Introduction

Technical vocabulary generally refers to words and phrases that are used and known mainly in a specific profession, trade, or, for simplicity purposes, subject area (henceforth, “subject area” or simply “subject” will be used as a generic term to refer to both “profession” and “trade”). Because of its subject-specific nature, technical vocabulary varies significantly from one subject area to another. Furthermore, technical words are ubiquitous and highly frequent in professional language. As such, technical vocabulary constitutes a very important and required knowledge for those who work directly or indirectly in a subject area as well as for students studying the subject. This chapter attempts to explore the key issues and topics related to technical vocabulary. Future directions in the work on technical vocabulary will also be briefly discussed.

Critical Issues and Topics

In this section, the following critical issues related to technical vocabulary will be addressed: (1) its definition, (2) its identification, (3) its role and importance in language use and learning, and (4) strategies and methods for learning and teaching technical vocabulary. However, the issue of the identification of technical vocabulary will be explored in much greater length due to its central importance in the study of technical vocabulary and its technical complexity.

Definition of Technical Vocabulary

On the surface, technical vocabulary seems to be a straightforward term referring to lexical items used with specialized meanings in a subject known mainly to a particular community of users. Yet, a close examination of the research on the topic reveals a lack of true consensus on not only how to define but also how label technical vocabulary. Regarding its labels, technical vocabulary has been referred to variously as “discipline specific vocabulary” (Woodward-Kron, 2008); “domain-specific glossaries” (Periñán-Pascual, 2015);
“scientific/technical terms” (Yang, 1986); “specialised lexis” (Baker, 1988); “specialist terms” (Woodward-Kron, 2008, p. 239); “specialized vocabulary” (Robinson, 1980); “terminological words”, “terms”, or simply “terminology” (Bečka, 1972, pp. 47–48; Kit & Liu, 2008; Peruzzo, 2014); and “terminological units” (Cabrè, 1999). However, more recently, as will be explained later in the section, “technical vocabulary” has become an established term in applied linguistics to refer to a specific category of specialized words different from academic vocabulary, another group of specialized words (Nation, 2013; Nation & Coxhead, 2012).

As for the variations in the definition of technical vocabulary, first, there appear to be two views on the scope of subject areas that may be considered technical: a broader view and a narrower view. The broader view holds that technical words may occur in any subject (be it an art, science, or engineering) and they include words of various types, ranging from those that are used almost exclusively in a subject area (e.g., arthrodesis and laparoscopy in surgery medicine) to those that boast a high-frequency in general language but are used with a subject-specific meaning (e.g., balance and interest in banking business) or are important concepts in a subject without a separate subject-specific meaning (e.g., neck and nose in anatomy). In short, in this broader view, “Technical vocabulary is subject related, occurs in a specialist domain, and is part of a system of subject knowledge” (Chung & Nation, 2004, p. 252).

In contrast, the narrower view assumes that technical vocabulary is confined to specialized words in hard science, engineering, medicine, or trades only, as is evidenced by the following quote on technical vocabulary learning from the webpage of the Applied Linguistics Program at the University of Warwick (Technical Vocabulary, 2017):

technical vocabulary is . . . often [found] in the fields of Science, Engineering and Medicine. . . . In Arts, Humanities and Social science disciplines, there will also be a requirement to use what may be termed “specialised” vocabulary, though this will not usually be deemed to be “technical”.

Based on this quote, words with specialized meanings in arts, humanities, and social science are simply specialized vocabulary, not technical vocabulary. This restricted view of technical vocabulary can also be seen in Brieger and Pohl’s (2002) textbook on technical vocabulary and Ardasheva and Tretter’s (2017) research on the teaching of technical vocabulary to high school ESL students because both studies focused exclusively on vocabulary in engineering, sciences, and trades.

This restricted view has several problems. First, in today’s world, the distinction between disciplines is not always clear-cut thanks to the increased interdisciplinary work across many different fields and also to the increased use of technology and scientific research methods in arts, humanities, and social sciences. Second, the use of the term “specialized vocabulary” to refer to technical terms in arts, humanities, and social sciences only adds to the confusion of the terms used in the discussion of technical vocabulary, because, as noted earlier, “specialized vocabulary” has actually often been used to refer to technical vocabulary. Furthermore, more recently, the term “specialized vocabulary” has been used as a generic term that covers both “academic vocabulary” and “technical vocabulary” (Nation, 2013; Nation & Coxhead, 2012). Although both academic and technical words are specialized vocabulary, they differ in that while academic vocabulary consists of words that occur across a wide range of subjects, technical vocabulary is composed of words with a specialized meaning used usually in one specific subject. Such a distinction is not really new, however, as it can be traced to
Bečka’s (1972, p. 48) classification of specialized terminology into “notional terms” and “descriptive and technical terms”: “Notional terms form the actual core of the terminology of scientific disciplines whilst descriptive and technical terms form the nomenclature of the respective branches of science”. By defining “notional terms” as “core of the terminology of scientific disciplines” and “descriptive and technical terms” as words of “specific branches of science”, Bečka appears to consider the former to be general academic vocabulary and the latter to be discipline-specific technical vocabulary.

However, it is worth noting that whereas the distinction between academic and technical vocabularies is clear and sensible, there are some overlapping items between the two groups thanks largely to the fact that sometimes a word may have different specialized meanings in different subjects. For example, the word tension may refer to electricity tension in physics, mental tension in psychology, and muscle tension in physiology (Bečka, 1972, p. 49). Because of its high dispersion across disciplines, tension is on Coxhead’s (2000) Academic Word List (AWL), but it could also be a technical word for physiology, physics, and psychology, respectively, due to its specialized meaning in each of the fields.

Another difference concerning the definition of technical vocabulary is that not all scholars agree that those high-frequency technical words used with or without a specialized meaning in a subject are technical terms. For example, while Ha and Hyland (2017) do not consider those without a specialized meaning technical words, Baker (1988) and Yang (1986) consider all high-frequency ones (with or without a specialized meaning) to be only “sub-/semi-technical words/terms”. To the latter scholars, only words that have an exclusive technical meaning, such as arthrodesis, deserve the label of technical vocabulary. However, such a view and the attempt to differentiate the two types of technical words may be problematic if examined closely. First the technical and semi-/sub-technical differentiation is arguably not particularly meaningful, especially for language learning and teaching purposes. In fact, the so-called sub-technical words are actually more difficult to learn due to their polysemous meanings and the fact that their technical meanings are often “opaque” according to Watson Todd (2017). Second, the differentiation may not always be easy to make because, as noted in the previous paragraph, many technical words actually have more than one specialized meaning and some may appear across several different subject areas. Such technical words are labeled as sub-technical by some scholars (Baker, 1988, pp. 96–97; Yang, 1986, p. 95) and many of them, such as accuracy and normal, are on Coxhead’s (2000) AWL.

As for the differentiation between those technical words with a specialized meaning and those without one, the problem is that such differentiation is sometimes gradable rather than binary, as shown in the Ha and Hyland (2017) study to be discussed later. Take for example the aforementioned words in the two categories (neck and nose as technical words in anatomy without a specialized meaning vs. balance and interest in banking business with a specialized meaning). Whereas it is fairly safe to say that neck and nose have no specialized meanings in anatomy and that interest has a specialized meaning in banking, we cannot really say for certain whether balance has a truly specialized meaning. This is because we can argue that the special meaning of balance in banking is actually a metaphorical extension of its common meaning. In fact, as contemporary cognitive linguistics has shown, the different meanings of a polysemous word are generally related, with most of its meanings being metaphorical extensions of its core meaning (Tyler, 2012).

If we agree that the two differentiations just mentioned are not particularly meaningful and feasible in the definition of technical vocabulary, then, technical vocabulary may indeed come not only from general common high-frequency words (as in the case of the aforementioned words balance and interest) but also from academic words (as in the case of tension).
In fact, a large majority of technical words are high-frequency and academic words. As reported in Nation (2013, p. 20), an analysis of the composition of the words in an academic textbook reveals that 68.5% of the words in the textbook belong to the first 2,000 words (i.e., those in the General Service List [GSL]), 6.9% are academic words (i.e., those in the AWL), 20.6% are technical words, and 4% are other words. Of the 20.6% technical words, 9.2% are those in the GSL, 6.2% are those in the AWL, and only 5.2% are not from the first two categories. This means that when the technical words are treated as a separate group with a 100% value, 45% of them are common high-frequency words used with a specialized meaning, 30% are AWL words, and only 25% are low-/lower-frequency technical words.

In short, technical vocabulary is subject-bound, referring to words used in a specific subject for communicating subject-specific knowledge. It includes both high-frequency and academic words that are used with a specialized meaning in a specific subject as well as those low-/lower-frequency words that appear almost exclusively in a subject. Together with academic words, technical words help form specialized vocabulary. Finally, before we move onto the next section, it is necessary to note that technical vocabulary is not limited to individual words. It includes multiword units and their acronyms, such as Acquired Immunodeficiency Syndrome (AIDS), flux theorem, and static electric field (Brieger & Pohl, 2002; Yang, 1986). In fact, multiword technical units are common and many studies on technical vocabulary have included multiword terms (e.g., Nazar, 2016; Periñán-Pascual, 2015; Yang, 1986).

Identification of Technical Vocabulary

Existing methods used for identifying technical vocabulary can be categorized into two major groups: (1) judgment-based (e.g., Brieger & Pohl, 2002; Chung & Nation, 2003) and (2) corpus-based (e.g., Bernier-Colborne & Drouin, 2014; Chung, 2003), although it is important to note that many studies have now incorporated both types of methods (e.g., Ha & Hyland, 2017; Kwary, 2011; Watson Todd, 2017). Judgment-based methods require subjective decisions made based on subject-domain knowledge and they may involve the use of a rating scale, a technical dictionary, and/or contextual clues of words in the text in which they appear, as illustrated in Chung and Nation (2004). The use of a dictionary is considered a judgment-based method because many traditional dictionaries have been compiled based on the compilers’ intuitive decisions. In contrast, corpus-based methods typically employ computerized search, analysis, and comparison of various frequency and dispersion (the number of occurrence of a word across texts/corpora) measures and linguistic features of lexical items for automatic or semi-automatic identification of technical terms (e.g., Bernier-Colborne & Drouin, 2014; Chung, 2003; Kit & Liu, 2008; Nazar, 2016; Periñán-Pascual, 2015). Although corpus-based methods still involve some judgment decisions, such as the selection of the features to consider and statistical formulas to use, such approaches, being largely quantitative in nature, are generally more objective and reliable. More importantly, thanks to the rapid advancements in computer technology and corpus linguistics, corpus-based methods have become increasingly more sophisticated and effective and are now widely used. Later, we describe the two groups of methods and their procedures and discuss their respective strengths and weaknesses where appropriate.

Before we proceed to the discussion of these methods, it is necessary to clarify the issue involving the use of “word family”, “lemma”, or “word type (form)” as the counting unit in the identification and reporting of technical vocabulary. As we know, in the development of word lists, some scholars (e.g., Coxhead in her 2000 AWL) use word family, others (e.g., Gardner and Davies in their 2014 Academic Vocabulary List) use lemma. In technical
vocabulary research, “word type” has also been used (e.g., Baker, 1988; Chung, 2003; Ha & Hyland, 2017). It seems that “word family” is not a very good option because often not all members of a word family are technical words (Chung & Nation, 2003) and a word-family-based selection threshold may miss a technical word in the family (Ha & Hyland, 2017).

Concerning judgment-based methods, as noted previously, such methods require experts in the subject area (sometimes researchers, material writers, and/or teachers) to determine, based on their knowledge of the subject domain, whether and/or to what extent a word is related to and has specific technical meaning in the subject in question. The reason for using experts in this method is that, as discussed earlier, technical vocabulary is subject-related; making sound judgment decisions regarding technical vocabulary, therefore, requires a solid knowledge of the subject domain in question. Furthermore, while some technical words – mainly those of Latin or Greek origin (e.g., *pectora* and *vertebrae* in anatomy or medicine in general) – are easy to identify with a good accuracy, most other technical words are actually often difficult to determine, especially those that are high-frequency words used with or without a specialized meaning in a subject, a problem that researchers have long noted (e.g., Bečka, 1972). Thus, as already mentioned, judgment decisions on technical vocabulary are largely intuition-based and may not have good reliability. Therefore, researchers (e.g., Chung & Nation, 2003; Watson Todd, 2017) have used rating scales to help make the judgment more valid and reliable. Chung and Nation (2003) used a four-step rating scale (shown in Table 8.1) to identify technical words in an anatomy textbook and an applied linguistics textbook.¹

Based on this rating scale, words belonging to Step 3 or 4 are considered technical terms while those in Step 1 or 2 are deemed nontechnical. The rating scale was found to be highly reliable with an inter-rater reliability of 0.95. By applying the rating scale, Chung and Nation (2004) identified 4,270 technical word types (37.6% of a total 11,305 word types) in the anatomy textbook and 835 technical word types (16.3% of total of 5,137 word types) in the applied

<table>
<thead>
<tr>
<th>Table 8.1 A rating scale for finding technical words (as applied to the anatomy text)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1</strong></td>
</tr>
<tr>
<td>Words such as function words that have a meaning that has no particular relationship with the field of anatomy, that is, words independent of the subject matter. Examples include the, is, between, adjacent, amounts.</td>
</tr>
<tr>
<td><strong>Step 2</strong></td>
</tr>
<tr>
<td>Words that have a meaning that is minimally related to the field of anatomy in that they describe the position, movements, or features of the body. Examples include <em>superior, part, forms, pairs, structures.</em></td>
</tr>
<tr>
<td><strong>Step 3</strong></td>
</tr>
<tr>
<td>Words that have a meaning that is closely related to the field of anatomy. They <em>refer</em> to parts, structures, or functions of the body, such as the regions of the body and systems of the body. The words may have some restrictions of usage depending on the subject field. Examples include <em>chest, trunk, neck, abdomen, breast.</em> Words in this category may be technical terms in a special field like anatomy and yet may occur with the same meaning in other fields and not be technical terms in those fields.</td>
</tr>
<tr>
<td><strong>Step 4</strong></td>
</tr>
<tr>
<td>Words that have a meaning specific to the field of anatomy and are not likely to be known in general language. They refer to structures and functions of the body. These words have clear restrictions of usage depending on the subject field. Examples include <em>thorax, sternum, costal, vertebrae, pectora.</em></td>
</tr>
</tbody>
</table>

¹ Source: Chung and Nation, 2003, p. 105, but with some examples omitted. Reprinted by permission of *Reading in a Foreign Language*
linguistics textbook. Their finding indicates that the percentage of technical words in a specialized text varies substantially across subject areas with those texts in sciences/engineering/trades boasting a much higher percentage than those in arts/humanities/social sciences.

Now, let us have a brief look at the other two judgment methods: using dictionaries and using contextual information. In the dictionary method, the first step is to choose a good, highly reputable technical dictionary. Usually the more widely and the longer a dictionary has been in use, the higher its quality will be. In terms of how to use a dictionary to identify technical words in a text, usually, for a word to be a technical item in the subject, it needs to be listed as a main or subentry in the dictionary (Chung & Nation, 2004). Regarding the use of contextual clues in a text to identify technical vocabulary, generally, there are three main types of clues provided by the authors of the text: “(i) definitions, (ii) typographical clues like bolding, italics, and brackets, and (iii) labels in diagrams or illustrations” (Chung & Nation, 2004, p. 256). As for the effectiveness and accuracy of the two methods, Chung and Nation’s (2004) comparison of the two methods against the use of a rating scale indicates that using a technical dictionary has a higher accuracy rate than using contextual clues. Ha and Hyland (2017) made the best use of dictionaries and contextual information (in the form of examining concordance lines), but their study used a combined approach involving both corpus analysis and judgment decisions, so it will be discussed later in the section on combined methods.

Regarding corpus-based methods, as already explained, they typically examine and compare the frequency and dispersion measures of words and use computer programs to process such information to automatically or semi-automatically identify technical words of interest. The automatic process so used has been known as “automatic term extraction/recognition”, among others (Kageura & Umino, 1996; Kit & Liu, 2008; Periñán-Pascual, 2015). Depending on the aspects of the information searched and analyzed and on the procedures used, corpus-based term extractions can be grouped into two approaches: “linguistically oriented” and “statistically oriented” although most studies employ both approaches simultaneously (Kageura & Umino, 1996, p. 271). Linguistically oriented approaches focus first on the formal features of words, such as parts of speech and grammatical structures (e.g., noun-noun or adjective-noun structure), and/or the semantic features, such as the semantic role of a lexical unit, semantic relationships, and semantic similarities. Earlier linguistically oriented approaches focused mostly on the formal aspects of words as reported in Kageura and Umino (1996), but the recent decade has seen studies focused on semantic features (e.g., Bertels & Speelman, 2014; Peruzzo, 2014). Usually, after the analysis and tagging of linguistic and/or semantic features, linguistically oriented approaches submit the features to various frequency statistical analysis. In some cases, after terms have been automatically extracted, they are then checked and verified by the researcher based on the relevant semantic features already specified.

In contrast, statistically oriented approaches take a straightforward frequency approach by simply and directly computing and comparing the frequency and dispersion measures of lexical items in a technical corpus of interest against a large general language corpus. This is because technical terms almost always occur much more frequently in a technical corpus than in a general corpus given that both corpora are large and representative enough of the type of language they each represent. This fact of technical terms occurring significantly more frequently in a technical corpus forms the basis of the statistical formulas or procedures used to identify and select technical terms. The most basic and straightforward formulas simply compute and compare the differences in the raw frequency and dispersion numbers of the words between the technical corpus and the general language corpus used in a study
Technical Vocabulary

(e.g., Baker, 1988). More sophisticated formulas involve the comparison of more complex measures, such as the ratio between the frequency or the percentage of frequency of a lexical item in a specialized corpus and its frequency or percentage of frequency in a large general language corpus (Ahmad, Davies, Fulford, & Rogers, 1994) and the “rank difference” of a word in the two comparison corpora (Kit & Liu, 2008). Other sophisticated formulas include even more complex measures, such as Nazar’s (2016) comparative “distributional analysis” of the characteristics of the co-occurrence patterns of words in a technical and a general corpus and Periñán-Pascual’s (2015) “composite measure” that combines well-weighted values of a lexical item’s “salience”, “relevance”, and “cohesion”.

Furthermore, the computer programs used in corpus-based approaches often include some facilitating components to enhance their extraction accuracy and efficiency, such as the inclusion of a “stop list” of words to automatically exclude nontechnical words, such as function words. Finally, it is important to reiterate that in corpus-based comparison approaches, the general corpus should be very large, usually much larger than the technical corpus and it should not include any texts from the latter (Nation & Coxhead, 2012). Below, we offer a brief chronological overview of the development of the corpus-based approaches and a basic description of the procedures used in the studies reviewed.

Bečka (1972) and Yang (1986) were among the earliest corpus-based studies on technical vocabulary, although Bečka’s main focus was the examination of vocabulary composition in technical texts, not technical vocabulary identification. Bečka’s (1972) compared the frequencies and dispersions (what he called “disponibility”) of three types of words in a balanced Czech technical corpus: grammatical (i.e., function), nontechnical, and technical. It was found that the words with the highest frequencies and dispersions were all function words, words with high to moderate frequencies and dispersions were mainly nontechnical words, and words with the lowest frequencies (one or two occurrences) and dispersions were mainly technical terms. It is important to note that, being among the earliest corpus-based work, Bečka’s study did not employ any true statistical analysis, nor did it seem to have been done using a computer program because some of the results were reported as “estimation” numbers (Bečka, 1972, p. 51). Also, the study failed to include a general language corpus for comparison purposes. In this sense, the study is not a true linguistically or statistically oriented corpus-based study.

Yang (1986) used computer programs to investigate the possibility of identifying technical vocabulary in a science/engineering corpus of nine texts plus a general language text (a novel) for comparison purposes. Via various frequency and distribution analysis, Yang demonstrated that it was possible to use corpus analysis to automatically identify what he labeled as “sub-technical words” and “technical terms”. Like Bečka’s results, Yang’s also showed that words with both a high frequency and dispersion tend to include many function words while words with a relatively low frequency but a high dispersion consist of many sub-technical words (such as accuracy, conclusion, and feature (many of which, as noted earlier, are actually AWL words). In comparison, words that exhibit a very low dispersion include many technical terms (such as bandwidth, carboxyl, and electrolysis). Yang also showed the possibility to identify multiword technical terms by tagging the parts of speech of the words in the corpus and by using a computerized program with a stop list. However, Yang did not give any definitive cutoff frequency and dispersion measures for identifying sub-technical and technical terms. Hence, although Yang employed both linguistic and statistical analyses, his study is not a true linguistically or statistically oriented study, either.

Baker (1988) and Chung (2003) were among the early studies that might be said to have used a true statistically oriented corpus approach. Using a medical and two comparison
corpora – a smaller medical corpus and the 7.3 million-word COBUILD Corpus of General English. Baker (1988) first identified the 218 most frequent words in the medical corpus and then compared the frequencies of the words in the medical corpus with their frequencies in the two comparison corpora. Due to a significant size difference among the corpora, Baker (1988, p. 95) used frequency percentage (number of occurrence divided by total number of words in the corpus) and “ratio of frequency percentages (RFP)” in her comparison. Words with a low RFP (defined as 5 or below) between the medical and General English corpora were considered general lexis items, such as the, whose RFP was 1.3 (5.57 FP in medical corpus/4.23 FP in general English corpus). In contrast, words with a very high RFP ratio (defined as 300 or above, i.e., an occurrence of 300 times more in the medical corpus than in the general corpus) and a similar ratio between the medical and small medical corpus were identified as technical terms. Of the 218 words, 92 were general lexis items, 65 were sub-technical terms, and 61 were technical terms.

To identify technical words in anatomy, Chung used an anatomy corpus (the same one with 425,192 tokens used in Chung and Nation (2003) mentioned earlier) and a comparison corpus of general language, three times larger, made up of approximately 1,892,000 tokens from the LOB and Wellington corpora (with all the texts in natural sciences in them excluded). Using the computer program RANGE, Chung identified a total of 66,223 word types in the two corpora. Then, she excluded the 54,867 types that appeared only in the general corpus and used the normalized frequencies of the remaining words and their frequency ratios in the two corpora to rank and group them. Using a frequency ratio of 300 (i.e., an occurrence of 300 times or more in the anatomy corpus than in the general corpus) as the cutoff, Chung identified 4,598 technical words in anatomy.

As an example of the more recent studies that used more sophisticated statistical measures, Periñán-Pascual (2015) proposed and tested an SRC “composite measure” formula that combines three statistical measures: “salience”, “relevance”, and “cohesion” in identifying technical terms. In simple words, “salience” assesses how prevalent a word is in a subject domain by measuring its frequency in a given text (known as term frequency or TF) in a specialized corpus compared with the inverse proportion of the word in the entire corpus (known as inverted document frequency or IDF). “Relevance” measures how likely the word may actually appear in the specialized domain by comparing its frequency in the specialized corpus with its frequency in a large comparison general corpus. “Cohesion” is used for assessing multiword terms as it measures the degree of association between the words in a multiword unit, i.e., how likely the words in the unit do co-occur. Periñán-Pascual tested the accuracy rate of the SRC composite measure formula coded in DEXTER (Discovering and Extracting TERminology) against the accuracy rate of the S, R, and C individual measures in two experiments. The first experiment involved an English electronics corpus of 520,383 tokens and the second one used a Spanish medical corpus of 197,812 tokens. The results from both experiments showed that the accuracy rate of the SRC composite measure was much higher than that of each of the single measures.

Finally, let us turn to studies that have combined judgment-based methods with corpus-based methods, a practice found in several recent studies (Ha & Hyland, 2017; Kwary, 2011; Peruzzo, 2014; Watson Todd, 2017). Of these studies, Ha and Hyland (2017), Kwary (2011), and Watson Todd (2017) were similar in that they all employed a corpus-based keyword analysis and a judgment rating. Because of their similarity, we look only at Ha and Hyland (2017). In examining technical vocabulary in finance, the two authors employed arguably the most sophisticated rating system so far labeled “Technicality Analysis Model” (TAM), which classifies words into five levels of technicality: “least technical” (words with no
Technical Vocabulary

specialized meaning in finance, e.g., comparable), “slightly technical” (words that have specialized finance meaning that is related to its general meanings, e.g., capital), “moderately technical” (words that have a specialized sense that may be remotely related or not related to its general senses, e.g., exposure), “very technical”, words that have “a specialised sense that is not related to any of its general senses”, e.g., facility used in the sense of a money borrowing arrangement), and “most technical” (words that have only a special sense without a general meaning, e.g., escrow) (Ha & Hyland, 2017, pp. 44–46). To trial this system, they first conducted a keyword analysis using a 6.7-million-word finance corpus with texts from four finance sectors and the academic subcorpus of the BNC Baby as a comparison corpus. It generated a list of 837 keywords that were each “specific to one or two financial sectors” (Ha & Hyland, 2017, p. 40).

Then, the two authors and several graduate students individually rated these keywords by using the TAM and checking the meanings of the words in a general English dictionary and a business English dictionary and, when necessary, reading concordance lines involving the words using WordSmith tools and WordBanks. Because they followed a rigorous level-by-level procedure to decide the technical level of a word by starting from the least to the most technical level, they attained a 95% inter-rater reliability on 100 words randomly from the 837 keywords. Based on their TAM rating, of the 837 words, 672 (82.29%) are least technical, 42 (5.09%) are slightly technical, 88 (10.51%) are moderately technical, 26 (3.11%) are very technical, and 9 (1.08%) are most technical (Ha & Hyland, 2017, pp. 44–45). Based on their finding “that a word specific to a financial sector is not necessarily technical”, Ha and Hyland (2017, p. 44) argue that “technicality and specificity are distinct concepts”. However, it is very important to note that while the use of a sophisticated rating scale helps make the identification decisions more valid and reliable, the decisions so made are still subjective and susceptible to errors as admitted by Ha and Hyland (2017, p. 45). Finally, a brief mention of Peruzzo (2014) is in order because her study did not use a rating scale in her combined method, but she incorporated the use of a semi-automatic term extraction with the use of “an event template” she developed based on a frame-based terminology model with the assistance of experts. The template covers all aspects related to a crime event, such as victim, offender, and harm.

The Role and Importance of Technical Vocabulary in Language Use and Learning

Corpus studies on vocabulary use in specialized texts have shown that a significant portion (between 20% and 30%) of the vocabulary in such writing is composed of technical vocabulary (Nation, 2013). Besides being shown in the studies already mentioned earlier (e.g., Chung & Nation, 2004), such a finding has also been reported in many other studies, including Lei and Liu’s (2016) study on the vocabulary use in a medical research article corpus and a medical textbook corpus. The results in Lei and Liu (2016, p. 49) reveal that medicine-specific vocabulary (including both high-frequency words with special medical meanings and low-frequency technical medical vocabulary) accounts for 31.75% of the tokens in the medical research article corpus and 30.44% in the medical textbook corpus. Given such a high percentage of technical words in specialized language, such as EAP and ESP, it is necessary for learners of such language to grasp these words, considering especially that research has shown that to fully understand a text, particularly an academic text, a reader needs to know about 98% of the running words in a text (Hu & Nation, 2000; Schmitt, Jiang, & Grabe, 2011).
Furthermore, research has also demonstrated that knowing technical vocabulary is indispensable for developing subject knowledge. For example, Woodward-Kron’s (2008) longitudinal study of undergraduate education majors’ learning and use of subject knowledge in education reveals the high importance of understanding technical vocabulary. The subject knowledge Woodward examined was related to child development and theories about thinking and learning in children. The major source of data were six students’ writing assignments on the aforementioned topics. Via a close analysis of the students’ writing, including their definitions and use of technical terms, such as egocentric, scaffolding, ZPD (Zone of Proximal Development), and Triarchic theory of intelligence, Woodward finds that understanding technical terms is imperative in students’ learning of specialist knowledge. She makes the point clear by stating that “learning specialist knowledge in pre-service teacher education involves adopting technical terms as well as coming to terms with the abstract dimension of the discourse” (Woodward-Kron, 2008, p. 234).

A good grasp of technical vocabulary is also essential in professional communication (e.g., Knoch, 2014). In a study on ESL aviation test takers’ technical English skills, Knoch (2014) had ten experienced native-English-speaking pilots listen and assess the recorded performance (speech samples) of nine non-native aviation test takers. The assessment included a writing and speaking component (a written questionnaire and a group oral interview). The results indicate that the understanding and use of technical vocabulary was considered by the rating pilots to be extremely important in aviation communication, as can be evidenced by the following quotes from two of the pilots in their assessment of the performance of the test takers being evaluated: “they [test takers] understood what they were talking about . . . and all the relevant aviation type terms were there [comment by #8 pilot-rater; underline added]. . . . I could understand him and he put enough [technical] information out to give me confidence . . . [comment by #3 pilot-rater]” (Knoch, 2014, p. 84).

Finally, there are two additional reasons that technical vocabulary deserves special attention in language learning. First, L2 learners often experience great difficulty in grasping such vocabulary due to, among other reasons, the low overall frequency of many technical words in general language and the specialized meanings of those polysemous technical words with a high-frequency use in nontechnical senses (Coxhead, Demecheleer, & McLaughlin, 2016; Watson Todd, 2017; Woodward-Kron, 2008). The latter technical words are especially challenging because learners often find their technical meanings opaque (Watson Todd, 2017). The other additional reason that technical vocabulary deserves special attention is the fact that with rapid advancements in various fields and technology in general, new technical words emerge constantly. In other words, learners and professionals alike have to learn technical words continuously. In fact, the study of technical terms is so important that “[t]erminology is [now] part of the programs of several university degrees and postgraduate courses (Translating and Interpreting, Applied Languages, Information Science)” (Alcina, 2009/2011, p. 1).

**Strategies and Methods for Learning and Teaching Technical Vocabulary**

Historically, some language educators (e.g., Cowan, 1974) did not think that it was a language teacher’s job and/or it was within his/her ability to teach technical vocabulary due to their lack of knowledge in the technical subject. However, today, with better understandings of technical vocabulary and its importance as well as more effective methods to teach such words, most experts and teachers believe that specialized words can and “should be taught and studied in a variety of complementary ways” and language teachers “may be able to
make a useful contribution" in this regard (Nation, 2013, pp. 32, 305; see also Fernández, Flórez de la Colina, & Peters, 2009). Through proper training and/or self-study, language teachers may be able to develop adequate knowledge in a technical field and can learn strategies and methods to help learners more effectively acquire technical vocabulary (Alcina, 2009/2011). In fact, there has been an increased interest in the teaching of technical vocabulary as evidenced by quite a few recent studies that explore effective ways for teaching technical vocabulary (Ardasheva & Tretter, 2017; Fernández et al., 2009; Gablasova, 2015) and a special issue of the journal Terminology devoted to the topic (Alcina, 2009/2011).

Based on recent research (Alcina, 2009; Fernández et al., 2009), the most useful strategy in teaching technical vocabulary is to engage learners in active meaningful learning activities, such as having learners identify technical terms and build term banks in their field on their own. The teacher only serves as a facilitator. Meaningful learning also calls for studying technical words in context, rather than in isolation (although the latter can be helpful in some ways and with some technical words). This is because contextual information may often enable learners to figure out and better understand the meanings of technical words (Nation, 2013). Also, studying technical words in context enables learners to learn the typical collocates of such words, which is very important because, like words in general, technical words often have their typical collocates (Nazar, 2016). Technical corpora and online technical data sources including terminology banks/bases have been found especially useful for context-based learning of the meanings and typical collocates of technical words (Alcina, 2009; Fernández et al., 2009). These tools can provide learners with the opportunity and resources to identify and compile their own technical word lists or banks (Fernández et al., 2009).

As another example of active learning, it will be very effective to explore and understand the connections between the technical meanings of polysemous technical words and their core meanings, connections that often exist in this group of technical words (Nation, 2013). For instance, the medical stoppage/cessation meaning of the word *arrest* in *cardiac arrest* and *respiratory arrest* is a metaphorical extension of its more commonly known meaning of seizing and putting a person in custody because when a person is arrested, his/her normal life activities come to a stop or are stopped. Teachers can have students examine and compare corpus examples of the two uses of such a word and then discuss and uncover the semantic connection between the two meanings. In the case of *arrest*, such a comparison should enable learners to see how the medical meaning of the word is a metaphorical extension of its more commonly known meaning. In fact, research on cognitive linguistics theory-inspired teaching has shown that a focus on the metaphorical extensions in the use of words can make vocabulary teaching significantly more effective (Boers & Demecheleer, 1998; Tyler, 2012). In their respective empirical studies on the teaching of prepositions, Boers and Demecheleer (1998) and Tyler (2012) found that exploring the metaphorical extensions in the use of prepositions significantly enhanced the participating students’ correct understanding and use of the prepositions they were learning. Similar positive results should be expected for using such an approach in teaching the meanings of many technical words.

There are also some tested or proven useful practices for learning and teaching technical terms. For example, the teaching of affixes of terms of Latin or Greek origin, such as the teaching of the prefixes *inter* vs. *intra* in terms like *internet* and *intranet* and teaching of the suffixes *cide* and *logy* in *herbicide* and *physiology* (Brieger & Pohl, 2002). One other long-time useful practice is the use of technical dictionaries and now technical word lists thanks to the recent publication of many new rigorously developed and pedagogically focused technical word lists in a variety of subjects, including Lei and Liu’s (2016) in medicine and Watson Todd’s (2017) in engineering. However, with limited study time, learners often need
to prioritize words on a list and focus on the most useful items, rather than covering all the items (Watson Todd, 2017). Another reason for doing so is that research has shown that a “deeper focus on fewer words” is much more effective than a general coverage of more words (Ardasheva & Tretter, 2017, p. 256).

Finally, recent research (Ardasheva & Tretter, 2017; Fernández et al., 2009; Gablasova, 2015) has shown the importance of using learner L1 or using a bilingual approach in L2 learning of technical vocabulary. Gablasova (2015) compared two groups of Slovak students learning technical terms: one group was instructed in their L1 and the other was taught in English. The results revealed a clear advantage for the L1 instructed group, as they made fewer errors and showed more complete understanding of the target words on the post-instruction test. The findings seem to suggest that when possible, in L2 learning of technical vocabulary, students may benefit from learning and checking the meanings of the target technical words in their L1.

### Future Directions

It can be seen in the preceding discussion that while scholars may continue debating on what technical vocabulary is and exploring its role in language, future work on technical vocabulary will likely focus mostly on how to accurately and efficiently identify technical words and how to effectively learn and teach such words. Next, we discuss some likely future developments related to the latter two issues.

First, regarding the issue of technical vocabulary identification, researchers will continue their effort to develop and find more rigorous, sophisticated, and effective methods of detection and extraction by applying new methods and technology developed in computational and corpus linguistics. There will also be more efforts to incorporate or combine computerized quantitative methods with new theories and tools emerging from qualitative research work, as has been done by Peruzzo (2014). In other words, there will likely be more use of combined methods to identify technical vocabulary. Furthermore, there may be continued work on identifying bilingual technical words using parallel corpora as was done by Macken, Lefever, and Hoste (2013). Also, there will likely be more effort to identify non-nominal technical words, especially technical verbs because, as Faber and L’Homme (2014) pointed out in their study on verbs in technical texts, verbs frame technical concepts (nouns) and are indispensable in our understanding of technical terms.

Concerning the learning and teaching of technical words, first, there surely will be continued endeavor to find more effective teaching strategies and methods, especially those that make use of new technology, corpora, and/or learners’ L1. Second, there will be more work in the development of pedagogy-oriented, discipline-specific technical word lists, including sub-discipline lists. This is because research (e.g., Bertels & Speelman, 2014; Grabowski, 2015) has shown that vocabulary use varies significantly even across different sub-field corpora of the same discipline (machining and pharmaceutical respectively in the two cited studies). Also, as bilingual lists have been found particularly useful for L2 learners (Fernández et al., 2009), many of the new word lists will likely be bilingual.

### Further Reading


This special issue consists of seven articles that introduce new theories and practices in the teaching and learning of terminology or technical terms. It covers the learning of technical terms in
several disciplines or areas of study, including architect and construction, law, and translation studies.


This recently published book provides a comprehensive coverage of research on the identification and categorization of specialized vocabulary (i.e., academic and technical vocabulary) for ESP as well as the value of specialized vocabulary lists for learning. Particularly worth mentioning is that it has a chapter (Chapter 8) devoted to technical vocabulary in the trades. While technical vocabulary abounds in the trades area, the topic has not received much attention until recently.

**Related Topics**

Academic vocabulary, high-, mid-, and low-frequency words, measuring depth of vocabulary knowledge, word lists

**Note**

1 This same rating scale was also employed in identifying technical words in the same anatomy text in Chung (2003), but it was included only for the purpose of evaluating a corpus comparison approach, the focus of this Chung study. In contrast, the use of the scale was the main focus of the Chung and Nation (2003) study. This is the reason why Chung and Nation (2003) is cited in this chapter as an example of a judgment-based method using a scale while Chung (2003) is given as an example of corpus-based method.

**References**


