AAA – Arbeiten aus Anglistik und Amerikanistik

Band 47 (2022) · Heft 1
Gunter Narr Verlag Tübingen
DOI 10.2357/AAA-2022-0005

Selecting ESP reading materials

Vocabulary suitability of science magazines for English for Science teaching and learning

Milica Vuković-Stamatović and Vesna Bratić

Bearing in mind that teachers often find themselves in a position where they have to produce their own teaching materials for English for Specific Purposes (ESP) classes, vocabulary profiling studies of certain genres may be of help in such situations. English for Science is an ESP field commonly taught around the world; however, despite this, the teaching resources for it are not as plentiful as the ESP teachers would like them to be. With this in mind, in this paper we study the vocabulary profile of science magazines, a genre that is generally written for non-expert audience and includes reports, news and opinions about science. We determine how complex the vocabulary of this genre is, using a corpus of approximately 230,000 running words, and define how many words are needed to reach the minimum reading comprehension level. We also determine how much high-frequency general, academic and scientific vocabulary this genre contains. Based on this, we draw conclusions on the target ESP audience these texts would be most useful for.

1. Introduction

Increasingly, English is considered as the *lingua franca* of science and university studies – it dominates the world's scientific, academic and technological communication (Gibson 2007; Tardy 2004; Dimova et al. 2015). Moreover, on account of the massive technological and scientific innovations introduced recently, the number of STEM[1] graduates has been on the

[1] Science, Technology, Engineering and Mathematics.

rise, whereas the demands for them have been increasing even faster[2]. In this context, English for Science (ES) has gained particular prominence, and along with English for Academic Purposes (EAP) courses (Parkinson 2013; Charles 2013), it is widely taught as part of university science curricula to non-native speakers of English.

As one of the branches of English for Specific Purposes (ESP), English for Science is taught separately from General English due to the fact that it has its own distinct methodology, a research agenda related to the disciplines that inform it, as well as a focus on practical outcomes (cf. Miller 2014). English for Science may subsume a number of more specific courses, such as courses for medical purposes, physical sciences, life science, etc., but it is sometimes taught as a whole, generally for groups of students with different majors. While some of the individual disciplines are well or at least modestly resourced with textbooks and other teaching and learning materials (e.g. English for Medicine), for many of them there are still insufficient up-to-date resources (e.g. for English for Mathematics or English for Physics), and that is also the case for English for Science taught as a whole. Moreover, the existing resources have primarily focused on prototypical scientific genres, such as research articles and lab reports, while a further incorporation and investigation of other genres related to science would be of value (Parkinson 2013).

In the light of the above, this study seeks to explore the vocabulary value of including the genre of science magazines in English for Science resources, a genre comprising reports, news and opinions on science, intended mostly for "lay persons", i.e. non-expert audience but also for scientists who wish to keep up-to-date with the advances from other scientific disciplines. Another feature of science magazine articles is that they are chiefly written by scientists themselves. This genre is, therefore, at least tangentially of value for real scientists and more so for science students, who still have not truly entered the world of science or settled on their majors. The content of science magazines is generally more interesting than that of scientific genres such as research articles, as it is written to both inform and entertain the scientific curiosity, and is richly illustrated with photos. The appeal of this genre could certainly recommend it for use in ESP resources. In addition, as English for Science teachers are generally disciplinary outsiders (they are not scientists) (Miller 2014), science magazines would certainly be more accessible to them contents-wise. These two arguments speak in favour of using science magazines in English for Science courses; however, before doing that, it is worth exploring just how valuable this genre is vocabulary-wise and which groups of English for Science students it would benefit the most.

[2] See, for instance, the 2018 report of the UK's Institute of Physics and the 2012 report of the US President's Council of Advisors on Science and Technology.

Bearing in mind the above, this paper examines the lexical profile of science magazines, i.e. the complexity of its vocabulary and the number of words needed to cross the comprehension thresholds as defined in the literature. A corpus of around 230,000 running words from this genre is also checked for the presence of general, academic and scientific vocabulary, as represented by the word lists of these types of vocabulary already established in the literature. Based on this, we draw pedagogy-related conclusions in terms of the value of this genre in English for Science courses, as well as the target audience that may benefit most from reading such texts.

2. Theoetical background

In this section we will discuss the importance of vocabulary knowledge for reading comprehension and go on to provide a brief review of some of the word lists produced for ESL and ESP purposes so far. We will also review the methodology used for lexical profiling that will be employed in this paper, as well as relate vocabulary sizes to various levels of ESL knowledge.

2.1. Vocabulary proficiency and reading comprehension

That vocabulary knowledge can predict the level of reading comprehension is a generally accepted notion (Laufer and Ravenhorst-Kalovski 2010). According to Nagy (1988), the knowledge of vocabulary is the single best predictor of a reader's level of understanding a text. Having this in mind, a number of scholars have tried to find out how many words a reader should be familiar with in order to understand certain types of texts at certain levels.

In her seminal study, Laufer (1989) determined that it is a 95%-vocabulary coverage threshold that would be needed to achieve a "reasonable" or "optimal" level of reading comprehension of a text, with the 5% remaining words assumed to be guessed from the context. Another study, that of Nation (2006), raised the threshold coverage, which should ensure an "optimum" reading of texts, achieved if a reader knows 98% of the words used in a text. Nation (2006) maintains that this 98%-threshold typically means the knowledge of approximately 9,000 word families. One word family comprises the root wood and all the words derived and inflected from it (for instance, *maintain, maintains, maintained, maintaining, maintenance, maintenances…*).

Both the cited thresholds seem quite high for non-native English speakers, for many of them virtually beyond reach. Nevertheless, some scholars have shown that the count of the words needed depends on the type of text and proven that a good selection of words for certain purposes may substantially reduce the number of the words needed. This insight inspired the birth of word lists, created for many different purposes, and subsequently the lexical profiling of various texts against these word lists.

2.2. Measuring vocabulary load and word lists

The Lexical Frequency Profiling (LFP) method, developed in 1995 by Laufer and Nation, is one of the widely used methods for quantifying the lexical complexity of a text and the receptive size of a learner's vocabulary. It relies on the following procedure: a corpus is loaded into a specialized programme, alongside one or more word lists (e.g. these lists could be those of the most frequent vocabulary, academic vocabulary, technical vocabulary, etc.). The programme then calculates the amount of coverage of each of the loaded lists in the corpus. The results can be compared to those for other corpora, which reveals the lexical richness and profile of a certain corpus when compared to others.

As one of the best-known frequency-based measures of vocabulary, LFP is very often used in ESL/EFL research and instruction. Even though this is not the only method for calculating lexical richness, the LFP method produces results that to a great extent match those obtained using the other methods (Lindqvist et al. 2013). The results obtained are quantitative, which contributes to their clarity and verifiability, but the method has still been criticized. The greatest fault found with it is its alleged bias to receptive knowledge of vocabulary. Also, when the lexical profile is "narrowed down" to just word frequencies, i.e. numbers, some "information" seems to be inevitably lost (Crossley et al. 2013). Nevertheless, this method has been widely used in the last twenty years, which corroborates its value (Morris and Cobb 2004; Read and Nation 2006; Douglas 2015, etc.).

There are a number of word lists available now but, in this paper, we will present only those that are applicable to our present study.

Word lists containing the most frequent general-purpose vocabulary are most typically used for the purpose of assisting General English instruction. Others are closely specialized for certain areas, i.e. specialized purposes, such as academic word lists for higher education students or ESP learners' lists. They are normally used as teaching and learning resources (Khani and Tazik 2013), but they are also useful as guidelines for developing textbooks and courses in the domain of EFL and ESP instruction (Wang et al. 2008; Jin et al. 2013).

The "pioneering" General Service List (GSL), developed by West in 1953, was influential for decades. This list was made before the invention of software and was taken out from a corpus manually. The GSL contains the most frequent 2,000 word families of English and was in wide use until quite recently, its updated replacements having been introduced only sixty years later. These GSLs are known as the NGSLs – the New General Service Lists (Brezina and Gablasova 2013, and Browne et al. 2013a) and they outdo the old GSL to a certain degree in modern corpora.

A very well-known word list is the Academic Word List (AWL) (Coxhead 2000). Coxhead generated it from a balanced academic corpus of 3.5 million words, after disregarding the words already contained in the GSL. With its 570 word families, the AWL generally accounts for about 10% of the words in most academic written corpora.

Coxhead and Hirsh (2007) went one step further and used the GSL and the AWL to develop the Science List (SL), from the science subsection of Coxhead's original academic corpus, after excluding the word families found in those two lists. The SL covers 3.79% of the words in their science corpus and it contains 388 word families. We could say that this is an academic science list – it is still academic as it was developed from an academic corpus, but it is more specialized towards science than the AWL.

Analogously to the AWL being built on the basis of the original 1953 GSL, the NAWL (New Academic Word List) was developed on top of the NGSL (Browne et al. 2013). These authors used a 288-million-word academic corpus to produce this list – this corpus is substantially larger than that used by Coxhead (2000), but it is not as balanced.

In the newest studies, both pairs of the lists are used (the GSL + the AWL, and the NGSL + the NAWL), the upgraded lists having the advantage of being more representative of modern authentic language and enabling a wider coverage in general, as they were extracted from larger corpora than the original lists. On the other hand, much of the data from the literature pre-dates the new lists and very often it comes in quite useful to compare the new results with the results from the earlier research using the earlier sets of word lists.

Another set of word lists that is widely used in lexical profiling is Nation's set (2012). A giant 450-million word corpus combining the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA), was used to develop this set of frequency-based word lists. It comprises 25 word lists, each of them having 1,000 word families. This set also contains four additional word lists with proper names, marginal words, non-hyphenated compounds and abbreviations. It is normally used to determine the lexical richness of a corpus and it is the largest word list set to date.

2.3. Vocabulary size and the Common European Framework of Reference (CEFR)

Milton (2010) attempts at correlating vocabulary sizes and the different levels of knowing a foreign language according to the Common European Framework of Reference (CEFR) (Council of Europe 2001). The author finds that to progress from A levels into B levels, according to the CEFR, students need around 3,000 words. For the B2 level, which presupposes reading with a substantial degree of independence, it is estimated that up to 5,000 words are needed. "For the highest level of fluency and understanding," Milton (2010: 202) concludes[3], the knowledge of 98% of the words used in a text would be necessary, which typically corresponds to 8,000–9,000 words for most genres, as Nation calculates (2013). The most advanced L2 speakers, such as those undergoing doctoral studies taught in English, will know about 9,000 word families of English (Nation 2013: 26).

In this study, we will calculate how many words are needed for reading science magazines and what level of the English language knowledge, according to CEFR, is required for this effort.

3. Corpus and methodology

The corpus employed in this study consists of 7 full issues of two reputable and very popular science magazines published in English – 4 issues of *New Scientist*, as well as 3 issues of *Scientific American*, published between February and June 2018. Both these subsets of our corpus contain a similar number of total running words (Table 1). *New Scientist* is based in London, whereas *Scientific American* is based in the USA. Both the magazines publish articles on a wide range of scientific topics and are sold globally.

The *pdf* files were converted into plain text files (.txt) by means of the programme *AntFileConverter 1.2.1* (Anthony 2017). The corpus so obtained was "cleaned", meaning *that* formulas and tables were removed, and that the conversion errors were addressed (most resulting from the words split at the end of a line, which were made whole again manually). This procedure generated a corpus whose details are presented below:

Science magazines	No. of issues	No. of tokens
New Scientist	4	118,400
Scientific American	3	108,068

Table 1. Details of the corpus

We then used *AntWordProfiler 1.4.0w* (Anthony 2014), a vocabulary profiling programme, which allows for corpora to be compared against the loaded word lists. We calculated the coverage of each of the loaded lists (the lists are those from the theoretical review), the results of which were compared against the other available data from the literature.

[3] The author, however, does not make a distinction between the C1 and the C2 levels, respectively.

4. Results and analysis

We first used the *NGSL 1.01* (Browne et al. 2013a), containing 2,818 lemmas (corresponding to 2,368 word families), as well as the *NAWL 1.0* (Browne et al. 2013b), containing 960 lemmas. As explained before, the NAWL was built on top of the NGSL, and so the two word lists are complementary (i.e. there are no overlaps) and can be used in conjunction. Their coverages in the corpus are presented in Table 2:

Word list	Token %	Cumulative %	Word list
NGSL	80.2	80.2	NGSL
NAWL	3.09	83.29	NAWL

Table 2. Coverage of the NGSL and the NAWL in the Science Magazine Corpus

The NGSL and the NAWL reach a combined average of 83.29% in our Science Magazine Corpus, leaving another 16.71%, or fewer than 3 in every 20 words (typically, two lines of a text), uncovered in the corpus. For an unassisted reading, students would need to know two more words in every 20 words of a text, if they were to rely solely on the NGSL and the NAWL.

These two word lists are recent, which means that not much research has been conducted using them and that there are not many data available regarding their coverage in various corpora, and so these results cannot be compared with those for other corpora (but might with some future findings). Bearing this in mind, we additionally calculated the coverage of the two older corresponding lists – the GSL and the AWL, for which there are extensive data available in the literature. As explained earlier, these two were also built in conjunction, i.e. on top of each other, and on top of the two, the Science List (SL) was also derived. The results presented in Table 3 include the coverages of these three complementary word lists:

Word list	Token %	Cumulative %	Word list	Token %
GSL	75.08	75.08	GSL	75.08
AWL	6.62	81.7	AWL	6.62

Table 3. Coverage of the GSL, the AWL and the SL in the Science Magazine Corpus

The combined coverage of the GSL and the AWL is 81.7%, which is somewhat lower than that obtained by the newer corresponding word lists (83.29% for the NGSL + the NAWL), which is understandable, given that the newer lists perform better in modern corpora. The coverage of the GSL in our corpus (75.08%) is comparable to that found by Coxhead (2000) for her mixed academic corpus (76.1%). However, the GSL covered just 72.9%

in Coxhead's science subsection of the corpus. As expected, we can see that our science magazine corpus contains more general-purpose words than it is the case with Coxhead's academic science corpus, which shows that the language of science magazines is less specialised than that of academic science.

The coverage of frequent academic words (as represented by the AWL) was 6.62% in our science magazine corpus, which is considerably lower than that found for Coxhead's academic corpus (2000), where its coverage was 10%, and for her science subsection, where its coverage was 9.1%. This, too, might have been expected but it needed an empirical confirmation and precise results – when compared to Coxhead's results for her academic science corpus, our Science Magazine Corpus features 27.26% less academic vocabulary. Therefore, for the students aiming at learning academic scientific English, the genre of science magazines as a source of vocabulary would not be fully appropriate. However, the level of academic vocabulary in science magazines is certainly not negligible and this finding does not exclude them as a genre to be incorporated, to some degree, in English for Science resources.

Our using the GSL and the AWL is also justified by the fact that the Science List – the SL, was built on top of them, i.e. in conjunction with them. In our Science Magazine Corpus, the SL covers almost 2%, i.e. one in every 50 words. For comparison, in the academic science corpus used by Coxhead and Hirsh (2007), its coverage was 3.79%. For the sake of precision, this means that we found 48.26% fewer scientific words in our Science Magazines Corpus than Coxhead and Hirsh found in their academic science corpus, which is a substantial difference. Still, science magazines do feature scientific words to some extent, and so the results do not point to their complete exclusion as a genre when considering sources for English for Science texts.

Bearing in mind that the target readers of science magazines are not exclusively scientists, it is expected that their vocabulary will contain fewer academic and technical words, as our results have corroborated. Having conducted this study, we know exactly how many fewer – the science magazine articles contained less academic vocabulary by about a third, and less scientific-technical words by about a half, in comparison to academic science texts. In addition, the selected science magazines also contained a somewhat simpler general vocabulary, making them more readable for L2 students.

The LFP method produces numerical results, for which it has sometimes been criticised, as suggested in the theoretical section of this paper. In order to overcome this disadvantage, we will illustrate what the texts of science magazine articles look like and show what words fall under the scope of the word lists referred to above. To this end, we will quote an extract from our corpus and mark the academic and the scientific words (the bold marks the words from the AWL and the underlined items are from the SL):

"This discovery triggered multiple campaigns of follow-up observations with radio telescopes worldwide. One of these used the Very Large Array (VLA) in New Mexico, a collection of 27 radio antennas observing in tandem, to regularly search for events on millisecond timescales in the same area of the sky as FRB 121102. This survey had the unique capability to pinpoint radio bursts' locations on the sky several orders of magnitude better than a single radio dish could. After roughly six months of observations, the team – led by Shami Chatterjee of Cornell University – discovered and localized a burst. Soon an even more precise location for this FRB came through the technique of very long baseline interferometry, where signals from multiple telescopes around the world are combined to synthesize a much larger virtual telescope with exquisite resolution on the sky. The finding, led by Benito Marcote of the Joint Institute for VLBI ERIC (JIVE) in the Netherlands and his colleagues, pinpointed the repeated bursts from FRB 121102 with an uncertainty of less than one arc second (1/3,600 of a degree)" (Lorimer and McLaughlin, 2018: 46).

In this extract, there are four words from the Science List: *array, magnitude,* synthesize and *arc*, whereas the academic words (the AWL), naturally, were more frequent – they included nine words: *multiple, radio, collection, burst, search, dish, signal, combined* and *repeated*. Most of the other words from the extract were frequent general-purpose words, which is why we decided not to mark them.

If we apply the word lists which Nation (2012) produced based on their frequency in the BNC and COCA corpora, we can calculate the vocabulary load of our Science Magazine Corpus and determine the target group of English for Science learners for whom they will work best. We present these results in Table 4:

BNC/COCA WORD LISTS	SCIENCE MAGAZINES	
	TOKEN%	CUMULATIVE %
Proper nouns	3.25	3.25
Marginal words and letters of the alphabet	0.74	3.99
Transparent compounds	0.45	4.44
Abbreviations	0.4	4.84
1st 1,000 word families	67.01	71.85
2nd 1,000 word families	10.77	82.62
3rd 1,000 word families	6.97	89.59
4th 1,000 word families	2.23	91.82
5th 1,000 word families	1.34	93.16
6th 1,000 word families	0.77	93.93

7th 1,000 word families	0.58	94.51
8th 1,000 word families	0.55	95.06

Table 4. Coverage of Nation's word lists in the Science Magazine Corpus

Considering the results presented in Table 4, we find that the coverage of 95%, needed for the minimum comprehension threshold (Laufer 1989), is reached at 8,000 most frequent word families of English. For comparison, we will note that Hirsh and Nation (1992) find that 5,000 words are needed to read novels for teenagers; in addition, Nation (2006) calculates that 4,000 words are required to read some novels and newspapers, as well as watch children's movies, whereas about 3,000 words are necessary to understand spoken English. Moreover, Coxhead and Walls (2012) calculate that 5,000 words are needed to listen to TED talks. In comparison with all these different genres, we may say that the genre of science magazines is substantially more demanding vocabulary-wise.

As suggested earlier, Milton (2010) establishes that the knowledge of 98% of the words used in a text would be necessary for the C levels, which typically corresponds to 8,000–9,000 words for most genres, as Nation calculates (2013). Bearing this in mind, we can conclude that the lexical demand of science magazines is rather high and that a strong vocabulary is needed to read them, one typically held by those at the C levels of mastery (according to the CEFR). Therefore, unadapted, the texts belonging to the genre of science magazines could only be read by the already very advanced and fluent learners, which certainly reduces the pool of the possible target learners for whom we might use this genre. Such learners will have to know the most frequent general-purpose, academic and scientific-technical vocabulary. Therefore, a safe conclusion is that unadapted texts of science magazines are accessible to C-level students.

B2-level students, with their assumed knowledge of up to 5,000 words (Milton 2010), will know up to 93% of the words of our corpus (Table 4 – with 5,000 word families, a coverage of 93.16% is attained). To reach Laufer's comprehension threshold (met at a 95%-coverage), at least 2% of the remaining words in science magazine articles would need to be adapted, i.e. replaced by some more common words, in order to make them more readable for B2 students. This means that 1 in every 50 words (or five lines of a text) would need to be simplified, explained or substituted.

On the other hand, for a B2 student to reach the ideal reading level (for which he/she should know 98% of the words used in the text, as per Nation (2006)), a teacher or a teaching material producer would need to adapt 5% of the uncovered words, i.e. 1 in every 20 words (or every two lines of a text). Though technically demanding, this still sounds feasible, and our conclusion is that, with some adaptation, the texts from science magazines can be used as supplementary teaching and learning resources for both upper-intermediate and advanced students, granted, though, that they would likely be more useful for the latter group.

To illustrate the vocabulary level of science magazine articles, we will quote another extract from our corpus and mark the level of words[4]:

> "WASPS[6] literally[3] drum[2] up[1] interest[1] in[1] food[1], banging[2] their[1] abdomens[5] against[1] the[1] walls[1] of[1] their[1] nest[2] to[1] inform[1] their[1] nestmates that[1] food[1] is[1] available[2]. We[1] have[1] known[1] since[1] the[1] 1960s that[1] several[1] species[2] of[1] wasp[6] perform[2] gastral drumming[2] from[1] time[1] to[1] time[1] banging[2] their[1] abdomens[5] against[1] their[1] nest[2] walls[1] in[1] a[1] series[2] of[1] short[1] bursts[2]. The[1] scientists[1] who[1] first[1] reported[1] this[1] behaviour[3] thought[1] it[1] may[1] be[1] a[1] signal[2] that[1] the[1] wasps[6] were[1] hungry[1]. Meanwhile[3], other[1] researchers[2] suggested[1] the[1] wasps[6] might[1] be[1] telling[1] nestmates about[1] food[1] sources[3]. Such[1] recruitment[3] behaviour[3] is[1] common[2] in[1] social[2] animals[1], from[1] house[1] sparrows[9] to[1] naked[3] mole[6] rats[2]. Benjamin[PW] Taylor[PW] at[1] the[1] City[1] University[2] of[1] New[1] York[PN] and[1] his[1] colleagues[3] have[1] now[1] put[1] the[1] two[1] ideas[1] to[1] the[1] test[1]. The[1] team[1] took[1] six[1] colonies[3] of[1] German[PW] yellowjacket wasps[6] (Vespula germanica) and[1] housed[1] them[1] in[1] artificial[4] nests[2]. The[1] wasps[6] were[1] allowed[1] to[1] freely[1] forage[7] for[1] a[1] day[1], but[1] the[1] next[1] day[1] they[1] were[1] shut[1] in[1] and[1] given[1] only[1] water[1], or[1] a[1] sucrose[18] solution[3]. On[1] the[1] third[1] day[1], the[1] exit[4] was[1] opened[1] again[1]. Drumming[2] declined[3] when[1] the[1] wasps[6] were[1] given[1] only[1] water[1], suggesting[1] it[1] was[1] not[1] a[1] signal[2] of[1] hunger[1]. The[1] wasps[6] drummed[2] more[1] when[1] sucrose[18] was[1] offered[1], and[1] the[1] levels[1] of[1] drumming[2] consistently[3] returned[1] to[1] a[1] baseline[TC] level[1] on[1] the[1] third[1] day[1]. This[1] suggests[1] that[1] the[1] wasps[6] drum[2] to[1] alert[3] each[1] other[1] to[1] the[1] presence[3] of[1] food[1] (The[1] Science[1] of[1] Nature[1], doi[PW].org[ABR]/cm4d)" (Kemmeny 2018).

As can be seen, most of the words from this extract belong to the first 2,000 words of English (marked with the superscripts 1 and 2). In fact, the vast majority of them comes from the first frequency band (1st 1,000 words in the BNC/COCA). A smaller group of words comes from the third band (*literally, meanwhile, behaviour, recruitment, naked, solution, declined, source*). Significantly fewer words are used from the higher frequency bands. For instance, there is just one word from the 18th 1,000 words – *sucrose,* which is a rather specialised and technical word; *sparrow* comes from the ninth band; *forage* is from the seventh, while *mole* and *wasp* are from the sixth band. The general knowledge of the words up to the ninth frequency band may be expected from an advanced student (of course, students of any level might know some individual words from the higher bands as well).

4 The legend: 1, 2, 3… mark the frequency band to which a word belongs, i.e. the 1st 1,000, the 2nd 1,000, the 3rd 1,000, etc.; ABR. stands for abbreviation, PW stands for proper word, while TC stands for a transparent (non-hyphenated) compound.

5. Pedagogical implications and recommendations

Based on the findings of the study, the following pedagogical implications and recommendations can be made:

- The target students' vocabulary size should first be tested (this is possible by means various Vocabulary Size Tests, some of which are available online). Those designing teaching materials who cannot test their target students, may follow the general guideline that they can include unadapted texts from science magazines for the target C-level groups, as well as adapted texts from science magazines for the target B2-level groups.
- Texts from science magazines contain less academic vocabulary than science genres but still sufficient to allow for the learning of some academic vocabulary in the ES and the EAP courses.
- Texts from science magazines contain 50% fewer scientific words than science genres, which means that they can only occasionally be used as a source of this type of vocabulary for English for Science students.

Overall, we recommend including texts from science magazines for the B2- and the C-level English learners on an occasional basis, preferably in an adapted, i.e. simplified form. The appeal of this genre, as suggested in the introduction, would be motivating for the students, while the teachers who are typically disciplinary outsiders should find it easier to teach such texts. However, as suggested, this will only work under the conditions presented above.

6. Limitations of the study

A limitation of this study lies in its somewhat limited corpus – for future research, we suggest employing a larger corpus containing a wider variety of science magazines. Also, the more specialised science magazines (Vuković-Stamatović 2020), i.e. those intended for a narrower field such as one discipline (i.e. physics, chemistry…), may be more lexically demanding and contain more academic and specialised vocabulary than is found in general science magazines, covering a wide variety of topics, such as the two selected for the purpose of this study. This means that, if they decide to use science magazines for their sources of materials, teachers and teaching materials producers should bear this consideration in mind – science magazines may vary vocabulary-wise according to how specialised they are. The more specialised magazines should be used for sources for the more specific English for Science branches (e.g. *English for Physics* or

English for Chemistry, etc.), while the broader, multidisciplinary ones are recommended for general English for Science courses.

7. Conclusion

In this paper we studied whether and how texts from science magazines can be used for English for Science classes. Our corpus contained about 230,000 running words coming from two science magazines – *New Scientist* and *Scientific American*.

We calculated the vocabulary load of the texts published in these two magazines. It turned out that the coverage of 95% of the words, needed for minimum comprehension (Laufer 1989), may require the knowledge of as many as 8,000 words. This means that the language of science magazines is rather lexically demanding and that the articles published in them may be used for advanced learners and, with some more adaptations, for upper-intermediate learners as well. Although rather lexically dense, the vocabulary of science magazines proved not to be as dense as that of academic scientific English. Naturally, it contained fewer academic and substantially fewer specialised, i.e. scientific-technical words. We showed that this genre could well be used for English for Science resources, under certain conditions.

For future research we also recommend exploring the vocabulary of other genres that may be employed as teaching and learning resources for English for Science classes.

References

Anthony, Laurence (2017). *AntFileConverter* (Version 1.2.1). Tokyo: Waseda University. [online] http://www.laurenceanthony.net/software.

Anthony, Laurence (2014). *AntWordProfiler* (Version 1.4.1). Tokyo: Waseda University. [online] http://www.laurenceanthony.net/software.

Brezina, Vaclav & Dana Gablasova (2013). "Is There a Core General Vocabulary? Introducing the New General Service List". *Applied Linguistics* 36 (1). 1–22.

Browne, Charles, Brent Culligan & Joseph Phillips (2013a). *The New General Service List*. [online] http://www.newgeneralservicelist.org/. https://doi.org/10.1093/applin%2famt018.

Browne, Charles, Brent Culligan & Joseph Phillips (2013b). *The New Academic World List 1.0*. [online] http://www.newgeneralservicelist.org/nawl-new-academic-word-list/.

Charles, Maggie. "English for academic purposes". In: Brian Paltrige & Sue Starfield (Eds.) *The handbook of English for specific purposes*. Chichester: Wiley-Blackwell. 137–154.

Council of Europe (2001). *The Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.

Coxhead, Averil (2000). "A New Academic Word List". *TESOL* 34 (2). 213–238. https://doi.org/10.2307/3587951.

Coxhead, Averil & David Hirsh (2007). "A pilot science-specific word list". *Revue Française de Linguistique Appliquée* 12 (2). 65–78. https://doi.org/10.3917/rfla.122.0065.

Coxhead, Averil & Roz Walls (2012). "TED Talks, vocabulary, and listening for EAP". *TESOLANZ Journal* 20 (1). 55–67.

Crossley, Scott A., Tom Cobb & Danielle S. McNamara (2013). "Comparing count-based and band-based indices of word frequency: Implications for active vocabulary research and pedagogical applications". *System* 41 (4). 965–981. https://doi.org/10.1016/j.system.2013.08.002.

Dimova, Slobodanka, Anna Kristina Hultgren & Christian Jensen (Eds.) (2012). *English-Medium Instruction in European Higher Education. Vol. 4*. Berlin: Walter de Gruyter.

Douglas, Scott Roy (2015). "The Relationship between Lexical Frequency Profiling Measures and Rater Judgements of Spoken and Written General English Language Proficiency on the CELPIP-General Test." *TESL Canada Journal* 32 (9). 43–64. https://doi.org/10.18806/tesl.v32i0.1217.

Ferguson, Gibson (2007). "The global spread of English, scientific communication and ESP: Questions of equity, access and domain loss." *Ibérica: Revista de la Asociación Europea de Lenguas para Fines Específicos (AELFE)* 13. 7–38.

Hirsh, David, and Paul Nation (1992). "What vocabulary size is needed to read unsimplified texts for pleasure?" *Reading in a Foreign Language* 8. 689–689.

Jin, Ng Yu, et al. (2013). "Development of the engineering technology word list for vocational schools in Malaysia". *International Education Research* 1 (1). 43–49. 10.12735/ier.v1i1p43.

Khani, Reza & Khalil Tazik (2013). "Towards the development of an academic word list for applied linguistics research articles." *RELC journal* 44 (2). 209–232. https://doi.org/10.1177/0033688213488432.

Laufer, Batia (1989). "What Percentage of Text-Lexis is Essential for Comprehension?" In: Christer Laurén & Marianne Nordman (Eds.). *Special Language: From Human Thinking to Thinking Machines*. Clevedon: Multilingual Matters. 316–336.

Laufer, Bakia & Geke C. Ravenhorst-Kalovski (2010). "Lexical Threshold Revisited: Lexical Text Coverage, Learners' Vocabulary Size and Reading Comprehension". *Reading in a Foreign Language* 22 (1). 15–30.

Laufer, Bakia & Paul Nation (1995). "Vocabulary Size and Use: Lexical Richness in L2 Written Production". *Applied Linguistics* 16 (3). 307–22. https://doi.org/10.1093/applin/16.3.307.

Lindqvist, Christina, Anna Gudmundson & Camilla Bardel (2013). "A New Approach to Measuring Lexical Sophistication in L2 Oral Production". In: Camilla Bardel, Christina Lindqvist & Bakia Laufer (Eds.). *L2 Vocabulary Acquisition, Knowledge and Use: New Perspectives on Assessment and Corpus Analysis*. European Second Language Association. 109–26.

Miller, Lindsay (2014). "English for science and technology". In: Bhatia Vijay & Stephen Bremner (Eds.). *The Routledge handbook of language and professional communication*. Routledge/Taylor Francis. 304–320.

Milton, James (2010). "The development of vocabulary breadth across the CEFR levels". In: Inge Bartning, Maisa Martin & Ineke Vedder (Eds.). *Communicative proficiency and linguistic development: Intersections between SLA and language testing research*. Eurosla Monograph Series I. 211–232.

Morris, Lori & Tom Cobb (2004). "Vocabulary Profiles as Predictors of the Academic Performance of Teaching English as a Second Language Trainees". *System* 32 (1). 75–87. https://doi.org/10.1016 /j.system.2003.05.001.
Nagy, William (1988). *Teaching vocabulary to improve reading comprehension*. Urbana: International Reading Association.
Nation, Paul (2012). *The BNC/COCA Word Family Lists*. [online] https://www.wgtn.ac.nz/data/assets/pdf_file/0005/1857641/about-bnc-coca-vocabulary-list.pdf [1 September 2019].
Nation, Paul (2006). "How large a vocabulary is needed for reading and listening?" *Canadian Modern Language Review* 63. 59–82.
Nation, Paul (2012). *The BNC/COCA word family lists*. [online] http://www.victoria.ac.nz/lals/about/staff/paul-nation [1 April 2018].
Nation, Paul & Robert Waring (1997). "Vocabulary Size, Text Coverage and Word Lists". In: Norbert Schmitt & Maria Pigada (Eds.). *Vocabulary: Description, Acquisition and Pedagogy.* Cambridge: Cambridge University Press. 6–19.
Nation, Paul (2013). *Learning Vocabulary in Another Language*. Second Edition. Cambridge: Cambridge University Press.
Parkinson, Jean (2013). "English for science and technology". In: Brian Paltridge & Sue Starfield (Eds.). *The Handbook of English for Specific Purposes*. Wiley-Blackwell. 156–173.
Read, John & Paul Nation (2006). "An investigation of the lexical dimension of the IELTS speaking test." *IELTS Research Reports* 6. 207–231.
Tardy, Christine (2004). "The role of English in scientific communication: lingua franca or Tyrannosaurus rex?" *Journal of English for Academic Purposes* 3 (3). 247–269. https://doi.org/10.1016/j.jeap.2003.10.001.
Vuković-Stamatović, Milica (2020). "Vocabulary complexity and reading and listening comprehension of various physics genres." *Corpus Linguistics and Linguistic Theory* 16.3. 481-514. https://doi.org/10.1515/cllt-2019-0022.
Wang, Jing, Shao-lan Liang & Guang-chun Ge (2008). "Establishment of a medical academic word list". *English for Specific Purposes* 27 (4). 442–458. https://doi.org/10.1016/j.esp.2008.05.003.
West, Michael (1953). *A general service list of English words: with semantic frequencies and a supplementary word-list for the writing of popular science and technology*. Green: Longman.

Milica Vuković-Stamatović
Vesna Bratić
University of Montenegro