally-used past participle *related* is in the NGSL, making a compound of which the meaning can be literally understood and does not alter the degree of technicality. In cases of having different degrees of technicality among the constituents in hyphenated words, the highest degree prevails. As a result, the adjectival compound *credit-related*, which does not stand as a separate entry in the *Cambridge Business English Dictionary*, was considered *slightly technical*.

If a hyphenated word has an entry in the *Cambridge Business English Dictionary* and the specialised sense contains more details than or is different from the literal meaning of the combined senses of the constituents in the hyphenated word, the hyphenated word was deemed at least *moderately technical*. The adjective *fixed-income*, for example, is “used to describe investments such as bonds that pay the same amount of money every month, year, etc.” (*Cambridge Business English Dictionary*, n.d.) and the specialised sense contains more details than the literal meaning of *fixed income* which usually means the compensation from work rather than return on an investment. The word type *fixed-income* was therefore categorised as *moderately technical*.

Other hyphenated word types such as *asset-backed*, *floating-rate*, and *risk-based* are examples of *moderately technical* words. The noun *market-making* refers to “the continuous buying and selling of shares in particular companies at particular prices” (*Cambridge Business English Dictionary*, n.d.) and the specialised sense is not closely connected to the literal meaning of *the making of a market* but successful guessing of the meaning is still possible with subject knowledge. In this regard, the word type *market-making* is remotely related to its literal meaning and subject knowledge is needed for guessing the meaning and *market-making* was therefore categorised as *very technical*. Likewise, another noun *pass-through* refers to “an arrangement in which a financial organization buys loans from a bank and sells bonds representing these loans to investors. The payments on the loans are then used to pay interest to the investors and pay back the bonds” (*Cambridge Business English Dictionary*, n.d.) and was also considered *very technical*.

The word *triple-net* is the only hyphenated word among the *most technical* words. As discussed earlier in this section, the adjective *triple-net* is usually used to modify the noun *lease* and basically refers to leases in which the lessee bears all the costs. The categorisation criteria can be readily

applied to the technicality assessment of *triple-net*, which is a monosemous word off the BNC/COCA sub-list with a specialised sense that cannot be understood or guessed literally.

## 6. Technicality in the Financial Corpus

Having elaborating the details of the TAM, we now present the results of applying this to the Financial Corpus. The TAM identified 837 technical words in the distributions shown in Table 11. Our findings show that a word specific to a financial sector is not necessarily technical, indicating that technicality and specificity are distinct concepts. The concordance analyses of financial-sector-specific words help to depict how the words reflect the business nature of the respective financial sector and illustrate their financial-sector-specificity. Polysemy was found to be pervasive in the financial-sector-specific vocabulary and 109 out of the 624 polysemes identified are technical to various degrees.

Table 11 Results of technicality analysis by degree of technicality

Degree of technicality	No. of word types	%	No. of tokens	%
TAM1 Least Technical	672	80.29%	1403782	20.79%
TAM2 Slightly Technical	42	5.02%	106633	1.58%
TAM3 Moderately Technical	88	10.51%	91382	1.35%
TAM4 Very Technical	26	3.11%	13644	0.20%
TAM5 Most Technical	9	1.08%	1780	0.03%
Grand Total	837	100.00%	1617221	23.95%

The most frequent word in each category (see Appendix A), by ascending technicality, are *so* (24,147 times), *assets* (17,497 times), *capital* (17,837 times), *equity* (10,174 times), *facilities* (1,358 times), and *mortgage-backed* (435 times).

To provide more detailed information for EAP teachers, we further categorised the 672 TAM 1 *least technical* words into *least technical – wide range* and *least technical – narrow range*, according to their distributions in the NGSL and the first and second BNC/COCA sub-lists. Looking at the 20 most frequent *least technical* words, finance-related words include *asset*, *assets*, *business*, *company*, *deposits*, *dividend*, *financial*, *interest*, *loans*, *management*, *risk*, *shareholders*, *value*, and so forth. Some words like *new*, *see*, and *so* do not have any specialised senses while corpus evidence reveals that other words such as *deferred* and *fair* actually form parts of technical multi-word units.

Regarding *slightly technical words*, each word has a general sense that is related to the specialised sense in finance but the latter sense contains more details. Thus the following examples of *capital* show how one of its specialised senses overlaps with a general sense.

- (1) ...we don't want to keep too much *capital* and if we cannot put the *capital* into good use...
- (2) We have, in our judgment, and will continue to have, access to the *capital* markets if we need more capital for major acquisitions.

When occurring in compounds in Examples 3 to 5, capital does not simply mean the money invested but functions to clarify the types of *projects*, *gains*, and *expenditures* referred to. The compound *capital projects* means that the company intended to make an investment that would lead to the company's ownership of the *564 sites* upon completion of the construction, exhibiting the first specialised sense as well. By the same token, *capital gains* means that the gains are derived from what the company owns whilst *capital expenditures* refers to the expenditures that are spent on what the company owns and on which value will be increased upon the completion of any work connected with the expenditures. Nevertheless, the specialised senses of *capital* can be regarded as simply contextualised senses having more details than the general sense, justifying *capital* to be *slightly technical*.

- (3) We spent about \$49 million on discretionary *capital* projects associated with the completion of the construction of 564 sites globally.
- (4) Exchange rate differences relating to the disposal of available-for-sale debt and equity securities are considered to be an inherent part of the *capital* gains and losses...
- (5) ...we may fund the *capital* expenditures for our triple-net leased properties through loans to the tenants or advances...

As for *moderately technical* vocabulary, every word has a specialised sense but may not have a general sense in the Cambridge Advanced Learners Dictionary & Thesaurus. The non-NGSL polysemous compound *goodwill*, for instance, has two general senses in the Cambridge Advanced Learners Dictionary & Thesaurus (n.d.) of which "part of a company's value that includes things that cannot be directly measured, for example, its good reputation or its customers' loyalty" is related to the specialised sense "the difference between the value of a company's assets and what profit it is expected to make in the future, which is included in the price paid when it is bought or sold" (*Cambridge Business English Dictionary*, n.d.). The latter appears to be a more detailed description which reflects the sense of *goodwill* in the more general dictionary, although it cannot be derived from the combined literal meaning of *good* and *will*, resulting in the categorisation into *moderately technical* instead of a higher degree of technicality.

Concerning *very technical words*, *facility* and *facilities* were categorised as *very technical* because they have a specialised sense in finance as "an arrangement that lets someone borrow money from a

bank or other financial institution for an agreed period of time or up to a particular amount” (*Cambridge Business English Dictionary*, n.d.). This specialised sense is not related to any of its general senses as can be seen in examples 6 and 7.

- (6) The liquidity risk associated with the potential drawdown on non-cancellable committed *facilities* is factored into our stressed scenarios and limits are set for these *facilities*.
- (7) Based on our current credit ratings, the amended *facility* bears interest annually at one-month LIBOR plus 1.075% and has a *facility* fee of 0.175%...

Regarding the *most technical vocabulary*, all the nine words are not in the NGSL or the AWL and are in the tenth BNC/COCA sub-list or beyond. A *most technical* word has only one specialised sense and cannot be understood literally.

## 7. Limitations

Investigating technicality is a fraught process and, like the previous studies discussed above, the TAM has its limitations. One is the selection of appropriate dictionaries for making judgements and EAP/ESP teachers need access to appropriate specialised dictionaries for their target discipline as well as good general dictionaries. A second important limitation is that the TAM requires users to make comparisons between general senses and specialised senses and between literal senses and specialised senses and this may be perceived as introducing subjectivity into judgements of technicality. We believe human decision-making is an inescapable part of any linguistic analysis – even those decisions made by disciplinary experts, but that the TAM minimises the personal judgment required with the aid of dictionaries and the design of the Technicality Analysis Model itself. In addition to the use of expert opinion in the form of specialist dictionaries, the model integrates frequency benchmarking of an item against established word lists and checking the number of entries of senses in authoritative dictionaries. This may suggest that our model is somewhat over-elaborate and complex, but we believe there is little prospect of further simplification without detriment to its analytical strength. The categorisation we propose provides explicit criteria which minimises the amount of judgement needed by teachers or analysts.

## 8. Conclusions

The Technical Analysis Model enables technicality to be systematically categorised and understood along a continuum. Unlike the methods employed in previous studies, our method considers not only the specialised sense of a word but also its general sense in order to assess the word’s technicality. In this study we have attempted to demonstrate the feasibility of categorising words along a continuum at five intervals of technicality in a principled and methodical way.

One point worth mentioning is the relationship between *technical* and *non-technical* words. The *least technical* vocabulary has minimal technicality and readers may be tempted to see them as *non-technical*. Nevertheless, caution must be exercised in any attempts to create a dichotomy between *technical* and *non-technical* words. At the initial stage of the analysis, only function words rather than all high-frequency words were excluded, assuming that function words do not have any specialised senses. Second thoughts need to be given to this assumption when considering function words such as *per* in the technical Multi-Word Unit *earnings per share*. The preposition *per* cannot be replaced with any other words for the word's conventional abbreviation (i.e. *EPS*) (*Cambridge Business English Dictionary*, n.d.) includes the initial letter P. In this regard, it is reasonable and prudent to categorise the preposition *per* as *least technical* instead of *non-technical*. This case demonstrates that even function words can have minimal degree of technicality. We argue that any sense in a word implies a certain degree of technicality. The technicality of content words should therefore be assessed along a cline and it would not be sensible to draw a line between *least technical* and *non-technical*.

We should also note the role of Multi Word Units (MWUs), where words take on technical meanings by virtue of their combination in particular domains, such as *mutual funds* and *capital expenditure*. This is already a long paper and we have no space to elaborate our treatment of these, particularly the semantically opaque *synergistic technical* MWUs where the combination of the literal senses of their constituents differs from the simple combination of individual terms (*common stock or carrying value*). More straightforward are Visible Technical MWUs, which have at least one technical word as a constituent, allowing the meaning to be identified without much difficulty. This is done by looking into the technicality of each constituent on a single word basis and categorising according to the highest degree of technicality among constituent. This produces 57 *slightly technical* MWUs, 55 *moderately technical* MWUs, and six *very technical* MWUs but no *most technical* MWUs (see appendix B).

Finally, we would like to suggest that the TAM might be used in future research to assess words in different disciplines. Disciplines diverge in terms of knowledge and discourse. The “discipline-based lexical repertoire” (Hyland & Tse, 2007, p. 235) is the direction to which the vocabulary research for ESP/EAP should be heading and to achieve the goal, more extensive discipline-based lexical research may be carried out in future. Such research could use exactly the same categorisation criteria and procedures, not only testing the robustness of the TAM itself, but building towards a repertoire of discipline-specific vocabulary lists for EAP teaching and learning. This will assist classroom teaching and allow the comparison of vocabulary technicality across genres and fields. The word lists may also be used to test students' vocabulary knowledge in particular fields (Nation, 2016). In addition, this

exploration of technical vocabulary need not be confined to single words as the TAM is sufficiently flexible to study multi-word units, further enriching the understanding of technical vocabulary.

## References

- Hyland, K. (1997). Language attitudes at the handover: communication and identity in 1997 Hong Kong. *English World Wide*, 18 (2): 191-210.
- Hyland, K. (2002). Specificity revisited: how far should we go now? *English for Specific Purposes*, 21 (4): 385-395
- Hyland, K. & Tse, P. (2007). Is there an 'academic Vocabulary'? *TESOL Quarterly*, 41 (2): 235-254
- Browne, C., Culligan, B., & Phillips, J. (2013). *A new general service list*. Retrieved from <http://www.newgeneralservicelist.org/>
- Cambridge advanced learners dictionary & thesaurus. Retrieved from <http://dictionary.cambridge.org/dictionary/english/>
- Cambridge business English dictionary. Retrieved from <http://dictionary.cambridge.org/dictionary/business-english/>
- Chung, T. M., & Nation, P. (2003). Technical vocabulary in specialised texts. *Reading in a Foreign Language*, 15(5), 103-116.
- Chung, T. M., & Nation, P. (2004). Identifying technical vocabulary. *System*, 32(2), 251-263.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238.
- Dow Jones Indexes. (2012). *Dow Jones Sector Titans Index: Selection List*. Retrieved November 20, 2012, from [http://www.djindexes.com/mdsidx/downloads/selectionlists/DJSectorTitans30\\_Selection.xls](http://www.djindexes.com/mdsidx/downloads/selectionlists/DJSectorTitans30_Selection.xls)
- Dudley-Evans, A., & St. John, A. M. (1998). *Developments in ESP: A multi-disciplinary approach*. Cambridge: Cambridge University Press.
- Durrant, P. (2014). Discipline and level specificity in university students' written vocabulary. *Applied Linguistics*, 35(3), 328-356.
- Evans, S., & Green, C. (2007). Why EAP is necessary: A survey of Hong Kong tertiary students. *Journal of English for Academic Purposes*, 6, 3-17.
- Evans, S., & Morrison, B. (2012). Learning and using English at university: Lessons from a longitudinal study in Hong Kong. *The Journal of Asia TEFL*, 9(2), 21-47.
- Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, 35(3), 305-327.
- Hsu, W. (2011). The vocabulary thresholds of business textbooks and business research articles for EFL learners. *English for Specific Purposes*, 30(4), 247-257.
- Investopedia dictionary. Retrieved from <http://www.investopedia.com/dictionary/>
- Kwary, D. A. (2011). A hybrid method for determining technical vocabulary. *System*, 39(2), 175-185.

- Martin, A. V. (1976). Teaching academic vocabulary to foreign graduate students. *TESOL Quarterly*, 10(1), 91-97.
- Moon, R. (2010). What can a corpus tell us about lexis? In A. O’Keeffe and M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 197-211). London: Routledge.
- Mudraya, O. (2006). Engineering English: A lexical frequency instructional model. *English for Specific Purposes*, 25(2), 235-256.
- Mukundan, J., & Jin, N. Y. (2012). Development of a technical nursing education word list (NEWL). *International Journal of Innovation in English Language Teaching and Research*, 1(2), 105-124.
- Nation, I. S. P. (2012). The BNC/COCA word family lists. Retrieved from [http://www.victoria.ac.nz/lals/about/staff/publications/paul-nation/Information-on-the-BNC\\_COCA-word-family-lists.pdf](http://www.victoria.ac.nz/lals/about/staff/publications/paul-nation/Information-on-the-BNC_COCA-word-family-lists.pdf)
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge: Cambridge University Press.
- Nation, I. S. P. (2016). *Making and using word lists for language learning and testing*. Amsterdam: John Benjamins.
- Nation, I. S. P., & Newton, J. (1997). Teaching vocabulary. In J. Coady and T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 238-254). Cambridge: Cambridge University Press.
- Paquot, M. (2010). *Academic vocabulary in learner writing: From extraction to analysis*. London: Continuum.
- Pearson, J. (1998). *Terms in context*. Amsterdam: Benjamins.
- Schmitt, N. (2010). *Research vocabulary: A vocabulary research manual*. Basingstoke: Palgrave Macmillan.
- Spack, R. (1988). Initiating ESL students into the academic discourse community: How far should we go? *TESOL Quarterly*, 22(1), 29-52.
- Sutarsyah, C., Nation, I. S. P., & Kennedy, G. (1994). How useful is EAP vocabulary for ESP? A corpus based case study. *RELC Journal*, 25(2), 34-50.
- Valipouri, L., & Nassaji, H. (2013). A corpus-based study of academic vocabulary in chemistry research articles. *Journal of English for Academic Purposes*, 12, 248-263.
- Ward, J. (2009). A basic engineering English word list for less proficient foundation engineering undergraduates. *English for Specific Purposes*, 28(3), 170-182.
- West, M. (1953). *A general service list of English words*. London: Longman, Green, & Company.
- Woodward-Kron, R. (2008). More than just jargon – the nature and the role of specialist language in learning disciplinary knowledge. *Journal of English for Specific Purposes*, 7(4), 234-249.
- WordBanks Online. Retrieved from <http://wordbanks.harpercollins.co.uk>
- Wu, W., & Hammond, M. (2011). Challenges of university adjustment in the UK: A study of East Asian master’s degree students. *Journal of Further and Higher Education*, 35(3), 423-438.



## Appendix A: Most Frequent Technical Words by Category

### Most frequent TAM 1 words in financial-sector corpora

Least technical - wide range				Least technical - narrow range				
	Word type	Tokens	Word type	Tokens	Word type	Tokens	Word type	Tokens
1	SO	24147	GROWTH	11506	ASSETS	17497	REVENUES	4178
2	BUSINESS	22181	RATE	11016	INVESTMENT	12309	ACQUISITION	3783
3	YEAR	22139	NEW	10684	PORTFOLIO	8229	TRANSACTIONS	3766
4	QUARTER	20957	SEE	10369	IMPACT	7253	REGULATORY	3659
5	FINANCIAL	18225	FAIR	10252	ASSET	6386	SHAREHOLDERS	3654
6	NET	17232	CASH	10065	REVENUE	6283	RETAIL	3478
7	VALUE	17200	YEARS	9626	CONSOLIDATED	6142	FEES	3419
8	INCOME	16927	BASIS	8655	INVESTMENTS	6051	ANNUAL	3400
9	RISK	16054	OPERATING	8189	CORPORATE	5106	OVERALL	3179
10	COMPANY	14505	BANK	8054	SIGNIFICANT	4983	CLIENTS	2981
11	MANAGEMENT	14031	SHARE	7914	RATIO	4651	DEPOSITS	2962
12	INTEREST	13911	RESULTS	7799	APPROXIMATELY	4318	CORPORATION	2938
13	JUST	13512	INSURANCE	7726	GLOBAL	4297	OFFSET	2884
14	TOTAL	12947	LOSS	7542	EXECUTIVE	4244	CONSUMER	2858
15	LOANS	11855	TAX	7501	PRIMARILY	4199	SEGMENT	2745

### Most frequent TAM2 – TAM 5 words in financial-sector corpora

	Slightly technical		Moderately technical		Very technical		Most technical	
	Word type	Tokens	Word type	Tokens	Word type	Tokens	Word type	Tokens
1	CAPITAL	17837	EQUITY	10174	FACILITIES	1358	MORTGAGE-BACKED	435
2	CREDIT	13919	LIABILITIES	6458	AMORTIZATION	1161	LIEN	352
3	SECURITIES	12308	MORTGAGE	4616	FACILITY	1058	SUBPRIME	177
4	DEBT	7275	LIQUIDITY	3517	COUNTERPARTY	961	TRIPLE-NET	174
5	OPERATIONS	6655	EXPOSURE	3259	SWAPS	915	ESCROW	146
6	BALANCE	5873	MARGIN	2929	SECURITIZATION	767	LIENS	132
7	BOARD	5347	IMPAIRMENT	2833	CHARGE-OFFS	712	ACCRETABLE	128
8	EXCHANGE	3571	DERIVATIVES	2729	COUNTERPARTIES	704	ACCRETIVE	124
9	RETURN	3038	UNDERLYING	2620	THIRD-PARTY	657	BANCASSURANCE	112
10	UNITS	2017	DERIVATIVE	2590	POLICYHOLDER	601		
11	RETURNS	1949	LIABILITY	2194	AMORTIZED	582		
12	ALLOWANCE	1945	REPURCHASE	1959	POLICYHOLDERS	573		
13	RESERVES	1835	OPTION	1835	SECURITIZATIONS	516		
14	HEDGE	1761	ATTRIBUTABLE	1714	FORECLOSURE	484		
15	TRUST	1758	GOODWILL	1526	COLLATERALIZED	411		
16	RECOGNIZED	1733	EXPOSURES	1489	UNIT-LINKED	349		
17	BALANCES	1667	MORTGAGES	1472	REINVESTMENT	309		
18	SETTLEMENT	1579	ADJUSTMENTS	1443	CHARGE-OFF	285		
19	BORROWINGS	1459	LEVERAGE	1423	DENOMINATED	250		
20	BONDS	1385	DEFAULT	1361	PROVISIONING	189		
21	HEDGING	1246	UNDERWRITING	1235	SECURITIZED	177		
22	SPREADS	1235	REINSURANCE	1194	GUARANTY	141		
23	SPREAD	1150	HEDGES	1168	PASS-THROUGH	140		
24	SECURED	1061	ADJUSTMENT	1163	MARKET-MAKING	129		
25	BOND	951	MARGINS	1151	DE-RISKING	122		
26	DEALS	771	FISCAL	1134	FORECLOSURES	93		
27	TRUSTS	747	DISCOUNT	1118				
28	OFFERING	721	UNREALIZED	1086				
29	HEDGED	494	RESTRUCTURING	1034				
30	OFFERINGS	402	INTANGIBLE	1023				

## Appendix B: Visibly Technical Multi Word Units

	Slightly technical MWUs	Specific sector(s)	Tokens in the FC		Slightly technical MWUs	Specific sector(s)	Tokens in the FC
1	CAPITAL MARKETS	4	1902	30	CREDIT CARDS	1	127
2	CREDIT RISK	3	1644	31	CAPITAL GENERATION	2	123
3	FOREIGN EXCHANGE	3	1232	32	MANAGEMENT BOARD	1	121
4	CREDIT CARD	2	1167	33	EXCHANGE RATE	1	114
5	CREDIT LOSSES	1	840	34	CAPITAL BASE	2	110
6	CAPITAL REQUIREMENTS	3	623	35	CAPITAL STRUCTURE	2	102
7	CREDIT QUALITY	2	512	36	EXCESS CAPITAL	2	102
8	CREDIT SPREADS	3	421	37	CREDIT SPREAD	1	81
9	REGULATORY CAPITAL	2	417	38	CREDITING RATES	1	80
10	CAPITAL MANAGEMENT	3	389	39	SECURED DEBT	1	73
11	LONG-TERM DEBT	1	382	40	SPREAD COMPRESSION	1	70
12	CAPITAL RATIO	3	377	41	NET DEBT	1	70
13	CAPITAL POSITION	3	372	42	CREDIT COSTS	1	69
14	CAPITAL RATIOS	3	370	43	HEDGING PROGRAM	1	64
15	INSURANCE OPERATIONS	1	343	44	INVESTMENT SECURITIES	1	63
16	INVESTMENT RETURNS	1	293	45	SECURITIES PORTFOLIO	1	61
17	CREDIT PORTFOLIO	1	211	46	CAPITAL TRANSACTIONS	1	52
18	SECURITIES LENDING	1	203	47	CREDIT TRENDS	1	49
19	CAPITAL PLAN	2	170	48	HEDGE FUND	1	48
20	SOVEREIGN DEBT	2	159	49	INVESTED CAPITAL	1	47
21	REQUIRED CAPITAL	1	156	50	CREDIT CARD PORTFOLIO	1	44
22	LOAN BALANCES	1	155	51	SECURITIES GAINS	1	44
23	STRONG CAPITAL	3	153	52	CREDIT MARKET	1	42
24	EXCHANGE RATES	1	153	53	CREDIT COST	1	42
25	GOVERNMENT BONDS	1	153	54	CAPITAL RULES	1	41
26	CREDIT PERFORMANCE	2	145	55	GOVERNMENT BOND	1	41
27	REVOLVING CREDIT	1	136	56	STATUTORY CAPITAL	1	39
28	CAPITAL LEVELS	3	132	57	CREDIT RATING	1	34
29	CAPITAL ALLOCATION	2	128				

	Moderately technical MWUs	Specific sector(s)	Tokens in the FC		Moderately technical MWUs	Specific sector(s)	Tokens in the FC
1	PRIVATE EQUITY	4	800	29	SHARE REPURCHASES	1	84
2	RISK-WEIGHTED ASSETS	2	495	30	RISK WEIGHTED	1	77
3	MORTGAGE LOANS	1	391	31	LEVERAGE RATIO	1	76
4	EQUITY MARKETS	3	382	32	INVESTMENT MARGIN	1	75
5	RESIDENTIAL MORTGAGE	1	360	33	UNDERWRITING RESULTS	1	72
6	SHAREHOLDERS EQUITY	1	329	34	IMPAIRMENT CHARGE	1	70
7	INTEREST MARGIN	2	313	35	NET EXPOSURE	1	64
8	OPERATING LEVERAGE	2	307	36	RISK-BASED CAPITAL	1	61
9	IMPAIRMENT CHARGES	1	261	37	UNDERLYING LOSS	1	60
10	FISCAL YEAR	1	257	38	RISK WEIGHTED ASSETS	1	59
11	MORTGAGE SERVICING	1	214	39	EQUITY RATIO	1	55
12	IMPAIRED LOANS	1	207	40	TANGIBLE COMMON EQUITY	1	52
13	MORTGAGE BANKING	1	186	41	COMMON EQUITY RATIO	1	51
14	EQUITY MARKET	2	179	42	SOLVENCY MARGIN	1	50
15	OPERATING MARGIN	1	169	43	LIABILITY MANAGEMENT	1	49
16	SHARE REPURCHASE	2	154	44	RISK-WEIGHTED ASSET	1	48
17	POSITIVE OPERATING LEVERAGE	1	147	45	DEBIT CARDS	1	45
18	RESIDENTIAL MORTGAGES	1	141	46	EQUITY TIER	1	44
19	UNREALIZED GAINS	1	135	47	DILUTED SHARE	1	44
20	CREDIT EXPOSURE	1	132	48	IMPAIRED LOAN	1	43
21	DISCOUNT RATE	1	123	49	UNDERWRITING STANDARDS	1	41
22	MORTGAGE PORTFOLIO	1	119	50	TANGIBLE BOOK VALUE	1	41
23	DILUTED EARNINGS	1	118	51	CREDIT METRICS	1	38
24	EQUITY FUNDS	1	113	52	MORTGAGE ORIGINATION	1	36
25	EQUITY INVESTMENTS	1	107	53	MORTGAGE BOOK	1	36
26	MORTGAGE BUSINESS	1	104	54	LIQUIDITY POSITION	1	35
27	MARGIN EXPANSION	2	96	55	MARGIN PRESSURE	1	34
28	GLOBAL EQUITY	1	95				

	Very technical MWUs	Specific sector(s)	Tokens in the FC
1	CREDIT FACILITY	1	313
2	NET CHARGE-OFFS	1	268
3	REVOLVING CREDIT FACILITY	1	106
4	CREDIT FACILITIES	1	83
5	POLICYHOLDER BEHAVIOR	1	54
6	REINVESTMENT RATE	1	41

