The Lexis of Maths Lectures: The Creation of a Pedagogic Corpus and Wordlist from a Series of Maths Lectures

Alister Drury

Rachel Perkins

Warren Sheard

Language Centre, School of Languages, Cultures and Societies, University of Leeds

ABSTRACT

Researchers have investigated discipline-specific and academic vocabulary in multiple academic disciplines through the creation of word lists (Dang, 2018; Gilmore and Millar, 2018; Watson Todd, 2017; Valipouri and Nassajii, 2013). However, there are currently no mathematics wordlists based on spoken corpora that are suitable for the context in question. The present study involved the creation of a context-specific corpus to investigate the frequency of technical and sub-technical vocabulary in a series of mathematics lectures through the creation of a keyword list. The keyword list was created using Sketch Engine from a 152,443-word corpus of 46 mathematics lectures. The final wordlist comprised 202 lemmas, covering 12.89% of the corpus. The benefit of creating a context-specific wordlist was clear. The New Academic Word List (NAWL) (Browne et al., 2013a) provided just 4.51% coverage of the corpus. Assuming students have knowledge of the first 2000 word families of the New General Service List (NGSL) (Browne et al., 2013b), which provided 84% coverage, total coverage was nearly 97% with the mathematics wordlist, compared to 88.67% with the general academic list. Of the 202 lemmas in the mathematics wordlist, 116 were sub-technical, meaning that they are polysemous, with both a general meaning and a mathematical meaning. This is a feature of the language of mathematics and may be an added challenge for students.

KEYWORDS: corpora, wordlists, technical vocabulary, maths, EMI, polysemous, sub-technical

INTRODUCTION

A growing number of students whose first language is not English are choosing to study on higher education courses in their own country which are taught in part or exclusively in English. These degree programmes are often provided by universities from English speaking countries that have established branch campuses overseas. The number of these transnational education providers (TNEs) has grown rapidly in recent years, with China being a major centre for this expansion (Fang and Wang, 2014).

Students studying on undergraduate degree programmes in these contexts often report difficulties in understanding their subject lectures due to a lack of discipline specific terminology (Soruç and Griffiths, 2018, p.42). In recent years researchers within the field of English for Specific Purposes have sought to address this challenge by producing word lists of discipline-specific corpora in order to better understand variations in lexis between different disciplines and to aid the development of English language teaching materials and curriculum design. Bondi (2010, p.3) defines keywords as 'those whose frequency (or infrequency) in a text or corpus is statistically significant, when compared to the standards set by a reference corpus'. While lists have been produced for a number of disciplines such as engineering, medicine and agriculture (Coxhead, 2018), the influence these lists have on curriculum design remains unclear (Nation, 2016, p.172). One reason for this may be that these keyword lists are generally derived from relatively large corpora which aim to represent a whole field of disciplinary study and hence do not adequately address the specific needs of curriculum designers. For example, a list of engineering vocabulary may not be of equal utility to electrical engineering as civil engineering.

The problem of disciplinary lexis can be exacerbated in the case of foundation year programmes in which English language instruction is often separated from content instruction delivered through an English Medium Instruction (EMI) model, meaning English teachers are often left unaware of the language challenges facing their learners in their content modules (Galloway and Ruegg, 2020, pp.5-6). Opportunities to provide language support to these learners who struggle to adapt to the high vocabulary load within the content-based courses can, therefore, be missed.

This study aims to bridge this gap through the creation of a key word list derived from a small corpus (152,443 tokens) of mathematics lectures in order to better understand the nature of the lexical challenge facing the students and to aid in the development of pedagogical materials and interventions in order to aid students' comprehension of these lectures. The corpus and keyword list were based on a series of recorded lectures delivered by an English-speaking lecturer to Chinese students studying at a TNE provider in China.

BACKGROUND

The context of the study

SWJTU-Leeds Joint School is a TNE partnership between the University of Leeds and South-West Jiaotong University based in Chengdu, China. Students study on undergraduate degree programmes in four subjects within the Faculty of Engineering and Physical Sciences (civil engineering with transport, mechanical engineering, computer science and electronic and electrical engineering). At the end of their studies students receive an undergraduate degree from both the University of Leeds and South-West Jiaotong University. Courses last for a period of four years with the first year being common to all strands before students specialise in year 2. During the first year, students study an English for Engineering module taught jointly by English teachers from the University of Leeds and SWJTU as well as content-based modules in physics and mathematics. The content-based modules are taught using English as a medium of instruction (EMI).

The present study focuses on the lectures delivered by a mathematics lecturer from the University of Leeds whose first language is English. Before 2020, this module was taught in person; however, the pandemic has resulted in the replacement of live lectures and classes with recorded lectures through which the content is exclusively delivered. Two of the stated course objectives/learning outcomes of the module are for students to both understand the language of Mathematics and to be able to use this language.

RATIONALE

Lexical Challenge

EMI has seen a rapid expansion in China and indeed throughout the globe in recent years (Dearden, 2015, p.2) and the number of students in transnational education (TNE) is now 1.4 times the number of international students studying in the UK (Universities UK, 2020, p.20). The rationale for EMI/TNE is that through studying a subject in English, students will make gains in both their content knowledge and their language proficiency. However, little research has been done to validate this claim in a HE environment (Macaro et al, 2018, p.57) and the question remains as to what extent studying in a second language impedes content knowledge acquisition (Chang, 2010; Shohamy, 2010, in Doiz et al., 2012). Studies investigating student perceptions of EMI courses have suggested that lack of vocabulary knowledge may be a significant impediment to the comprehension of subject course materials and lectures (Evans and Green, 2007; Chang, 2010; Tatzl, 2011, cited in Harada and Uchihara, 2018).

Moreover, research suggests that adequate comprehension of spoken discourse depends on students knowing a high percentage of the vocabulary in a text. Stær (2009) found that 94% coverage was required to gain a 60% comprehension of advanced level listening texts while 98% coverage was needed for students to obtain a 70% score in a comprehension test. Similarly, van Zeeland and Schmitt (2013) found that adequate comprehension of spoken discourse required students to know between 95% and 98% of the words in a text. Coverage also appears to vary depending on the text type. For example, lower levels of coverage have been found for audio-visual material (Durbahn et al., 2020), dialogues (Giordano, 2021) and texts graded for intermediate level language learners (Noreillie et al., 2018). Another consideration in assessing the difficulty of academic spoken discourse is the number of words learners require to achieve a certain degree of coverage. Dang et al. (2014) found significant variation between disciplines with medical and life sciences requiring 5,000 words to achieve 95% coverage, compared with 3,000 words for social sciences. The present study used screen-capture video lectures. While the visual dimension of the lectures may help to reduce the coverage required for adequate comprehension, the content of the lectures was clearly technical, and they were delivered in the form of monologues. Hence, coverage of between 95% to 98% for the mathematics lectures was assumed would provide adequate comprehension.

Defining General, Academic and Technical Vocabulary

There are different categories of vocabulary in the corpus used in the present study. General, high-frequency vocabulary in English is commonly defined as the most frequent 2000 word families in English, although there is an argument for increasing this to 3000 word families (Schmitt and Schmitt, 2014). A word family consists of the headword (e.g., 'power'), inflected forms (e.g., 'powered') and derived forms (e.g., 'powerless'). Academic vocabulary, defined as words that 'have wide range and high frequency in academic texts' (Dang et al., 2017, p.6), such as 'analyse', has been seen by some researchers as separate from these high-frequency words; for example, Coxhead's Academic Word List (2000), which omitted words included in the General Service List (West, 1953). However, this does not take into account those high frequency words that also have a specialised meaning in a discipline, such as 'interest'. Therefore, some general academic wordlists; for example, the Academic Vocabulary List (Gardner and Davies, 2014) include this high-frequency vocabulary.

General academic vocabulary is distinct from technical vocabulary. As Ha and Hyland (2017) note, there is no single precise definition of technical language, although there is general agreement on a very specialised use within a discipline and infrequency outside the discipline. There are different categories of technical vocabulary (Dang, 2020, p.439); fully technical words that are known and used by specialists in the field (e.g., 'arctangent' in mathematics), lay-technical words that are understood by non-experts (e.g., 'multiply'), and polysemous words with both a general meaning and a specialised meaning in the discipline, like 'factor', whose mathematical meaning is very different from the general meaning. Technical vocabulary is often cited as a barrier to understanding lectures and coursebooks (Evans and Morrison, 2011, p.154; Evans and Green, 2007, p.13). Indeed, being able to understand technical vocabulary is key to understanding of the discipline (Woodward-Kron, 2008, p.246)

As mentioned above, vocabulary items that are commonly found on general service lists of the top 2000 or 3000 most common words in English, or indeed general academic word lists, may also have very specialised meanings in particular disciplines. This type of polysemous lexis, which is referred to as sub-technical (Mudraya, 2006) or crypto-technical (Fraser, 2009), may pose significant difficulties for learners; such technical words may be disregarded by learners if the general meaning is already known (Fraser, 2009, p.157). Accessing the technical meaning of polysemous words may also be problematic (Watson Todd, 2017).

Therefore, in order to fully understand the lexical challenge facing learners, the degree of technicality of vocabulary items in a corpus needs to be analysed (Ha and Hyland, 2017, p.36) in addition to consideration of the number of lexical items or word families that students need to know to achieve sufficient coverage of a text or corpus. An example of a study that has done this would be Fraser (2009), who identified and produced counts for what he termed crypto-technical vocabulary across the various sections of IMRaD-style (introduction, methods, results and discussion) medical research articles.

Identifying technical vocabulary

A number of methods have been employed to identify technical vocabulary in a corpus (Ha and Hyland, 2017). One way is to use the corpus-comparison approach. This method involves comparing a specialised corpus with a general English corpus. Words appearing only in the specialised corpus are assumed to be technical and words that meet a certain threshold of comparative frequency are likely to be technical (Coxhead, 2018, p.8).   However, some technical words may not be included using this approach, as polysemous words that have a technical meaning but also have a general high-frequency meaning may not be identified as relatively frequent in a discipline (Ha and Hyland, 2017, p.37). Collocations with a technical meaning may also be omitted (Kwary, 2011). Another method for identifying technical vocabulary is to conduct a keyword analysis. This method also involves the comparison of two corpora. However, in the case of keyword analysis a statistical measure is used to compare the frequencies of words in order to determine words that occur with an unusually high frequency (Coxhead, 2018, p.9). This method generates a keyness score which allows vocabulary items to be ranked. A further method involves making reference to specialised vocabulary knowledge. This may involve consulting specialists in the field or referring to specialist dictionaries in order to distinguish between technical and general or sub-technical vocabulary (Coxhead, 2018; Ha and Hyland, 2017). However, the criteria for inclusion of a word may not be transparent (Nation et al., 2016, p.147), and is subjective (Ha and Hyland, 2017).

The current study used a combination of keyword analysis and consulting specialist vocabulary knowledge. A key word list was first generated to broadly identify likely technical vocabulary before the list was categorised into different subtypes (technical, sub-technical and lay technical vocabulary). It was felt this method was efficient and was sufficiently objective.

Technical and sub-technical language in the discourse of mathematics

While polysemous language is a feature of academic texts in general (Coxhead 2016, p.181), it is particularly relevant to the discourse of mathematics. O'Halloran (2015) characterises this discourse as multimodal, integrating the language of mathematics, its symbolic notation (including superscript and subscript notation), and its graphs and diagrams. While symbols and diagrams may be universally understood, the complexity and technicality of the language of mathematics may present a greater challenge. Grammatical aspects of mathematical discourse include the way that logical relationships are expressed, using conjunctions in a precise way and the use of complex noun phrases, resulting in lexical density (Wilkinson, 2019, p.88). Regarding lexis, Halliday (1975, cited in Wilkinson, 2019, p.88) described the mathematical register as including phrases, such as 'complete the square', compound words, such as 'output', and words of Greek and Latin origin such as 'parabola'. Most significantly for the current study, there are also 'everyday words interpreted in the context of mathematics' (Wilkinson, 2019, p.88).

It is important to provide support regarding technical and sub-technical language. Students who are not proficient speakers and who may not have a good understanding of technical vocabulary in the context of mathematics may face difficulties on their university courses (Bedore, Pena, and Boerger, 2011, cited in Wilkinson, 2019, p.88). Text comprehension (in both listening and reading) is linked to knowledge of technical language, with students in some disciplines facing considerable challenge; Chung and Nation (2003) identified 37.6% of an anatomy text as being technical, and 16.3% of an applied linguistics text. There is also a link between knowledge of technical vocabulary and content knowledge. Bond (2020, p.104) notes that acquisition of disciplinary vocabulary is 'a key aspect of gaining access to target knowledge'. As well as having receptive knowledge of this vocabulary, students also need to be able to use it to become part of their discourse community (Wray, 2002, cited in Coxhead, 2016, p.178; Szudarski, 2018, p.140).

Specialised key word lists

The present study aims to help students with the high vocabulary load of their mathematics lectures through the creation of a key word list. The main purpose of wordlists is to identify key vocabulary and provide a target for vocabulary learning, allowing students to see their progress, which can be very motivating (Coxhead 2016, p.180). This is true of all wordlists, including interdisciplinary

academic wordlists, such as the Academic Word List (Coxhead, 2000), the Academic Vocabulary List (Gardner and Davies, 2014), and the Academic Spoken Word List (Dang et al., 2017). These interdisciplinary wordlists assume a common core of academic vocabulary across disciplines and can be useful in EGAP contexts (Dang et al., 2017, p.2; Coxhead 2018, p.22). In contrast, a more specialised wordlist focusing on a particular discipline can be even more beneficial for specific contexts. Such a wordlist may be seen as clearly relevant to the student's course and result in increased motivation (Hyland, 2016, p.20). It provides greater coverage of discipline-specific lexis (Fraser, 2009; Ward, 2009; Dang, 2018) and may result in more efficient learning of lexis (Hyland 2016, p.20).

A wordlist can be specialised in different ways: discipline- or genre-specific, or based on either a spoken or written corpus, for example. Using Becher's classification of academic disciplines into hard and soft (Becher, 1989), there are several wordlists available for hard sciences, such as Lei and Liu's (2016) New Medical Academic Wordlist, the Pharmacology Word List (Fraser, 2009), the Basic Engineering Wordlist (Ward, 2009), and the Engineering English Word List (Hsu, 2014). The closest list to one for mathematics is a list of academic lexical bundles (Alasmary, 2019) but this consists of phrases such as *if and only if* or *a set of all* rather than individual items of lexis.

While the above are derived from written corpora, there are far fewer wordlists based on spoken corpora. Examples include the Academic Spoken Wordlist (Dang et al., 2017), and the Academic Formulas List (Ellis and Simpson-Vlach, 2010), neither of which are discipline-specific. At the time of the current study, there is one wordlist based on a spoken corpus that is specific to hard science disciplines. Dang (2018) created the Hard Science Spoken Wordlist, based on a corpus of 6.5 million words from six hard-pure disciplines and six hard-applied disciplines, from a range of lectures, seminars, labs and tutorials. Regarding discipline-specific spoken wordlists, there is the Medical Spoken Word List (Dang, 2020), but there is very little that is specific to mathematics derived from either spoken or written corpora.

Bond (2020) discusses the need for specificity in EAP language support, working with disciplines in the university, studying the discourse of these communities and increasing awareness of the importance of discipline-specific language. It was felt that a general wordlist would not meet the needs of our specific context. Therefore, the primary aim of the study was to create a 'pedagogic corpus' (Willis, 1990, cited in Szudarski, 2018, p.108), comprising all the language that students

encounter on a module. Such a context-specific corpus is the most effective way to target key lexis and reduce vocabulary load for the students (Hyland and Tse, 2007, p.251). A study by Hou (2014) showed that a pedagogic corpus in a different context was used in the creation of learning materials which were successful in improving students' understanding of disciplinary vocabulary. From our pedagogic corpus, a keyword list was created to form the basis of materials designed to help SWJTU-Leeds Joint School students with the linguistic demands of their mathematics lectures.

METHODOLOGY

Creation of the corpus

Due to Covid, all the lectures used in the creation of the corpus we pre-recorded and thus mp4s were readily available. Permission to use these was obtained from the lecturer involved. Transcripts of the video lectures were created using Otter.ai (a commercially available online transcription application). These were then manually checked against the recordings for lexical and orthographic errors. Punctuation was only corrected in cases where errors impeded comprehension.

The 46 lectures were divided almost equally between the three researchers and each lecture was listened to while reading the transcript. Once corrected, the transcripts were searched using a text editor, Notepad ++, to correct typographic errors and to eliminate inconsistencies with hyphenation (pre-factor, three-dimensional), UK/US spellings (meter/metre, recognise/recognize, labeled/labelled), and compounds of words (arcsecant/arc secant, workout/ work out).

Finally, Sketch Engine was chosen to analyse the corpus data and produce a keyword lemma list, conforming to Bauer and Nation's level 2 of word families (1993). A lemma consists of a headword and inflections that are the same part of speech. For example, the lemma 'factor' as a noun would include 'factors' (plural). 'Factor' as a verb would include 'factors', 'factored' and 'factoring' in the lemma. Sketch Engine, however, combines headwords that are the same form but a different part of speech, so 'factor' as a verb and a noun are the same lemma.

Creation of the keyword list

Webb (2021) suggests that choice of unit of counting for wordlists (word type, lemma, word family) should be made based on the purpose of the list. As such, the lemma was chosen as the unit of counting when analysing and creating our keyword list. This is because research suggests L2 learners have sufficient knowledge to cope with lemmas but often lack morphological awareness to deal with word families (Brown et al., 2022, p.600; Gardner and Davies, 2014, p.30). Lemma was also helpful as several words in the list only occurred in a derived form, such as variable(vary), recursion(recur) and decomposition(decompose). The list would also be needed for both productive and receptive uses, in which case, lemmas are the preferred unit (Nation, 2016, p.26).

The simple maths formula (Kilgarriff et al., 2014) used in the keyword analysis feature of Sketch Engine allows for a variable to be set to change the focus of the list between more common and rarer items (Sketch Engine, no date). Higher values will bias the results in favour of frequency. In our corpus a list created with the variable set at one gives 'sine', which has a frequency in the focus corpus of 607, as the word with the highest keyness score. However, if the variable is set to 1000, 'minus', whose frequency is 1381, appears first in the keyword list and 'sine' appears eighth. The default score in Sketch Engine of one was chosen for our analysis.

As the focus of this research was on technical language rather than features of spoken English, a corpus containing only samples of spoken discourse was selected as the reference corpus, using the spoken component of the British National Corpus (BNC) (2014). Experiments using the full BNC corpus revealed a bias towards idiomatic language and colloquialisms, whereas using the spoken corpus removed these items.

The first version of the list contained a significant number of items of mathematical notation (x, y and dx for example) as well as proper nouns (Chengdu, Leeds) and numbers (one, two, three). As these either did not represent English words or would already be familiar to the students, they were added to a non-word list and excluded from the analysis. A keyword list containing only real words was then obtained.

Refining and categorising the keyword list

Once the keyword list was generated in Sketch Engine it was categorised into four groups: technical vocabulary, sub-technical words with both technical and non-technical meanings, lay-technical words, and a final group comprising high-frequency words with no technical meaning (see appendix 2). This was initially done individually by the three researchers and then any words where there was disagreement were checked together. To be categorised as technical, words had to be monosemous and included in the mathematical dictionary. These included terms such as cosecant, cotangent, and calculus. Words were categorised as sub-technical if they were assigned a general meaning in the Oxford Learners Dictionary (Oxford, 2021) and also had a technical meaning in the Oxford Concise Dictionary of Mathematics (Clapham and Nicholson, 2014). Examples of these terms include derivation, dummy, and log. Finally, words were categorised as lay-technical if they had a mathematical meaning but would be easily understood by a general audience. These words appeared in both the general and mathematical dictionaries and the mathematical meaning was the most common one. Examples of lay-technical items include multiply, radius, and decimal. Words that did not fit into the technical, sub-technical or lay-technical groups and were to be found on the top 2000 most frequent words in the Browne NGSL were then excluded from the list. The decision to exclude general words was justified on the basis that most students would be familiar with these vocabulary items, or they would have been required to learn them in order to pass the College English Test, which requires a knowledge of 4,200 words (Wei, 2004). It should be noted that much research (for example, Lu and Dang, 2022; Sun and Dang, 2020) suggests students at this level in China have not mastered the first 2000 most frequent words. However, as this study was aimed at helping them with their mathematics vocabulary knowledge, it was decided any generally frequent vocabulary would be better dealt with separately in their English lessons. Browne's New General Service list (2013b) was preferred over other lists as it was readily available in lemmas and in a format which could be used with Range and Ant Word Profiler without any further manipulation. Finally, vocabulary items that appeared five times or less in the corpus were excluded from the list as it was felt that learning these would not be beneficial for students.

The process of categorising the words on the wordlist was not straightforward. Some words found in both the technical and non-technical dictionaries appeared to make reference to the same concept with the main difference being the degree of precision of the definition. An example of this would be *variable* which is defined as 'a situation, number or quantity that can vary or be varied' in the Oxford Learner's Dictionary but in the Oxford Dictionary of Mathematics as:

an expression, usually denoted by a letter, that is defined for values within a given set. Can be used to represent elements of sets which are not numbers but frequently it relates to numerical quantities and functions defined in them together with the relationship between them.

This problem was further compounded by the fact that some words did not appear in the same form in the mathematical dictionary as in the corpus. For example, the corpus has examples of *primed*, while the mathematical dictionary only has *prime*. Another complicating factor was that words often occurred as part of collocations in the mathematical dictionary making direct comparison with the general dictionary difficult.

Once a definitive list of categorised key words had been obtained, it was decided to sub-divide the list. The reason for this was to spread the load of learning for the students. It was felt that words that occurred in all three lecture topics (series, derivatives and integrals) would be best presented to students at the beginning of the course. In order to accomplish this task Range (Heatley et al., 2002) was used to provide data on the number of occurrences in each sub-section of the corpus. The remaining words on the list were then assigned to one of three lists corresponding to each lecture series (see appendix 3).

RESULTS AND DISCUSSION

Categories of vocabulary in the keyword list

The final keyword list comprised 202 lemmas. This were categorised into four categories: technical terms, sub-technical, lay-technical and non-technical. In total 116 words were identified as sub-technical, 35 were found to be technical terms and 20 were designated as lay-technical. The other 31 words on the list were non-technical.

Analysis of the corpus

The mathematics lecture corpus contains 152,443 word tokens (a single occurrence of a word in a text), comprising 2,125 word types (1,496 lemmas).

Table 1 below shows the percentage of words the learners would potentially be able to understand if they knew the first 2000 words of the NGSL (Browne et al., 2013b) and the NAWL (Browne et al., 2013a). The first 1000 would cover just over 79% of the corpus and combined with the second 1000 that would rise to just over 84%. By learning the NAWL they could increase their understanding to nearly 89% of the words in the lectures. Again, the Browne NAWL list was selected because it was readily available in a format which could easily be imported into Range.

These results are broadly in line with other research which shows that a combination of knowledge of the NGSL and the NAWL provides very good coverage of academic texts. However, this does not take into account sub-technical words that possess both a general and a technical meaning; i.e., words where it is unlikely they would know the mathematical meaning despite knowing the more general meaning.

	TOKEN	TOKEN%	CUMTOKEN%
1st 1000	120642	79.12	79.12
2nd 1000	7680	5.04	84.16
NAWL	6879	4.51	88.67

Table 1: Browne NGSL + NAWL

Interestingly, the sub-technical vocabulary on our wordlist accounts for 8.82% of the tokens in the corpus (see Table 2), significantly higher than the coverage provided by the NAWL which suggests that lack of knowledge of this vocabulary may cause learners significant comprehension problems. However, it must be acknowledged that the sub-technical vocabulary was not always used in a technical sense in the corpus. For example, the word 'even' was most commonly found in the corpus in its general meaning, while 'product' was found exclusively in its technical sense.

	TOKENS/%	TYPES
Sub-technical	13450/ 8.82	116
Technical	2882/ 1.89	35
Lay-technical	2552/ 1.67	20
Nontechnical	766/ 0.50	31

Table 2: Coverage by word type

Coverage of keyword list

The maths wordlist (as it stands) covers nearly 13% of the vocabulary they encounter in the lectures (see table 3). This means that by learning the first 2000 word families from one of the NGSL lists plus our list will give them approximately 97% of the words needed to understand the lectures. This is a much better coverage than using the NAWL and the NGSL combined.

WORD LIST	TOKENS/%	TYPES/%	LEMMAS
Maths list	19650/12.89	359/16.93	202

Table 3: Coverage of whole list

Sub-technical vocabulary in the corpus

The sub-technical vocabulary in the corpus comprises vocabulary on both the NGSL and the New NAWL, while a number of words are not present in either of these two lists. In total, 56 terms were found to be present in the NGSL, while 33 were in the NAWL. 29 terms were not present on either list (see table 4).

Examples:

Sub-technical vocabulary also on the NAWL: derivative, integral, substitution

Sub-technical vocabulary on the NGSL: function, square, power, value

Neither list: alternate, diverge, inverse

29 were not present in the NAWL or NGSL

56 were in NGSL

33 were in NAWL

Table 4: Breakdown of sub-technical words by NGSL and NAWL

During the process of checking the lecture transcriptions on Otter.ai, it became clear that there was a high number of these 'everyday words interpreted in the context of mathematics' (Wilkinson, 2019, p.88). This sub-technical language comprised 8.82% of the tokens in the corpus, compared to just 1.89% for fully technical language. It is very difficult for people with limited understanding of mathematics to judge how similar a word is in its mathematical sense to its general meaning, but some general categories could be discerned.

- High-frequency words (in the 2000 most common word families in English) whose mathematical meanings may be inferred from their general meanings outside the context of mathematics; examples include 'limit' and 'boundary'. (Word frequency was checked using the VocabProfiler function of the Compleat Lexical Tutor site (Cobb, no date), which uses the BNC-COCA frequency lists.)

- Lower frequency words whose mathematical meanings are connected to their general meanings outside the context of mathematics. An example is 'cusp'.

- 'Opaque' (Watson Todd, 2017) or 'cryptotechnical' (Fraser, 2009) words which are high frequency outside the discipline. These words have a very different meaning in mathematics, so their meaning cannot be easily understood from their common meanings outside the discipline. (E.g., 'square')

- 'Opaque' words that are lower frequency, such as 'differentiate'.

- Words that are used as a different part of speech in mathematics; for example, 'constant' and 'bound', which are both used as nouns.

- Words with two meanings in mathematics. 'Prime', for example, may refer to prime number, or the prime symbol.

- Opaque words that form parts of collocations/multi-word items, like 'rational function'.

Potential challenges for students and tutors regarding polysemous lexis

There is a lack of literature on the potential language difficulties encountered by Mathematics students, such as those caused by polysemy. In comparing mathematical and general meanings of polysemous lexis for the current study, various potential issues for both students and tutors presented themselves. Students may find searching for mathematical definitions challenging. Often a mathematical definition does not appear in Google Translate or general dictionaries like the Oxford Advanced Learner's Dictionary or is near the end of the definitions. Specialised mathematics dictionaries such as the Oxford Concise Dictionary of Mathematics include all terms but are not written for English learners. Another issue is collocations and multi-word items. 'Common denominator' occurs in Google Translate and is translated into Mandarin, but 'arbitrary constant' does not; Google Translate translates each word separately.

Another difficulty is how to approach technical language when the language tutor has no content knowledge. The different categories of sub-technical lexis in mathematics suggest different approaches; knowledge of the non-mathematical meaning can be helpful when the word is frequent and has a connection with the mathematical meaning. Nation discusses core meanings of words and how meanings are often more specific in their technical sense (Nation 2013, p.295 and p.306). Bond (2020, p.101) suggests that such vocabulary may not present significant difficulties for students. However, comparison with the general meaning(s) is less helpful when the word is low frequency and perhaps unknown, or when meanings are very different. If an everyday meaning of a specialised word is known, it may affect comprehension of the same word with a different meaning in a different context (Coxhead, 2018 p.32).

Future directions

One aim of the current study was the indirect application of the corpus in creating teaching materials (as opposed to direct application, when data driven learning is used in the classroom (Flowerdew, 2009; Rohmer, 2011, cited in Szudarski, 2018, p.141)). According to Coxhead (2016, p.117), few studies 'go beyond simple frequency counts and also consider learnability and teachability'. Teachability is a concern in this mathematics context; using Nation's (2007) four strands of language learning, tutors working with subject specialists can provide some practice in comprehensible meaning-focused input, using lecture notes and concordance lines to provide context and encourage noticing. Language-focused learning is also possible, with a focus on collocation or pronunciation, for example. A lack of content knowledge on the part of the language

tutor means that meaning-focused output and fluency development would require collaboration with subject specialists.

Providing a glossary might be an efficient way of dealing with the issue of technical and sub-technical language in mathematics and might indeed be necessary with lexis that is not easily searchable. However, we hope to use our insights regarding polysemous language in mathematics in the creation of materials to raise awareness of aspects such as opacity and collocation. This awareness would be transferable to other contexts and would hopefully be beneficial to students later in their courses. A further benefit of these materials would be their potential to raise staff awareness of possible language issues for students.

It became apparent during this project that our assumption that the learners would already know the most frequent 2000 words may be incorrect. This has been addressed initially by focusing solely on maths vocabulary, however, we intend to test this assumption during the next academic year as it would have an impact on the content of their English lessons.

Our choice of the Browne NGSL and NAWL is also possibly questionable. It was mainly a choice based on the ease of availability as both lists were available as .xml files in lemma format with inflections included meaning they could be used instantly without needing any further manipulation. As they were used to remove only words of a non-mathematical meaning, we do not believe the choice of list had a meaningful effect on the result.

CONCLUSION

The aim of this project was to build a corpus and use that to analyse the lexical challenges faced by our learners. Clearly, they face a difficult task understanding the lectures from a purely lexical perspective. If we take van Zeeland and Schmitt's (2013) estimate that learners need to know a minimum of 95% of the words from a lecture to understand it, our learners will not come close to this by learning words from the NGSL plus NAWL (88.67%). We can solve this issue by combining the NGSL with our wordlist, which would allow them to achieve nearly 97% coverage.

The wordlist has allowed the separation of words into those which are technical and therefore outside the scope of EAP tutors and those which are nontechnical or lay-technical that can be taught by EAP tutors. The final category of sub-technical highlights the difficulty of exactly where to draw the line between subject knowledge and language knowledge. Here, the best approach appears to be one of consciousness raising of the learners to the high quantity of potentially opaque language within their discipline and strategies to cope with it.

Addresses for correspondence: a.j.drury@leeds.ac.uk; r.a.perkins@leeds.ac.uk; w.e.sheard@leeds.ac.uk

REFERENCES

Alasmary, A. 2019. Academic lexical bundles in graduate-level math texts: a corpus-based expert-approved list. *Language Teaching Research*. **26**(1), pp. 99-123.

Bauer, L. and Nation, P. 1993. Word families. *International Journal of Lexicography.* **6**(4), pp.253–279.

Becher, T. 1989. *Academic tribes and territories: intellectual enquiry and the cultures of disciplines*. Milton Keynes: Society for Research into Higher Education and Open University Press.

Bond, B. 2020. *Making language visible in the university: English for academic purposes and internationalisation*. Bristol: Channel View Publications.

Bondi, M. 2010. Perspectives on keywords and keyness: an introduction. In: Bondi, M. and Scott, M. eds. *Keyness in Texts*. Amsterdam/Philadelphia: John Benjamins Publishing Company, p.3.

The British National Corpus. 2014. [Online]. [Accessed 20 January 2022]. Available from: http://www.natcorp.ox.ac.uk/

Brown, D., Stoeckel, T., Mclean, S., and Stewart, J. 2022. The most appropriate lexical unit for L2 vocabulary research and pedagogy: a brief review of the evidence. *Applied Linguistics.* **43**(3), pp. 596-602.

Browne, C. Culligan, B. and Phillips, J. 2013a. *New academic word list*. [Online]. [Accessed 14 July 2021]. Available from: http://www.newgeneralservicelist.org/nawl-new-academic-word-list

Browne, C., Culligan, B. and Phillips, J. 2013b. *New general service list*. [Online]. [Accessed 14 July 2021]. Available from: http://www.newgeneralservicelist.org/

Chang, Y.Y. 2010. English-medium instruction for subject courses in tertiary education: reactions from Taiwanese undergraduate students. *Taiwan International ESP Journal.* **2,** pp.55-84.

Chung, T. and Nation, P. 2003. Technical vocabulary in specialised texts. *Reading in a Foreign Language*. **15**(2), pp.103-116.

Clapham, C. and Nicholson, J. 2014. *The concise Oxford dictionary of mathematics.* 5th ed. Oxford: Oxford University Press.

Cobb, T. [no date]. *Compleat lexical tutor*. [Online]. [Accessed 27 October 2022]. Available from: https://www.lextutor.ca/

Coxhead, A. 2000. A new academic wordlist. *TESOL Quarterly.* **34**, pp.213-238.

Coxhead, A. 2016. Acquiring academic and disciplinary vocabulary. In: Hyland, K. ed. *The Routledge handbook of English for academic purposes*. London: Routledge, pp.177-190.

Coxhead, A. 2018. *Vocabulary and English for specific purposes research: quantitative and qualitative perspectives.* London: Routledge.

Dang, T.N.Y. and Webb, S. 2014. The lexical profile of academic spoken English. *English for Specific Purposes*. **33**, pp.66-76

Dang, T. N. Y. 2018. A hard science spoken wordlist. *International Journal of Applied Linguistics*. **169**(1), pp.44-71.

Dang, T.N.Y. 2020. The potential for learning specialized vocabulary of university lectures and seminars through watching discipline-related TV programs: insights from medical corpora. *TESOL Quarterly.* **54**(2), pp.436-459.

Dang, T.N.Y., Coxhead, A. and Webb, S. 2017. The academic spoken word List. *Language Learning*. **67**(4), pp.959-997.

Dearden, J., 2015. *English as a medium of instruction: a growing global phenomenon*. London: British Council.

Doiz, A., Lasagabaster, D. and Sierra, J. M. 2012. *English-medium instruction at universities: global challenges*. Bristol: Multilingual Matters.

Durbahn, M., Rodgers, M. and Peters, E. 2020. The relationship between vocabulary and viewing comprehension. *System.* **88**, pp.1-13.

Ellis, N.C. and Simpson-Vlach, R. 2010. An academic formulas list: new methods in phraseology research. *Applied Linguistics*. **31**(4), pp.487-512.

Evans, S. and Green, C. 2007. Why EAP is necessary: a survey of Hong Kong tertiary students. *Journal of English for Academic Purposes.* **6,** pp.3-17.

Evans, S. and Morrison, B. 2011. The student experience of English-medium higher education in Hong Kong. *Language and Education.* **25**(2), pp.147–162.

Fang, W. and Wang, S. 2014. Chinese students' choice of transnational higher education in a globalized higher education market: a case study of W university. *Journal of Studies in International Education*. **18**(5), pp.475-494.

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

Fraser, S. 2009. Breaking down the divisions between general, academic, and technical vocabulary: the establishment of a single, discipline-based word list for ESP Learners. *Hiroshima Studies in Language and Language Education*. **12**, pp.151-167.

Galloway, N. and Ruegg, R. 2020. The provision of student support on English medium instruction programmes in Japan and China. *The Journal of English for Academic Purposes*. **45**(2020), pp.1-14.

Gardner, D. and Davies, M. 2014. A new academic vocabulary list. *Applied Linguistics.* **35**(3), pp. 305–327.

Gilmore, A. and Millar, N. 2018. The language of civil engineering research articles: a corpus-based approach. *English for specific purposes*. **51**, pp.1-17.

Giordano, M., J. 2021. [Forthcoming]. Lexical coverage in dialogue listening. *Language Teaching Research*. [Online]. [Accessed 11 October 2022]. Available from: https://journals.sagepub.com/doi/full/10.1177/1362168821989869

Ha, A.Y.H. and Hyland, K. 2017. What is technicality? A technicality analysis model for EAP vocabulary. *Journal of English for Academic Purposes*. **28**, pp.35-49.

Harada, T. and Uchihara, T. and 2018. Roles of vocabulary knowledge for success in English-medium instruction: self-perceptions and academic outcomes of Japanese undergraduates. *TESOL Quarterly*. **52**(3), pp.564-587.

Heatley, A., Nation, I. S. P. and Coxhead, A. 2002. *Range* (version 1.32). [software]. [Accessed 22nd April 2022]. Available from: https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/vocabulary-analysis-programs

Hou, H., I. 2014. Teaching specialized vocabulary by integrating a corpus-based approach: implications for ESP course design at the university level. *English Language Teaching*. **7**(5), pp.26-37.

Hsu, W. 2014. Measuring the vocabulary load of engineering textbooks for EFL undergraduates. *English for Specific Purposes.* **33,** pp.54-65.

Hyland, K. 2016. General and specific EAP. In: Hyland, K. and Shaw, P. eds. *The Routledge handbook of English for academic purposes*. Abingdon: Routledge, pp.17-29.

Hyland, K. and Tse, P. 2007. Is there an "academic vocabulary"? *TESOL Quarterly*. **41**(2), pp.235-253.

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P. and Suchomel, V. 2014. The Sketch Engine: ten years on. *Lexicography* (Berlin). **1**(1), pp. 7–36.

Kwary, D. A. 2011. A hybrid method for determining technical vocabulary. *System*. **39**(2), pp.175-185.

Lei, L. and Liu, D. 2016. A new medical academic word list: a corpus-based study with enhanced methodology. *Journal of English for Academic Purposes*. **22**, pp.42-53.

Lu, C., and Dang, T. N. Y. 2022. Vocabulary in EAP learning materials: what can we learn from teachers, learners, and corpora? *System.* **106**, pp.1-13.

Macaro, E., Curle, L., Pun, J., An, J. and Dearden, J. 2018. A systematic review of English medium instruction in higher education. *Language Teaching*, **51**(1), pp.36-76.

Mudraya, O. 2006. Engineering English: a lexical frequency instructional model. *English for Specific Purposes*. **25**, pp.235-256.

Nation, I. 2007. The four strands of learning. *Innovation in Language Learning*. **1**(1), pp.1-12.

Nation, I. 2013. *Learning vocabulary in another language.* 2nd ed. Cambridge: Cambridge University Press.

Nation, I. 2016*. Making and using word lists for language learning and testing*. Amsterdam: John Benjamin's Publishing Company.

Nation, P., Coxhead, A., Chung, T.M. and Quero, B. 2016. Specialized word lists. In: Nation, P. ed. *Making and using word lists for language learning and testing*. Amsterdam: John Benjamin's Publishing Company, pp. 145-152.

Noreillie, A., Kestemont, B., Heylen, K., Desmet, P. and Peters, E. 2018. Vocabulary knowledge and listening comprehension at an intermediate level in English and French as foreign languages an approximate replication study of Stæhr (2009). *International Journal of Applied Linguistics.* **169** (1), pp.212 – 231.

O'Halloran, K. 2015. The language of learning mathematics: a multimodal perspective. *Journal of Mathematical Behavior*. **40**, pp.63-74.

Oxford Learner's Dictionaries. [online]. 10th ed. 2021. [Accessed 12 July 2021]. Available from: www.oxfordlearnersdictionaries.com

Schmitt, N., and Schmitt, D. 2014. A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching.* **47**(4), pp.484–503.

Sketch Engine. [No date]. [Online]. [Accessed 20th May 2021]. Available from: https://www.sketchengine.eu/documentation/simple-maths/

Soruç, A. and Griffiths, C. 2018. English as a medium of instruction: students' strategies. *ELT Journal*. **72**(1), pp.38-48.

Stær, L., S. 2009. Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*. **31**(4), pp.577-607.

Sun, Y., and Dang, T. N. Y. 2020. Vocabulary in high-school EFL textbooks: texts and learner knowledge. *System.* **93**, pp.1-13.

Szudarski, P. 2017. *Corpus linguistics for vocabulary: a guide for research*. London: Routledge.

Universities UK International. 2020. International facts and figures 2020. London: Universities UK International.

Valipouri, L. and Nassaji, H. 2013. A corpus-based study of academic vocabulary in chemistry research articles. *Journal of English for Academic Purposes*. **12**(4), pp.248-263

van Zeeland, H. and Schmitt, N., 2013. Lexical coverage in L1 and L2 listening comprehension: the same or different from reading comprehension? *Applied Linguistics*. **34**(4), pp.457-479.

Ward, J. 2009. A basic engineering English word list for less proficient foundation engineering undergraduates. *English for Specific Purposes*. **28**, pp.170–182.

Watson Todd, R. 2017. An opaque engineering word list: which words should a teacher focus on? *English for Specific Purposes*. **45**, pp.31-39.

Webb, S. 2021. Word families and lemmas, not a real dilemma. *Studies in Second Language Acquisition.* **43**(5), pp.973-984.

Wei, D. 2004. Reflections on vocabulary size of Chinese university students. *International Education Journal*. **5**(4), pp.571-581.

West, M. 1953. *A general service list of English words*. London: Longman.

Wilkinson, L. 2019. Learning language and mathematics: a perspective from linguistics and education. *Linguistics and Education*. **49**, pp.86-95.

Woodward-Kron, R. 2008. More than just jargon – the nature and role of specialist language in learning disciplinary knowledge. *Journal of English for Academic Purposes.* **7**(4), pp.234–249.

A list of further reading is available from the authors.

APPENDIX 1

Full wordlist by frequency

0-51		52-102		103-153		154-202	
minus	1381	coefficient	75	simplify	32	manipulation	13
square	1263	pi	74	rotate	31	intermediate	13
derivative	1002	recursion	71	sub	31	cotangent	12
function	864	exponential	71	increment	31	unknown	11
series	750	separately	70	extend	30	calculus	11
integral	739	true	70	finite	30	outermost	11
times	627	graph	69	quantity	30	sufficiently	11
sine	607	irreducible	68	subtract	29	min	10
term	496	absolute (value)	68	equivalent	28	max	10
cosine	475	slope	68	explicitly	26	exact	10
factor	453	chain	66	inner	26	circumference	10
power	367	arctangent	66	individually	25	diagram	10
value	336	quadratic	65	implicitly	24	systematic	10
evaluate	317	equation	65	geometric	23	stack	10
give	298	convergence	64	trigonometric	22	bound	9
dot	278	agree	64	essentially	22	decimal	9
cube	263	polynomial	62	proper	20	diagonal	9
constant	227	logarithm	61	corresponding	20	statement	9
limit	225	rational	58	geometrical	20	branch	8
converge	218	identity	57	raise	19	parabola	8
log	212	delta	54	straightforward	19	arcsecant	8
root	212	segment	53	verify	19	algebra	8
taylor	206	exponent	53	obtain	19	calculator	8
formula	202	arc	51	expansion	18	bracket	8
curve	186	contour	50	expand	18	legitimate	8
even	185	plot	49	machinery	18	argument	8
integrate	183	alternate	49	outer	18	transform	7
substitution	178	sequence	49	namely	18	cusp	7
expression	178	rectangle	49	differentiate	18	inflection	7
prime	170	strip	49	infinitely	17	divergence	7

tangent	168	notation	48	cosecant	17	condition	7
integration	166	degree	45	local (max/min)	17	accuracy	7
denominator	160	integer	45	p-series	17	geometrically	7
infinity	155	repeat	45	angle	17	geometry	7
multiply	151	definite	44	valid	17	multiplication	7
secant	147	variable	43	approximation	16	compute	7
theta	145	rid	42	continuous	16	tricky	7
fraction	140	numerator	40	calculation	16	specify	7
product	135	derive	40	chop	16	precise	7
factorial	133	differentiation	39	strictly	16	improper	6
partial	132	inverse	38	maximum	15	approximate	6
convert	132	implicit	38	radius	15	composite	6
insert	127	arcsine	38	satisfy	15	representation	6
cancel	126	cone	37	parameter	15	proof	6
infinite	120	endpoint	36	conversion	14	accurate	6
interval	117	quotient	35	common (denominator)	14	correspond	6
ratio	104	inequality	33	triangle	14	non-zero	6
trig	96	substitute	33	correctly	14	parametric	6
decomposition	95	axis	33	preliminary	14	manipulate	6
diverge	91	arbitrary	33	indefinite	13		
linear	89	odd	32	minimum	13		

Wordlist by word type and frequency

Technical		Sub-technical		Lay-technical		General	
sine	607	square	1263	minus	1381	insert	127
cosine	475	derivative	1002	times	183	separately	70
Taylor	206	function	864	multiply	151	true	70
denominator	160	series	750	fraction	140	strip	49
secant	147	integral	739	rectangle	49	notation	48
theta	145	term	496	cone	37	rid	42
factorial	133	factor	453	subtract	29	inner	26
trig	96	power	367	angle	17	individually	25
coefficient	75	value	336	calculation	16	essentially	22
pi	74	evaluate	317	radius	15	straightforward	19
recursion	71	give	298	triangle	14	obtain	19
absolute	68	dot	278	circumference	10	machinery	18
arctangent	66	cube	263	diagram	10	outer	18
quadratic	65	constant	227	decimal	9	namely	18
polynomial	62	limit	225	diagonal	9	valid	17
logarithm	61	converge	218	algebra	8	chop	16
arc	51	log	212	calculator	8	strictly	16
integer	45	root	212	bracket	8	satisfy	15
numerator	40	formula	202	multiplication	7	correctly	14
arcsine	38	curve	186	compute	7	preliminary	14
geometric	23	even	185			manipulation	13
trigonometric	22	integrate	183			intermediate	13
geometrical	20	substitution	178			outermost	11
cosecant	17	expression	178			sufficiently	11
local (minimum/maximum)	17	prime	170			systematic	10
p-series	17	tangent	168			stack	10
Common (denominator)	14	integration	166			legitimate	8
cotangent	12	infinity	155			tricky	7
calculus	11	product	135			specify	7
parabola	8	partial	132			precise	7

arcsecant	8	convert	132	manipulate	6
geometrically	7	cancel	126		
geometry	7	infinite	120		
non-zero	6	interval	117		
parametric	6	ratio	104		
		decomposition	95		
		diverge	91		
		linear	89		
		exponential	71		
		graph	69		
		irreducible	68		
		slope	68		
		chain	66		
		equation	65		
		convergence	64		
		agree	64		
		rational	58		
		identity	57		
		delta	54		
		segment	53		
		exponent	53		
		contour	50		
		plot	49		
		alternate	49		
		sequence	49		
		degree	45		
		repeat	45		
		definite	44		
		variable	43		
		derive	40		
		differentiation	39		
		inverse	38		
		implicit	38		
		endpoint	36		
		quotient	35		

inequality	33
substitute	33
axis	33
arbitrary	33
odd	32
simplify	32
sub	31
rotate	31
increment	31
extend	30
finite	30
quantity	30
equivalent	28
explicitly	26
implicitly	24
proper	20
corresponding	20
raise	19
verify	19
expansion	18
expand	18
differentiate	18
infinitely	17
approximation	16
continuous	16
maximum	15
parameter	15
conversion	14
indefinite	13
minimum	13
unknown	11
min	10
max	10
exact	10
bound	9

statement	9
branch	8
argument	8
transform	7
cusp	7
inflection	7
divergence	7
condition	7
accuracy	7
improper	6
approximate	6
composite	6
representation	6
proof	6
accurate	6
correspond	6

APPENDIX 3

Wordlist by subject and frequency

Useful to all		Derivatives		Integrals		Series	
minus	1381	integral	739	series	750	Taylor	206
square	1263	tangent	168	converge	218	factorial	133
derivative	1002	secant	147	integrate	183	infinite	120
function	864	theta	145	substitution	178	diverge	91
sine	607	quadratic	65	integration	166	convergence	64
term	496	identity	57	partial	132	alternate	49
cosine	475	delta	54	interval	117	sequence	49
factor	453	segment	53	decomposition	95	geometric	23
power	367	arc	51	linear	89	expand	18
value	336	contour	50	recursion	71	p-series	17
evaluate	317	plot	49	irreducible	68	bound	9
give	298	repeat	45	absolute	68	divergence	7
dot	278	differentiation	39	polynomial	62	condition	7
cube	263	inverse	38	rational	58	accurate	6
constant	227	implicit	38	strip	49		
limit	225	arcsine	38	rectangle	49		
root	212	substitute	33	degree	45		
log	212	inner	26	cone	37		
formula	202	implicitly	24	endpoint	36		
curve	186	trigonometric	22	inequality	33		
even	185	geometrical	20	odd	32		
times	183	outer	18	sub	31		
expression	178	angle	17	rotate	31		

Word	Count	Word	Count	Word	Count
prime	170	local	17	finite	30
denominator	160	chop	16	proper	20
infinity	155	calculation	16	machinery	18
multiply	151	radius	15	expansion	18
fraction	140	maximum	15	infinitely	17
product	135	preliminary	14	cosecant	17
convert	132	triangle	14	approximation	16
insert	127	conversion	14	continuous	16
cancel	126	minimum	13	correctly	14
ratio	104	cotangent	12	manipulation	13
trig	96	outermost	11	indefinite	13
coefficient	75	unknown	11	systematic	10
pi	74	calculus	11	circumference	10
exponential	71	stack	10	diagram	10
separately	70	diagonal	9	min	10
true	70	legitimate	8	max	10
graph	69	branch	8	exact	10
slope	68	argument	8	decimal	9
chain	66	arcsecant	8	algebra	8
arctangent	66	specify	7	calculator	8
equation	65	cusp	7	parabola	8
agree	64	inflection	7	tricky	7
logarithm	61	geometrically	7	transform	7
exponent	53	geometry	7	improper	6
notation	48	manipulate	6	approximate	6
integer	45	composite	6	non-zero	6
definite	44	representation	6		
variable	43	proof	6		

rid	42	parametric	6
derive	40		
numerator	40		
quotient	35		
axis	33		
arbitrary	33		
simplify	32		
increment	31		
quantity	30		
extend	30		
subtract	29		
equivalent	28		
explicitly	26		
individually	25		
essentially	22		
corresponding	20		
obtain	19		
straightforward	19		
verify	19		
raise	19		
namely	18		
differentiate	18		
valid	17		
strictly	16		
satisfy	15		
parameter	15		
common	14		
intermediate	13		

sufficiently	11
statement	9
bracket	8
precise	7
multiplication	7
compute	7
accuracy	7
correspond	6