ARTÍCULOS

Círculo de Lingüística Aplicada a la Comunicación
ISSN: 1576-4737

EDICIONES COMPLUTENSE

https://dx.doi.org/10.5209/clac.78300

How much and what kind of vocabulary do marine engineers need for adequate comprehension of ship instruction books and manuals?

Zorica Đurović[1]; Milica Vuković-Stamatović[2]; Miroslav Vukičević[3]

Received: 18 September 2019 / Accepted: 30 January 2021

Abstract. Considering the importance of adequate understanding of instruction books and manuals on board vessels all over the world, as well as the challenges it imposes to the English language teachers and course designers, this paper aims to answer important research questions in relation to the quantity and type of vocabulary required for their adequate reading comprehension. In this study we use the method of Lexical Frequency Profiling and the software developed by Anthony Laurence – AntWordProfiler 1.4.0w. The corpus is comprised of 1,769,821 running words obtained from instruction books and manuals of various ship and machinery types. The results of this study point to the high technicality and lexical demand of the corpus, which calls for a highly technical English courses' design and further research in marine engineering (English) vocabulary. Additionally, the research findings point to the need of creating a marine engineering-specific word list.
Keywords: vocabulary; lexical profiling; English for Marine Engineering Purposes (EMEP); instruction books and manuals.

Cómo citar: Đurović, Z., Vuković-Stamatović, M., Vukičević, M. (2021). How much and what kind of vocabulary do marine engineers need for adequate comprehension of ship instruction books and manuals?. *Círculo de Lingüística Aplicada a la Comunicación* 88, 123-133, https://dx.doi.org/10.5209/clac.78300

1. Introduction

Technical manuals are an essential part of shipboard equipment and are of vital importance to ship operators. Although pieces of equipment may originate from all over the world, virtually all of them have an English version of the technical manual, which is commonly the most widely used and referred to. In addition to more or less simple everyday conversation and dealing with uniform forms and documents, everyday job requirements of marine engineers also include having to read and use numerous and rather demanding instruction books and manuals related to various ship's systems, both in terms of technical knowledge and English vocabulary. This is of particular significance since instruction books are especially consulted in case of errors, failures, breakdowns or emergency, when consulting an instruction book or manual should be as quick and efficient as possible. Moreover, reading technical manuals is also the main means of detailed familiarisation with particular ship's systems once a marine engineer signs on to a new vessel. Bearing all this in mind, it is crucially important that marine engineers who are non-native to English are able to read them at an adequate level of reading comprehension without outside assistance. As knowledge vocabulary is generally considered a strong predictor of reading comprehension, in this paper we set out to explore the lexical profile of this genre, with the goal to determine how much and what type of vocabulary prospective and practitioning marine engineers need for reading ship instruction books and manuals.

[1] University of Montenegro, Faculty of Maritime Studies Kotor, E-mail: zoricag@ucg.ac.me ORCID: 0000-0002-8668-0005
[2] University of Montenegro, Faculty of Philology, E-mail: vmilica@ucg.ac.me ORCID: 0000-0001-5497-1361
[3] University of Montenegro, Faculty of Maritime Studies Kotor, E-mail: miroslav.v@ucg.ac.me ORCID: 0000-0002-9930-5650

2. Theoretical background

In the theoretical section of the paper, we briefly present English for Maritime and English for Marine Engineering Purposes as ESP (English for Specific Purposes) fields, and then dwell on the relation between vocabulary, as the meaning-carrier constituent of language (Prichard, 2011: 3) and reading comprehension, as established in the literature. The second part of this review deals with word lists, with a special focus on engineering word lists produced to date, as the most relevant lists to the study at hand.

2.1. English for Maritime Purposes and English for Marine Engineering Purposes

In the effort to define English for Specific Purposes (ESP), linguists agree on one – its purpose is to serve the specific language needs of its learners (Gollin-Kies, Hall and Moore; 2015, Bhatia, 1993, Charles, 1996), generally related to a certain field or discipline they deal with. In view of that, their requirements can be related to spoken or written texts or the combination of the two, whereas their relevant literature may be of secondary or tertiary level of education, including various genres such as textbooks, academic articles, spoken trials, etc.

Among ESP fields, English for Maritime Purposes (EMP) poses a special challenge to both English learners and teachers. In maritime industry, English is the leading language of communication of all the states and navies of the world. Just as the shipping industry presents a multidisciplinary and interdisciplinary scientific and professional field, comprising a vast area of human activities, so too maritime English represents a language of many narrower scientific disciplines and professional settings, such as marine engineering, navigation, marine telecommunications, shipbuilding, maritime law, maritime insurance, maritime economics, maritime ecology, maritime medicine, oceanography and many others (Đurović, 2019, 2014; Dževerdanović, 2008; Bocanegra Valle, 2013).

In this paper, the focus will be on English for Marine Engineering Purposes (EMES), as a specific ESP field, as we intend to explore the vocabulary profile of one marine engineering genre – that of ship instruction books and manuals.

Considering the challenges and great responsibility of their everyday performance on-board ships worldwide, the demands of marine engineering in terms of the English language skills are growing on a daily basis. Bearing in mind the fact that English for Maritime Purposes in general aims at accuracy, precision and conciseness, so that communication can be as efficient as possible, the learning burden in EMES, too, is rather placed on vocabulary and relaying a clear message than on the beauty of the language expression and its grammatical complexity. As a matter of fact, this engineering field has been shown to be one of the most demanding ones in terms of lexical burden (Hsu, 2014), which is why focusing on it in EMES instruction is of great importance.

3. Vocabulary level and reading comprehension

The relationship between reading comprehension and knowledge of vocabulary has been the topic of various studies. Laufer (1992) finds that by measuring vocabulary size, one can predict reading scores. In their 2010 study, Laufer and Ravenhorst-Kalovski find that almost two thirds of variance in reading scores could be accounted for by vocabulary. Of course, vocabulary is not the only factor contributing to reading comprehension, but it seems to be the most powerful one (Nagy, 1988).

How much vocabulary is needed for reading comprehension? Two answers prevail in the literature – the first one refers to a 95%-lexical threshold, which Laufer (1989) found necessary for adequate reading. The other was determined by Nation (2006), who argued that as much as 98% would be necessary for the "ideal" coverage. The remaining percentages of unknown words in a text (2%-5%), readers are expected to be able to guess from the context or check in dictionaries, without outside assistance. If their lexical thresholds, however, are below the set levels, reading becomes increasingly difficult and their comprehension is generally limited.

This type of research has many practical implications for ESL instruction – for instance, researchers can calculate how many words are needed to reach the set thresholds in particular texts and thus determine how lexically demanding these texts are. Additionally, using vocabulary size tests, such as those freely available on Nation's website (available at: https://www.victoria.ac.nz/lals/about/staff/paul-nation), teachers can determine how many words their learners know and compare these results with how many words are needed for adequate and ideal reading comprehension in particular types of texts.

Bearing in mind a number of facilitating factors present in the situation in which a marine engineer reads instruction books and manuals, such as their technical knowledge of the field and ship's systems, additional training courses, as well as the presence of pictures and diagrams in this particular genre, which also contribute to comprehension, for the purposes of this particular study we will settle for the lower threshold cited in the literature, that of 95%, which is needed for adequate reading comprehension.

4. Word lists

Applied and corpus linguistics have shown an increased interest and effort in selecting words and phrases to which instruction should be devoted in foreign language courses, especially professional and technical English language courses. These efforts have given rise to the creation of various word lists.

4.1. General Service List(s)

Despite the fact that several frequency-based lists of general English vocabulary have been created to date, one of the most influential and widely used is still West's General Service List (GSL), produced in 1953. It contains 2,000 word families (the headword with all its inflected and derived forms) which are considered to be in wide and general use, i.e. regardless of the specific purpose of the English language use (Chung and Nation, 2003: 104). The list was extracted from a corpus of 5 million words, bearing in mind the needs of learners of English as a Second/Foreign Language (ESL/EFL). Around 165 word families are functional words, such as *a*, *some*, *because* and *to*, and the rest are content words such as nouns, verbs, adjectives and adverbs (Coxhead, 2018: 22). This list is now seen as somewhat outdated, bearing in mind that the corpus it was derived from is about 70 years old. In addition, it was extracted manually from a pre-computer era corpus, considered small by today's standards for the corpora used for producing general word lists, and another problem is that not a clear definition of what makes a 'word' was employed (Browne, 2013: 14).

Two recent general word lists which feature prominently in the literature are both called the New General Service List (NGSL). One was created by Brezina and Gablasova (2013) and the other by Browne, Culligan and Phillips (2013). Both were formed from vast corpora – the former from a 12-billion-word corpus and the latter from a 273-million-word corpus (a subsection of the Cambridge English Corpus), and they both report somewhat better coverages than the GSL. The corpora used for making them, although vast, are biased towards written texts and the resulting lists are therefore more reflective of written English.

A special contribution in the creation of high-frequency word lists was given by Paul Nation. He first produced a set of 14 lists (Nation, 2004), each comprising 1,000 word families, from the master list of the British National Corpus (BNC), a vast reference corpus made up of 100 million words, consisting of 90% written and 10% spoken text (available at http://www.comp.lancs.ac.uk/ucrel/bncfreq/flists.html). In order to enhance the universality of his lists, Nation later combined the BNC with the Corpus of Contemporary American English (COCA) (available at: http://corpus.byu.edu/coca), including the following genres/registers: spoken, novels, popular magazines, newspapers and magazines. Based on this immense and more balanced 450-million-word combined corpus (BNC/COCA), Nation (2012) then made a new set of 25 lists, again each comprising 1,000 word families, which are more widely applicable. In addition, he produced four more lists – proper names, abbreviations, marginal words (letters of the alphabet, swear words, exclamations, etc.), and transparent compounds (those written without a hyphen). These lists provide good basic vocabulary for various types of texts and also for identification of specific types of vocabulary (Nation and Hwang, 1995; Brezina and Gablasova, 2013). As for all the lists, some criticism has been directed at Nation's word list sets as well, including the fact that their purpose and target audience have not been clearly defined (Nation, 2016), as well as the imperfect composition of the reference corpora it was derived from (the same criticism stands for virtually all word lists).

GSL and both NGSL's provide just the most frequent 2,000 words, and so cannot provide information we need beyond this basic level. On the other hand, the BNC/COCA lists can provide insight into the vocabulary sizes needed, as well as the types of words making up a particular vocabulary (Nation, 2016: 142), which is why we have opted to use this set of words for the task at hand.

Apart from creating general-purpose word lists, researchers have also been focused on creating academic word lists, as well as various discipline-specific word lists.

4.2. Academic Word List

Once the English learners have mastered the first 2,000-3,000 general words, language learning should become more field-oriented, depending on the needs and goals of the learners (Nation, 2000: 302). The next step could be a more specialised one but still common to a series of various fields – such as students' learning of academic vocabulary, which is common to a wide variety of academic texts from various disciplines. For non-native speakers, intending to study and/or do research in English, this type of vocabulary is of great significance, especially at tertiary level of education. For them, academic vocabulary can be a bridge between general and technical vocabulary, which is why it is sometimes referred to as the generally useful scientific vocabulary (Barber, 1962), sub-technical vocabulary (Cowan, 1974; Yang, 1986; Anderson, 1980) or semi-technical vocabulary (Farrel, 1990).

The best known and most commonly applied academic word list to date has been the Academic Word List (AWL), built by Averil Coxhead (2000), from her academic corpus of 3.5 million words. The AWL comprises 570 word families and covers approximately 10% of words in academic texts outside the 2,000 most frequent words of English (as represented by West's GSL), which are common to a wide range of academic texts (Coxhead, 2000).

There have been several other attempts at making new academic word lists, from larger corpora and employing different methods, such as the Vocabulary Academic List by Gardner and Davies (2014), and New Academic Word List by Browne et al. (2013), as well as some more discipline-specific ones, such as the medical academic word list (Wang et al., 2008), the applied linguistics academic word list (Khani and Tazik, 2013), the environmental academic word list (Liu & Han, 2015), and others. However, Coxhead's AWL is considered very practical and it is still very widely used in both research and in practice. It is short enough to make a feasible learning target and at the same

time it provides a respectable coverage of the words used in various types of academic texts (Vuković-Stamatović, 2020). In lexical profiling of various texts, the AWL has been used very frequently, together with the GSL, on top of which it was built. Using these lists allows for a comparison of the results of the new studies with the previous ones. Therefore, in addition with the BNC/COCA lists, we will use the AWL and the GSL, to the said advantage.

In the next section, we present the remaining lists to be used for the lexical profiling of ship instruction books and manuals, undertaken in this study.

4.3. Engineering word lists

Among the discipline-specific word lists particularly relevant to our study are engineering word lists, four of which exist, to our best knowledge. No marine engineering word lists have been built to date and so these general engineering word lists are currently the most applicable to the genre we study in this paper.

The most notable list among these is Ward's Basic Engineering List (BEL) (2009), which was derived from a corpus of 271,000 words consisting of 25 undergraduate textbooks. The list contains 299 word types and covers 16.4% in the said corpus. Ward used the following method: he first excluded function words and then applied frequency and range criteria. His list was built "from scratch", which means that no general words were excluded, bearing in mind that it is intended for weaker learners, with very limited vocabulary.

On the other hand, Jin et al. (2013) made the Engineering Technology Word List (ETWL), by first excluding the GSL and the AWL vocabulary, as this list is intended for students who already know these words. In fact, it is a technical word list, containing 313 items with a 8.7%-coverage in the corpus it was derived from – two textbooks used in Malayan upper-secondary education. This list is intended for vocational secondary education, as opposed to Ward's list, which is intended for undergraduates.

Hsu (2014) studied the lexical profile of engineering undergraduate textbooks and found that 5,000 words are needed for the 95%-lexical threshold. In addition, she created the English Engineering Word List (EEWL) from a corpus of 100 engineering textbooks by first excluding the most frequent 2,000 words, following which she applied strict range and frequency criteria. The result is a list of 729 word families, covering 14.3% in the textbook corpus it was created from. Like Ward's list, this list is intended for undergraduates, but the more proficient ones, as they are expected to know high-frequency words.

Another available word list is Todd's (2017), called Opaque Engineering Word List (OEWL), containing 186 items. It was produced from 27 textbooks. By "opaque" words, Todd refers to the words that learners will likely have problems with – polysemous words, where the meaning required in the engineering field is not the usual meaning of those words. This is a list of words which teachers should especially focus on in teaching and it is not intended to be used for measuring coverages.

Bearing all this in mind, in this study we will use Ward's (2009) and Hsu's (2014) word lists, to check the coverage of general engineering vocabulary in the genre of marine engineering instruction books and manuals. Jin et al.'s list (2013) is not applicable bearing in mind that our target group are undergraduates and graduates, who are most likely to use these instruction books and manuals, whereas Todd's list is not intended for lexical profiling.

5. Research questions

In this study, we aim to answer the following research questions:

1. How much high-frequency general-purpose and academic vocabulary is used in marine engineering instruction books and manuals?
2. How much general engineering vocabulary is used in marine engineering instruction books and manuals?
3. How many words are needed to read marine engineering instruction books and manuals at an adequate level?

Answering these questions will help us establish the lexical profile of marine engineering instruction books and manuals, and derive important pedagogical implications for ESP instruction to users of this genre.

6. Corpus and methodology

In order to answer the research questions set out above, we will use the corpus and methodology detailed in the following text.

6.1. Corpus details

Besides hand and power tools which are used on-board ships on a daily basis, the most useful tools to marine engineers, once they sign on to a vessel, are instruction books and manuals. They are indispensable for familiarisation with the specific vessel's machinery, devices, plants and systems, especially with regard to their maintenance and

repairs. Therefore, adequate reading comprehension of these materials is essential in shipping industry, which is why instruction books and manuals are the focus of investigation here.

In addition to the fact that there are many types of vessels of today, resulting in their complex classifications, the array of instruction books and manuals themselves on-board each ship is vast. We invested our best efforts into making a rational selection of them, according to the frequency of ship types and taking into account their diversity and comprehensiveness. In the process of collecting and selecting the material, a number of active seafarers were consulted, primarily Chief Engineers. The final corpus comprises instruction books and manuals from a container ship, tanker ship, offshore supply vessel and passenger ship.

Following the expert advice, additional material was added to accommodate for the diversity and comprehensiveness criteria – for instance, we included manuals for various types of main engines, including those most common ones, but also some of the most modern engines, such as the dual propulsion (fuel oil and gas) and the electric-driven ones. Apart from the instruction books and manuals intended for maintenance, operation and repairs of main engines, we included the technical manuals for various machinery and systems estimated to be the most essential on specific types of ships, such as auxiliary generators, separators and purifiers, boilers, pumps, lubricating, cooling, ventilation and air conditioning systems, desalination plants, air compressors, steering gears, sewage treatment plants, incinerators, filters, alarms, transmitters, propellers, etc.

The selected corpus comprised hundreds of pages of the material in the electronic form, which was converted to plain text without images and diagrams, using AntFileConverter (Anthony, 2017). The corpus so obtained needed to be manually cleaned, by removing tables, figures, lists of parts, lists of proper names, references etc., and remedying some conversion, typesetting and typing errors, as is typically done in lexical profiling studies (Coxhead, 2000), bearing in mind that the corpus needs to be "as clean as possible" (Nation, 2016: 62). This painstaking process is needed for more accuracy of the results (Nation, 2016: 224), so that fewer items may result off-lists, and was undertaken by the authors to the best of their abilities, considering that we were dealing with a vast collection of materials. We may say that the corpus was sufficiently 'cleaned' for our purposes and that any residual errors cannot significantly affect our results. In addition, for the reasons of discretion and avoidance of commerciality, we omitted the names of brands and companies.

Using the above detailed procedure, we produced a corpus of ship instruction books and manuals (CSIBM), comprising 1,769,821 running words, which is made up of texts varying in length, depending on the size and complexity of the plant or system they refer to. Bearing in mind that this is a discipline- and genre-specific corpus, we may say that our corpus is of a substantial size, which should warrant the representativeness of our results and the generalisation of the conclusions we will make.

6.2. Methodology

The method used in this study is Lexical Frequency Profiling – LFP (Laufer and Nation, 1995), which is used to measure vocabulary level of texts. In this method, software is used to classify words based on word lists they belong to and calculate the coverages of those word lists in texts. For this purpose, we will use the software developed by Anthony Laurence – AntWordProfiler 1.4.0w (2014), which can lexically analyse texts against the loaded word lists.

The word lists used are those presented in the introductory section. To answer the first research question, we will use the General Service List (West, 1953) and the Academic Word List (Coxhead, 2000), as these lists are complementary (the AWL was built on top of the GSL). This will allow for a comparison with the previous studies on other genres and disciplines, as most studies have used these two lists for lexical profiling.

To answer the second research question, we will use, as suggested above, the Ward's (2009) and Hsu's (2014) engineering word lists. These are only available as headwords and for lexical profiling they have to be expanded into all-family-member form, which we will do using Cobb's programme called Familizer+Lemmatizer v2.0 (2018).

In order to answer the last research question, we will use Nation's BNC/COCA lists (2012), as the most complete set of general word lists available.

Pedagogical implications will be discussed in section 8.

7. Results

The results obtained through the above explained methodology will be presented in tables and discussed by answering the research questions posed above.

7.1. How much high-frequency general-purpose and academic vocabulary is used in marine engineering instruction books and manuals?

By analysing the finalised corpus against the General Service List (GSL) and the Academic Word List (AWL), we obtained the results regarding the frequency of general-service and academic words in the Corpus of Ship Instruction Books and Manuals – CSIBM (Table 1).

Table 1. Coverage of GSL and AWL in CSIBM.

Word lists	Tokens	Coverage (%)
GSL	1,263,472	71.39
AWL	142,761	8.07
-	363,588	20.54
Total	1,769,821	100

The coverage of general-purpose words in the corpus is 71.39%, which is, not surprisingly, below the expected coverage of 78%-92%, cited for 'various written texts' (Nation and Waring, 1997), although similar to the coverage of 70%-71.9% in various academic texts, as reported by Coxhead (2000). Basically, this means that, with just the knowledge of 2,000 most common English words, approximately every fifth word of a text of this genre will not be familiar, which would make reading and understanding very difficult and slow.

Not surprisingly, the corpus also features a somewhat lower coverage of academic vocabulary (8.07%), compared to the expected average of around 10% (Coxhead, 2000) in typical academic texts, such as research articles and textbooks (e.g. 10% in medicine (Cheng and Ge, 2007), 11.17% in applied linguistics (Vongpumivitch et al. 2009), 9.47% in pharmacology (Fraser, 2007), etc.). The level of academic vocabulary varies amongst academic genres themselves (Vuković-Stamatović, 2020) and, strictly speaking, the genre we study here is not considered as academic. However, although a non-academic genre, ship instruction books and manuals still contain a substantial amount of academic vocabulary compared to some other non-academic genres, such as newspapers, where the AWL's coverage is 4% (Nation, 2016). To a large degree, the AWL covers words belonging to formal register (Nation, 2016), which is why some coverage of it is expected in all types of formal texts. The coverage we found is substantial enough that we may conclude that learning the AWL during the studies or for the needs of a formal training after secondary education is not only useful for reading academic texts, such as textbooks, but that it will also be useful in a future career of a marine engineer serving on board, which is an important pedagogical implication.

The cumulative coverage of the GSL and the AWL (71.39% + 8.07% = 79.46%) is lower than the average of 86.1% found in academic texts (Nation, 2000: 27). These kinds of results were to be expected, since ship instruction books and manuals are highly-technical in nature and meet the specific needs of the technical 'community' of marine engineers, and having in mind that they are not overly narrative. These results point to the fact that a substantial percentage of the words used in ship technical manuals belongs to technical vocabulary, which is prominent and frequent in them.

To illustrate the kind of words which are represented by the said lists, we present an excerpt from our corpus, taken from a boiler instruction manual used on a container ship (Figure 1). The words marked red represent the first 1,000 words of the GSL, whereas those marked green are from the second 1,000 words of the GSL. The words marked blue belong to formal academic vocabulary, as represented by the AWL. The word marked black are outside the GSL and the AWL.

The working principle of the rotary cup burner is based on atomising by centrifugal force. The atomising cup is driven at high speed by an electric motor through belts. The oil is gently positioned at low pressure into the spinning cup in which it moves gradually forward forced by the centrifugal action of the cup until it is thrown off the cup rim as a very fine, uniform film. The high velocity primary air discharge around the cup strikes the oil film, breaks it up and converts it into a mist of fine particles which are injected into the combustion zone and burned.

The total consumption of air needed for a complete combustion is supplied by the combustion air fan which is mounted on top of the wind box. The primary air represents approximately of the total air required for the combustion. The remaining is represented by the secondary air and tertiary air.

The primary air is used not only to atomise the oil film, but also to direct the resulting cone of oil into the combustion chamber. For this purpose a swirl ring is fitted in the cup shroud which will guide the primary air in the same rotating direction as the rotary cup.

The combustion air passes a distribution ring before it reaches the combustion air opening which is provided with air vanes. Furthermore, a part of the combustion air is led through adjustable tertiary air gaps making a balanced air flow through swirler and combustion air opening.

Figure 1. Coverage of GSL and AWL in an excerpt from CSBIM.

7.2. How much general engineering vocabulary is used in marine engineering instruction books and manuals?

In this part of our study we used Ward's Basic Engineering English Word List (BEEWL), and Hsu's Engineering English Word List (EEWL), as indicated in the introductory section of the paper (4.3).

As explained earlier, Ward's target population were low-proficiency students, which is why he did not pre-exclude any number of high-frequency words. Due to this, we cannot use this list together with any general-purpose list. Table 2 presents the coverage of BEEWL in our corpus.

Table 2. Coverage of BEEWL in CSBIM.

Word list	Tokens	Coverage (%)
BEEWL (Ward)	293,377	13.53

Ward's list covered 16.4% in his engineering corpus and compared to that, our result is somewhat lower (13.53%). This was to be expected, as our corpus is more discipline-focused than the one Ward used, which comprised many engineering fields. In addition, the genre used for making that list was different from ours (BEEWL was based on textbooks). Still, Ward's list is rather short (it contains 299 types), and the coverage it makes is substantial relative to that. Thus, for low-proficiency students of marine engineering, Ward's list might be of value. However, without expanding their vocabulary knowledge further, prospective marine engineers will not even remotely be able to cope with the English texts of this genre.

We will use the same excerpt as in section 7.1 to illustrate which words are covered by Ward's list (Figure 2). The words from BEEWL are marked red.

The working principle of the rotary cup burner is based on atomising by centrifugal force. The atomising cup is driven at high speed by an electric motor through belts. The oil is gently positioned at low pressure into the spinning cup in which it moves gradually forward forced by the centrifugal action of the cup until it is thrown off the cup rim as a very fine, uniform film. The high velocity primary air discharge around the cup strikes the oil film, breaks it up and converts it into a mist of fine particles which are injected into the combustion zone and burned.

The total consumption of air needed for a complete combustion is supplied by the combustion air fan which is mounted on top of the wind box. The primary air represents approximately of the total air required for the combustion. The remaining is represented by the secondary air and tertiary air.

The primary air is used not only to atomise the oil film, but also to direct the resulting cone of oil into the combustion chamber. For this purpose a swirl ring is fitted in the cup shroud which will guide the primary air in the same rotating direction as the rotary cup.

The combustion air passes a distribution ring before it reaches the combustion air opening which is provided with air vanes. Furthermore, a part of the combustion air is led through adjustable tertiary air gaps making a balanced air flow through swirler and combustion air opening.

Figure 2. Coverage of BEEWL in an excerpt from CSBIM

Table 3 presents the coverage of Hsu's EEWL in our corpus, together with the coverage of the 2,000 most frequent words, as represented by the BNC/COCA lists, bearing in mind that Hsu's list was built on top of them and does not overlap with them.

Table 3. Coverage of EEWL in CSBIM.

Word lists	Tokens	Coverage (%)
BNC/COCA 2,000	1,307,381	73.87
EEWL (Hsu)	178,857	10.11
Total	1,486,238	83.98

Not surprisingly, again, the coverage of 14.3% in the original corpus was considerably higher than the one we obtained in our corpus (10.11%). As said earlier, Hsu's corpus comprised undergraduate textbooks from 20 engineering subfields, marine engineering being just one of them. So, in addition to the fact that Hsu's list is more general and made from a different genre, we should bear in mind Hsu's finding (2014) that marine engineering proved to be one of the most demanding engineering fields vocabulary-wise.

Figure 3 represents the same excerpt we used before, only this time the words marked red belong to BNC/COCA first 1,000 words, whereas those marked green belong to the second BNC/COCA list. The words marked blue belong to Hsu's EEWL.

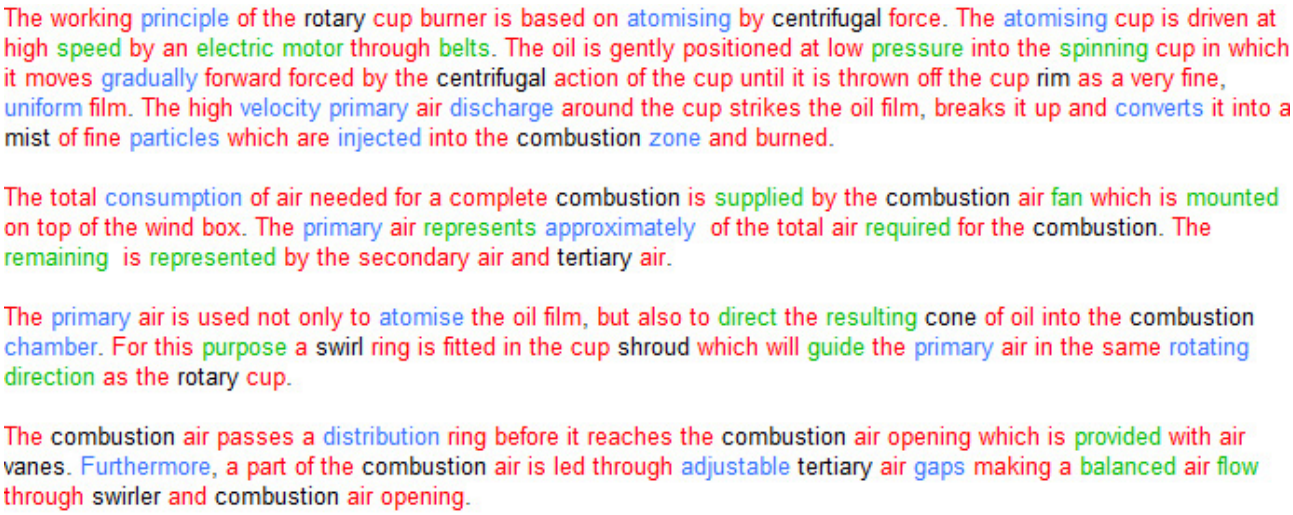

Figure 3. Coverage of BNC/COCA 2,000 and EEWL in an excerpt from CSBIM.

In the absence of a marine engineering word list, the lists presented above are all of some value and applicable to an extent, even though they all perform considerably lower in our corpus than in their original corpora. This calls for a creation of a marine engineering word list, which will be of more value to marine engineering students and practitioners.

7.3. How many words are needed to read marine engineering instruction books and manuals at an adequate level?

Finally, we come to the question of how much vocabulary is actually needed to adequately read and understand ship instruction books and manuals. Table 4 presents the coverage of Nation's BNC/COCA lists in our corpus. At each level, the knowledge of proper nouns, abbreviations and marginal words is assumed, as these are words which bear no or minimum learning load. Namely, proper names are generally easily recognised in a text; abbreviations are typically explained the first time they are used in a text; whereas marginal words contain the letters of the alphabet (typically also used as abbreviations), exclamations and swear words – the former are generally easily recognised, while the latter are not used in the genre we study.

Table 3. Coverage of BNC/COCA lists in CSBIM.

BNC/COCA word lists	Coverage (%)
2,000 + proper n., abbrev. and marginal words	78.12
3,000 + proper n., abbrev. and marginal words	86.34
4,000 + proper n., abbrev. and marginal words	89.19
5,000 + proper n., abbrev. and marginal words	91.53
6,000 + proper n., abbrev. and marginal words	92.42
7,000 + proper n., abbrev. and marginal words	93.31
8,000 + proper n., abbrev. and marginal words	93.88
9,000 + proper n., abbrev. and marginal words	94.27
10,000 + proper n., abbrev. and marginal words	94.62

BNC/COCA word lists	Coverage (%)
11,000 + proper n., abbrev. and marginal words	94.82
12,000 + proper n., abbrev. and marginal words	95.12
25,000 + proper n., abbrev. and marginal words	95.85

The results presented in Table 4 confirm that ship instruction books and manuals are lexically very demanding. The coverage of 95%, required for adequate reading comprehension (Laufer, 1989, 1992) is reached with the knowledge of 12,000 most frequent words of English. The ideal coverage, that of 98%, is not even reachable if only general vocabulary is considered.

Compared to the results obtained from marine engineering textbooks (Hsu, 2013), which is a corpus from the same discipline but of a different genre, where the threshold of 95% was reached at the level of 8,500 most common words, we can, once again, confirm the high lexical demand of our corpus. This does not come as a surprise considering the fact that textbooks are usually more narrative, which requires more high-frequency words, whereas instruction books and manuals are more technical and specific in nature. Further comparison with textbooks from other engineering fields points to more differences. For instance, Hsu (2014) finds that the same adequate reading threshold is reached at the level of 5,500 words for electrical and chemical engineering textbooks, and at the level of as few as 3,500 words for civil and mechanical engineering textbooks.

The first 4,000-5,000 words of English cover about 90% of the corpus and the additional thousands do not contribute much to a further substantial rise of the cumulative coverage, i.e. the additional coverage of 5% needed is reached only with some additional 7,000 words. Therefore, after mastering some level of general vocabulary, it would make sense to turn attention to the field-oriented vocabulary, where a much smaller selection of words could bring about a substantial coverage, as suggested by Nation (2,000). This would drastically reduce the learning load of prospective marine engineers.

For illustration purposes, we present the same excerpt with markings which show what level list particular words belong to. The legend is as follows: red – 1st 1,000 words, green – 2nd 1,000 words, blue – 3rd 1,000 words, pink – 4th 1,000 words, violet – 5th 1,000 words, ochre – 6th 1,000 words, brown – 7th 1,000 words, black – words outside these lists.

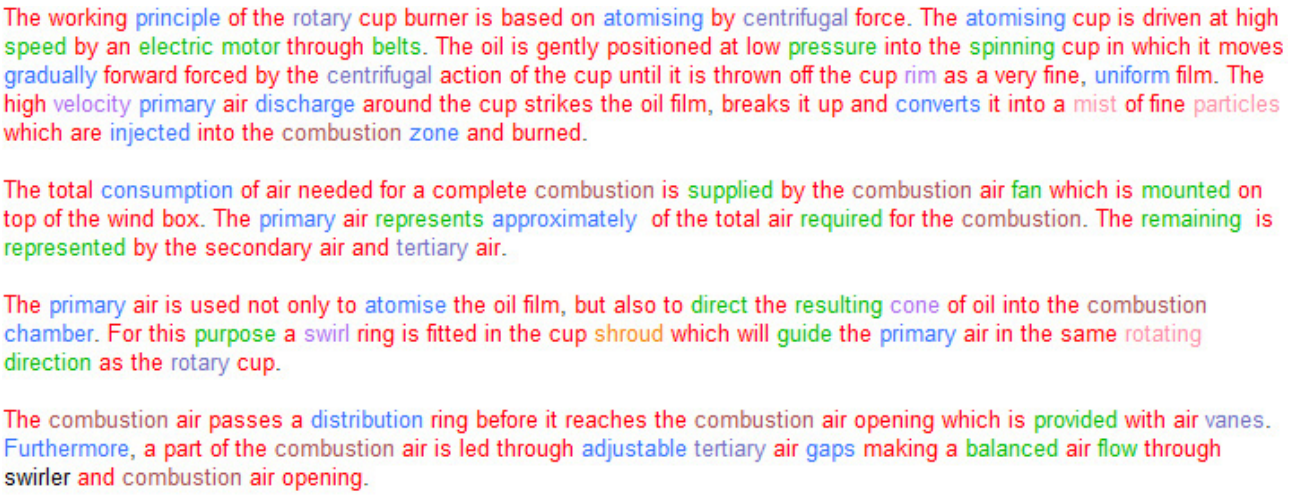

Figure 4. Coverage of BNC/COCA lists in an excerpt from CSBIM.

As can be seen in Figure 4, the words from the first three BNC/COCA word lists pervade the excerpt, whereas the other word lists are represented by just a few or only one item.

8. Pedagogical implications

The results of this paper have important pedagogical implications for ESP instruction of prospective marine engineers, particularly with a view to enable them to adequately read ship instruction books and manuals. Namely, the results point to the need to create a marine-engineering-specific word list, so as to reduce the learning load of the learners – without using discipline-specific word lists, these ESP learners would need to learn about 12,000 most common words of English, which is a daunting task. However, until such a list is created, learners and teachers may rely on some corpus-based solutions, i.e. alternative word lists, the value of which we have attempted to estimate

in this paper. We found that the existing general engineering lists are of some value – for low-proficiency students, there is Ward's list of some 300 words, which can provide a solid start to their English learning. For those who have mastered the first 2,000 words of English, Hsu's list is recommended. Still, both will not be enough to reach the level of adequate reading comprehension and will need to be supplemented by other word selections, including those not derived from corpora, in order to master the needed vocabulary, until the creation of a marine engineering word list.

9. Limitations of the study

As pointed out in the introduction, the knowledge of vocabulary is a strong predictor of reading comprehension but is definitely not the single predictor and factor having an impact on it. The knowledge of grammar would need to be adequate, to some extent, for adequate reading comprehension, and, on the other hand, some facilitating factors, which we briefly mentioned in the introduction, need also to be taken into account. Namely, being very familiar with the subject may compensate for lack of words and enable comprehension even below the set thresholds (Schmitt et al., 2015), i.e. a prospective or practising marine engineer may be very familiar with ship's systems and be in a position to understand ship instruction books and manuals even with some gaps in his/her vocabulary. In practice, we expect very few of them to be familiar with the most frequent 12,000 words of English, which means that they have found ways to learn specialised vocabulary. Another facilitating factor is that ship instruction books and manuals are abundant with images, diagrams and schemes, which certainly add to the overall readability of the texts. On the other hand, the learning of vocabulary from this field cannot be easily estimated or measured due to polysemous words (for example *average*, which, in Maritime English can also mean *loss* or *damage to maritime property*). Still, employing corpus-based solutions, as suggested here, can immensely contribute to their efforts.

10. Conclusion

In this paper we studied the lexical profile of ship instruction books and manuals with the aim to answer three research questions.

First, we determined that this genre contains somewhat less high-frequency general-purpose vocabulary than is the case with some other genres, which points to the fact that a substantial remainder of the uncovered words could in fact be technical words. Academic vocabulary was not as present as in academic texts, but its presence was still substantial enough to justify and recommend including English for Academic Purposes courses as part of education to be received by a marine engineer.

Secondly, we dealt with general engineering vocabulary in our corpus, as represented by two currently available word lists. The results point to a considerable presence of this vocabulary, too, and show that these lists have some value in the ESP instruction for prospective marine engineers. However, the lists proved insufficient, even when combined with lists of high-frequency general vocabulary.

Thirdly, we determined that the most common 12,000 words of English were needed to adequately read ship instruction books and manuals. This figure is beyond the reach of many learners, which, together with the previous findings (insufficient applicability of the existing engineering word lists), points to the need of creating a marine engineering-specific word list.

The results of this study point to the high technicality and lexical demand of ship instruction books and manuals, which calls for a highly technical English courses' design and further research in marine engineering (English) vocabulary.

References

Anderson, J. (1980). The lexical difficulties of English medical discourse for Egyptian students. *English for Specific Purposes* (Oregon State University), 37, 4.

Anthony, L. (2014). AntWordProfiler (Version 1.4.1) [computer software]. Tokyo: Waseda University. http://www.laurenceanthony.net/software (accessed 1 July 2019).

Anthony, L. (2017). AntFileConverter (Version 1.2.1) [computer software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software

Barber, C. L. (1962). Some measurable characteristics of modern scientific prose, in *Contributions to English Syntax and Philology*, 21-43. Goteburg: Acta Universitatis Gothoburgensis.

Bhatia, V. (1993). *Analysing Genre: Language Use in Professional Settings*. London: Longman

Bocanegra-Valle A. (2013). Maritime English, *The Encyclopedia of Applied Linguistics*, Blackwell Publishing Ltd., 3570-3583.

Brezina, V. & Gablasova, D. (2015). Is there a core general vocabulary. Introducing the new general service list, *Applied Linguistics*, 36 (1), pp. 1-22. https://doi.org/10.1093/applin/amt018

Browne, C. (2013). *The New General Service List: Celebrating 60 years of Vocabulary Learning*. JALT Publications: The Language Teacher.

Browne, C., Culligan, B. & J. Phillips. 2013. The new academic word list. http://www.negeneralservicelist.org/nawl-new-academic-word-list/ (accessed 1 April, 2018)
Chen, Q. & Ge, G.C. (2007). A corpus-based lexical study on frequency and distribution of Coxhead's AWL word families in medical research articles. *English for Specific Purposes* 26, 502-514. https://doi.org/10.1016/j.esp.2007.04.003
Chung, T. & Nation, I. S. P. (2003). Technical vocabulary in specialized texts, *Reading in a foreign language*, Vol. 15, No. 2.
Cobb, T. (2018). Familiser + Lemmatiser v.2.0. [computer software]. https://www.lextutor.ca/ Accessed 1 July 2019.
Cowan, J. R. (1974). Lexical and syntactic research for the design of EFL reading materials. *TESOL Quarterly*, 8(4), 389-400
Coxhead, A. (2000). A New Academic Word List. *TESOL* 34(2), 213-238. https://doi.org/10.2307/3587951 .
Coxhead, A. (2018). *Vocabulary and English for Specific Purposes research: Quantitative and qualitative perspectives*. London: Routledge.
Dževerdanović, M. (2008). *Discourse structures of maritime reports*, Master's thesis, Belgrade, Faculty of Philology.
Đurović, Z. (2019). Maritime and other discourse communities, The example of Charter Parties, 8th International Conference of Maritime Sciences Book of Proceedings.
Gollin-Kies, S., Hall, D. R. & Moore, S. H. (2015). *Language for Specific Purposes*, Basingstoke: Palgrave Macmillan
Gardner D. & Davies M. (2014). A new academic vocabulary list. *Applied Linguistics* 35(3), 305-327. https://doi.org/10.1093/applin/amt015
Grego, Z. (2014). *Discourse analysis of Charter Parties*, Master's thesis, Nikšić, Faculty of Philology
Farrell, P. (1990). Vocabulary in ESP: A lexical analysis of the English of electronics and a study of semi-technical vocabulary, *CLCS Occasional Paper* No. 25. Trinity College.
Fraser, S. (2007). Providing ESP learners with the vocabulary they need: Corpora and the creation of specialized word lists. *Hiroshima Studies in Language and Language Education* 10, 127-143.
Hsu, W. (2014). Measuring the vocabulary load of engineering textbooks for EFL undergraduates. *English for Specific Purposes* 33, 54-65. https://doi.org/10.1016/j.esp.2013.07.001
Jin, N. Y., Ling, L. Y., Tong, C. S., Sahiddan, N., Philip, A., Azmi, N. H. N., & Tarmizi, M. A. (2013). Development of the Engineering Technology Word List for vocational schools in Malaysia. *International Education Research*, 1(1), 43-49.
Khani R. & Tazik, K. (2013). Towards the development of an academic word list for applied linguistics research articles. *RELC Journal* 44(2), pp. 209-232. https://doi.org/10.1177/0033688213488432
Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In Christer Lauren & Marianne Nordman (eds.), *Special Language: From Humans Thinking to Thinking Machines*, 316-323. Clevedon: Multilingual Matters.
Laufer, B. (1992). How much lexis is necessary for reading comprehension? In Pierre Arnaud & Henri Béjoint (eds.), *Vocabulary and applied linguistics*, 126-132. London: Palgrave Macmillan. https://doi.org/10.1007/978-1-349-12396-4_12
Laufer, B & Nation I. S. P. (1995). Vocabulary size and use: lexical richness in L2 written production. *Applied linguistics* 16.3, 307-322. https://doi.org/10.1093/applin/16.3.307
Laufer, B. & Ravenhorst-Kalovski, Geke C. (2010). Lexical Threshold Revisited: Lexical Text Coverage, Learners' Vocabulary Size and Reading Comprehension. Reading in a foreign language 22(1). 15-30.
Liu, J. & Han, L. (2015). A corpus/based environmental academic word list building and its validity test. *English for Specific Purposes* 39: 1-11. https://doi.org/10.1016/j.esp.2015.03.001
Nagy, W. E. (1988). *Teaching vocabulary to improve reading comprehension*. Newark, DE: International Reading Association.
Nation, I. S. P. & Hwang K. (1995). Where would general service vocabulary stop and special purposes vocabulary begin? *System*, Vol. 23, 35-41. https://doi.org/10.1016/0346-251X(94)00050-G
Nation, I. S. P. & Waring, R. (1997). Vocabulary size, text coverage and word lists. *Vocabulary: description, acquisition and pedagogy,* 14, 6-19.
Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press
Nation, I. S. P. (2004). A study of the most frequent word families in the British National Corpus. *Vocabulary in a second language: Selection, acquisition, and testing* 10, 3-13. https://doi.org/10.1075/lllt.10.03nat
Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review* 63. 59-82.
Nation, I. S. P. (2012). The BNC/COCA word family lists. http://www.victoria.ac.nz/lals/about/staff/paul-nation (accessed 1 July 2019).
Nation, I. S. P. (2016). *Making and using word list for language learning and testing*. Amsterdam: John Benjamins.
Pritchard, B. (2011). Minimum (technical) vocabulary – some issues in Maritime English, *ESP Maritime English Journal*, Vol 1, No. 1.
Schmitt, N., T. Cobb, M. H. & Schmitt D. (2015). How much vocabulary is needed to use English? Replication of van Zeeland and Schmitt (2012), Nation (2006) and Cobb (2007). Language Teaching 50(2), pp. 212-226. https://doi.org/10.1017/S0261444815000075
Todd, R. W. (2017). An opaque engineering word list: Which words should a teacher focus on? *English for Specific Purposes* 45, 31-39. https://doi.org/10.1016/j.esp.2016.08.003 .
Vongpumivitch, V, Huang, J.Y. & Chang , Y. C. (2009). Frequency analysis of the words in the Academic Word List (AWL) and non-AWL content words in applied linguistics research papers. *English for Specific Purposes* 28, 33-41.
Vuković-Stamatović, M. (2020) Vocabulary complexity and reading and listening comprehension of various physics genres, *Corpus Linguistics and Linguistic Theory* 16(3), 487-514. DOI: https://doi.org/10.1515/cllt-2019-0022 .
Wang J., Liang, S. I. & Ge, G. C. (2008). Establishment of a medical word list. *English for Specific Purposes* 27, 442-458.
Ward, J. (2009). A basic engineering English word list for less proficient foundation engineering undergraduates. *English for Specific Purposes*, 28(3), 170-182.
West, M. (1953). *A General Service List of English Words*. London: Longman, Green and Co.
Yang, H. (1986). A new technique for identifying scientific/technical terms and describing science texts. *Literary and Linguistic Computing*, 1, 93-103.