



## Do textbooks support incidental vocabulary learning? – a corpus-based study of Swedish intermediate EFL materials

Denise Bergström, Cathrine Norberg & Marie Nordlund

To cite this article: Denise Bergström, Cathrine Norberg & Marie Nordlund (2025) Do textbooks support incidental vocabulary learning? – a corpus-based study of Swedish intermediate EFL materials, *Education Inquiry*, 16:1, 69-87, DOI: [10.1080/20004508.2022.2163050](https://doi.org/10.1080/20004508.2022.2163050)

To link to this article: <https://doi.org/10.1080/20004508.2022.2163050>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 28 Dec 2022.



Submit your article to this journal [↗](#)



Article views: 3352



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 5 View citing articles [↗](#)

# Do textbooks support incidental vocabulary learning? – a corpus-based study of Swedish intermediate EFL materials

Denise Bergström , Cathrine Norberg  and Marie Nordlund 

Department of Health, Education and Technology, Luleå University of Technology, Luleå, Sweden

## ABSTRACT

Learning vocabulary is a central but yet complex aspect of learning a language. Hence, researchers stress the importance of facilitating vocabulary development via a structured approach to target words and recycling. While teaching materials have the potential to provide this structure to all students in a classroom, few studies have investigated the vocabulary component of textbooks and the learning opportunities they provide. In the present study, the texts in five series of EFL materials aimed at intermediate learners in Swedish secondary school (years 7–9) were investigated, using corpus-based methods. The results indicate that the texts encompass a suitable amount of unknown vocabulary for vocabulary learning from reading and provide exposure to mid-frequency vocabulary. However, it was also found that these items are not recycled sufficiently. Rather, the materials mainly recycle lexical items that students are likely to know already. It is therefore concluded that although the materials offer input suitable for the target students, they are not structured in a way that supports vocabulary development.

## KEYWORDS

Vocabulary development; teaching materials; EFL; recycling; mid-frequency vocabulary

## 1. Introduction

Language learners face a considerable vocabulary learning task as many words are needed for communication and comprehension. It has, for example, been found that a learner needs to know about 98% of the words in a text to fully comprehend it (e.g. Hu & Nation, 2000; Schmitt, Jiang, & Grabe, 2011). Nation (2006) has suggested that knowledge of about 8–9,000 word families<sup>1</sup> is needed to understand texts in English, such as newspapers or novels. Furthermore, Nation and Coxhead (2021) report that 13- to 14-year-old native-speakers have a vocabulary size comprising 9,000 word families. These figures show that the vocabulary target when learning English has to be high, if learners are to become proficient language users. Besides knowing many words, students also need to know the most useful words in the language, as these equip them for encounters with authentic language (cf. Nation, 2013). Moreover, if students are to acquire new words, they have to encounter these words many times (cf. Ellis, 2002). These findings demonstrate that vocabulary development is of utmost importance for

**CONTACT** Denise Bergström  [denise.bergstrom@ltu.se](mailto:denise.bergstrom@ltu.se)  Department of Health, Education and Technology, Luleå University of Technology, Luleå 97187, Sweden

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

a language learner and that it is a complicated endeavour. Researchers, such as Newton (2020), have hence argued that students need structured support to succeed in their vocabulary learning, for instance via planning and recycling of target vocabulary, and that the classroom plays a crucial part in providing this structure. It is therefore important that the EFL classroom and teaching materials are structured in a way that facilitates word learning.

Although a substantial amount of linguistic research has been dedicated to developing principles for successful vocabulary teaching (Nation, 2011), it remains unclear whether EFL instruction is structured in a way that supports students' vocabulary learning. Studies have, for example, found that EFL teachers do not perceive vocabulary as a distinct learning objective in the classroom (e.g. Gao & Ma, 2011; Hermagustiana, Hamra, Rahman, & Salija, 2017) and that they generally do not structure their teaching with vocabulary in mind (e.g. Bergström, Norberg, and Nordlund's (2022) interview study with Swedish EFL teachers). This absence of vocabulary focus has also been identified in relation to the production of teaching materials. A recent interview study with Swedish developers of EFL materials, for example, indicates that vocabulary is not a primary concern during the design process (Bergström, Norberg, & Nordlund, 2021). Similar results were found when Prowse (2011) surveyed textbook authors from different countries and he concludes that they seem to rely on their intuition and focus on content rather than language input. These studies suggest a lack of awareness regarding the importance of structured vocabulary, which raises the question of whether there is sufficient support for vocabulary development in the EFL classroom.

Textbooks are widely used by EFL teachers and students (e.g. Gray, 2016) and may thus function as a main provider of vocabulary input in the classroom. It has been found that vocabulary learning from reading is facilitated by adapted texts (e.g. graded readers), where the amount of unknown vocabulary is regulated (Nation, 2013; Schmitt, 2008). Researchers have suggested that a text suitable for vocabulary development should include at least 95% already known vocabulary (Nation & Anthony, 2013; Schmitt & Schmitt, 2020) and recycle appropriate target vocabulary (e.g. Waring & Takaki, 2003). If these conditions are met, a learner is likely to be able to read and comprehend the text and acquire vocabulary whilst reading. Teaching materials could hence play an important role in facilitating vocabulary development by providing supportive input, as pointed out by Schmitt (2019), especially as textbook authors have the opportunity of structuring the input far more than the individual teacher may have time or resources to do (see also Hughes, 2019). Since teaching materials have this potential and adapted texts support vocabulary development, it could be expected that materials encompass texts adapted to the target students.

Although teaching materials are central in language classrooms and have the potential to offer a principled approach to vocabulary, including both text adaptation and recycling, few studies have investigated the vocabulary learning opportunities in EFL teaching materials. This study maps the lexical component of Swedish intermediate teaching materials, in terms of the lexical profile of the texts and the extent to which words are recycled. So far, studies conducted on intermediate textbooks are all in Asian contexts (e.g. Sun & Dang, 2020; Yang & Coxhead, 2020) and in a Swedish context, the focus has primarily been on books for young learners (e.g. Norberg & Nordlund, 2018; Nordlund, 2015a). The present study is thus likely to add valuable insights into an area of research, which so far has been given little attention.

## 1.1 Literature review

### 1.1.1 The role of word frequency in vocabulary development

As mentioned, the vocabulary learning task does not only entail learning many words. It also means acquiring the most useful words for comprehension and communication. Given the nature of language, not all words occur as frequently in a discourse of a general kind. Rather, a few words actually make up a large proportion of the language. It has been estimated that the thousand most frequent word families in English cover about 85% of general English (Webb & Nation, 2017), which renders these words very important to learn initially. Vocabulary researchers (e.g. Nation, 2013) therefore argue that vocabulary learning should be structured following a frequency principle where target words are chosen on a frequency basis. Empirical research has found that knowing high-frequency items has a positive impact on general language proficiency as well. Sakata (2019), for example, investigated how 52 Japanese EFL students' language proficiency was impacted by what words they already knew. The results indicate that when students know the same number of words, the ones who know more frequently occurring words have a higher general language proficiency than those who know fewer. This suggests that it is important that teaching materials and language teachers take the frequency profile of target vocabulary into consideration.

In order to facilitate a frequency approach to vocabulary learning, frequency-based word lists have been constructed (Nation, 2016) where words are categorised based on their frequency in general corpora like the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA). In such word lists, the first frequency band comprises the thousand most frequent word families in the corpus, the second band the 1,001–2,000 most frequent ones and so on. As frequency lists reflect common language use, they can function as a structuring tool in the classroom and for materials development. The lists can also be used to compare lexical input to general English and to calculate the lexical demands posed by different texts. For instance, Coxhead, Stevens, and Tinkle (2010) used wordlists to calculate the difficulty of science textbooks by identifying what vocabulary size was needed to reach 98% text coverage, and Webb and Rodgers (2009) analysed the lexical demands of English-speaking movies using a similar method. Word lists can thus facilitate the analysis of educational texts and assess their appropriateness for an intended group of learners.

Concerning the vocabulary target in the EFL classroom, it has traditionally been suggested that students need to focus on high-frequency words and academic vocabulary (e.g. Nation, 2013). Schmitt and Schmitt (2014), however, stress the importance of also focusing on mid-frequency vocabulary, because these words equip students for encounters with authentic language, such as English-medium instruction and leisure activities like watching TV. They therefore argue that, like high-frequency vocabulary, mid-frequency words require pedagogical attention to facilitate acquisition (see also Qian & Lin, 2020). In Schmitt and Schmitt's (2014) now widely used framework, the first 3,000 word families are considered high-frequency and the 3,001 to 9,000 mid-frequency. Further, relevant to this study is the expected level of Swedish secondary school students. The National Agency for Education (Skolverket, 2012) states that a passing grade in primary school English is equivalent to being on an A2 level (i.e. basic use) in the Common European Framework of Reference for Languages (CEFR;

Council of Europe, 2001). Students starting year 7 can thus be expected to master basic vocabulary in English. As teaching materials should encompass input that matches the students' needs, EFL teaching materials for Swedish secondary school need to provide input of not only high-frequency vocabulary, but also sufficient mid-frequency vocabulary.

### 1.1.2 *The impact of recycling on incidental vocabulary learning*

Researchers (e.g. Ellis, 2002; Webb, 2014) are in agreement concerning the importance of repeated encounters for vocabulary learning, but, depending on the learning activity, the extent of recycling needed for acquisition differs. Vocabulary learning activities can be divided into two main kinds, intentional and incidental, where the former entails that the main focus is to acquire vocabulary and the latter that vocabulary development is a by-product of another activity, such as reading for content (e.g. Hulstijn, 2003). Textbooks, which are the focus in the present study, can provide both incidental and intentional vocabulary learning opportunities, from the texts and vocabulary-focused exercises respectively. Here, the texts in teaching materials are studied and, accordingly, the incidental vocabulary learning opportunities. In relation to recycling, incidental vocabulary learning requires more encounters than intentional focused learning (Webb, 2020). In order for students to be able to develop vocabulary from reading, it is thus important that textbooks provide repeated encounters with target words.

Studies seeking to identify a sufficient number of repetitions for incidental vocabulary learning from reading have rendered different results and no general figure is agreed on. For example, Pigada and Schmitt's (2006) case study, although with only one Greek-speaking participant, indicates that more than 20 encounters are necessary to acquire deep knowledge of words. Similar results are reported in an investigation by Waring and Takaki (2003) with 15 Japanese EFL university students, where the long-term retention rate is low (10–15%) even when words are encountered more than 18 times. In Webb's (2007) experiment with 121 Japanese EFL students, it was, on the other hand, shown that ten encounters can lead to substantial vocabulary gains, but he also points out that more encounters may be needed to develop full knowledge of a word (see also Chen and Truscott's (2010) study with 72 Taiwanese students, which reports similar findings). Along the same lines, Pellicer-Sánchez and Schmitt (2010) investigated the vocabulary gains when 20 Spanish EFL learners read a novel. They found a general increase in learning after 5–8 repetitions, but that more than ten encounters are probably necessary for sustainable vocabulary gains (see also Uchihara, Wee, and Yanagisawa's (2019) meta-analysis of 26 studies concerning the effects of recycling on incidental vocabulary learning). It has also been suggested that it is likely that the number of repetitions required will vary between words as they may differ in difficulty and salience and that the number of repetitions also depends on the kind of knowledge learners are expected to attain (Ellis & Wulff, 2015; Webb, 2014). Hence, even though no general figure can be firmly established, it is clear that a word needs to be encountered many times if a learner is to develop lasting, productive knowledge of it. This has also been found in classroom-oriented research, such as Brown's (2021) study at a Japanese university where he found that the frequency of target words has an impact on students' incidental vocabulary gains. Recycling of lexical

items is therefore an important principle to take into account when constructing materials, as argued by Schmitt (2019).

### 1.1.3 *Previous studies of word frequency and recycling in textbooks*

Although research focused on the vocabulary component of EFL teaching materials is scarce, there are some studies that have investigated the opportunities for incidental vocabulary learning in textbooks. These studies have mainly been conducted in an Asian context, as previously mentioned. They have identified issues concerning primarily the nature of the input, in terms of frequency, and the extent of recycling in the books. Alsaif and Milton (2012), as an example, investigated 22 EFL textbooks used in Saudi Arabia (years 6–12) and found that the books include over 80% of the 2,000 most frequent word families in English, but only about half of the 5,000 most frequent ones. Similarly, Sun and Dang (2020) analysed eleven EFL materials used in Chinese high schools. Their analysis showed that the textbooks cover many of the thousand most frequent word families and that these words are recycled to a high extent. They also established, however, that the exposure to vocabulary from the second and third frequency bands is insufficient in the materials and that the words from those bands are not frequently recycled. They argue that this could explain why Chinese high school students in general do not master the first 3,000 words when finishing high school. Worth mentioning in this context is also Yang and Coxhead's (2020) analysis of two Chinese EFL textbooks in a series and the lexical demands posed by them. Their study reveals that 98% text coverage is achieved in the fifth and sixth frequency band respectively (see also Coxhead and Boutorwick's (2018) analysis of German English textbooks which shows that high school students need to know the 8,000 most frequent word families to comprehend the texts). They also found that 65% of the mid-frequency items occur only once in the books and therefore conclude that the learning conditions for mid-frequency vocabulary are unsatisfactory. Matsuoka and Hirsh (2010) conducted a similar study on one upper-intermediate textbook and discovered that 31% of the words from the second frequency band are recycled five or more times, which they judge as sufficient recycling. Academic words (defined as part of the Academic Word List [Coxhead, 2000]) and vocabulary below the second frequency band, however, do not occur frequently enough to be acquired incidentally from the exposure in the books. On a general level, the studies presented above indicate serious flaws in intermediate EFL textbooks in terms of primarily recycling, but also exposure to vocabulary from suitable frequency bands.

In a Swedish context, the evaluation of EFL materials has focused on textbooks aimed at young learners (see, however, Ljung, 1990 analysis of the vocabulary in Swedish upper-secondary school textbooks). Nordlund (2015a), for example, analysed three textbooks aimed at pupils aged 10–12 and found very low levels of recycling. Only 4% of nouns and adjectives and 12.5% of lexical verbs occur more than twelve times in the books. Moreover, a large proportion of the lexical content is words from the third frequency band or lower. Similarly, Norberg and Nordlund (2018) evaluated seven EFL textbooks for elementary students and showed that only about 40% of the content words come from the first frequency band, a number surprisingly low considering that the target pupils are young learners and thus in need of exposure to high-frequency vocabulary. They also identified, based on type-token ratio, that recycling is not

a prominent feature in the books (see also Nordlund, 2015b, 2016). The results from the above-mentioned studies suggest that the lexical input provided in Swedish textbooks for young learners of English does not appear to be systematically planned in terms of frequency and recycling. This raises the question of whether the same pattern is present in intermediate EFL materials used in Swedish secondary schools.

Although the studies mentioned above are similar to the present investigation, the approaches vary, especially for characterising the lexical input. They typically analyse the lexical component by calculating the lexical demands for reading comprehension or the extent to which different frequency bands are represented in textbooks. This study instead focuses on characterising what vocabulary is likely to be learned from reading the texts included in teaching materials. As teaching materials have a central role in the classroom and can be a facilitative tool for principled vocabulary development, their vocabulary component is of interest. Intermediate students face a considerable learning task because mid-frequency vocabulary is less frequent than beginner-level vocabulary in general discourse and may thus not be encountered as often outside of the classroom. Like young learners, they need teaching materials constructed in a way that support vocabulary acquisition at their level of learning.

## 1.2 Research questions

The aim of this study is to determine to what extent the lexical input in Swedish EFL textbooks for secondary school supports incidental vocabulary learning for intermediate learners. The research questions addressed are:

- (1) What characterises the lexical input in terms of frequency distribution and suitability for the target students?
- (2) To what extent are lexical items recycled in the textbooks?
- (3) To what extent are mid-frequency words recycled in the textbooks?

## 2. Material and methods

### 2.1 Corpus compilation

The study analyzes five series of EFL teaching materials in Swedish secondary school (years 7–9, students aged 13–15). The materials come from four prominent publishers in Sweden and were also found to be frequently used via an informal survey among teachers. By using both digital and printed materials from different publishers, the intention is to represent a range of the materials available in the local context. The printed textbooks series *Awesome* (Childs-Cutler & Gentili Cronholm, 2017, 2018; Childs-Cutler, Gentili Cronholm, & Niemann, 2016), *Good Stuff Gold* (Coombs, Bayard, Hagvörn, & Johansson, 2012, 2013, 2014) and *Wings* (Frato et al., 2015, 2016, 2017) were studied as well as the digital teaching materials *Digilär<sup>2</sup>* (Gode, 2014) and *Gleerups Engelska* (Taylor, Carlsson, & Smith, n.d.).

A corpus was compiled of the reading texts from the five series of the teaching materials. As the focus of the study is incidental learning through reading, word lists,



**Table 1.** Texts, types and tokens in the sub-corpora.

	Awesome	Digilär	Gleerups	GS* Gold	Wings
Texts	89	74	92	161	134
Types	9,101	6,443	7,943	7,700	9,545
Tokens	77,165	39,230	74,660	76,033	74,560

\* GS = Good Stuff

comprehension questions and exercise instructions were excluded from the corpus. While the reading texts are likely to be used for meaning-focused activities, instructions and word lists are less likely to be part of the meaning-focused input. It should be noted, however, that teachers and learners may use and approach reading texts in different ways and that the texts thus may be used as language-focused input in classrooms as well. Nevertheless, the focus of the present study is to investigate whether the texts support students when reading them for content. The results thus illuminate their incidental language learning function.

The reading texts are very varied in terms of genre and length. All the materials encompass a considerable variation in this respect, including both text extracts from novels and texts written for learning purposes. As the materials are aimed at years 7–9, each series comprises three parts (one per school year), either as distinct books or three parts of an online digital resource. The corpus consequently consists of five sub-corpora, one for each series, each comprising texts from three books (see Table 1).

The corpus was compiled by scanning the physical textbooks and downloading the texts from the digital teaching materials.<sup>3</sup> The scanned files were manually examined to make sure that the texts corresponded to the physical copy. All the files were then converted into.txt-files. To facilitate searches and analyses, corpus files are typically part-of-speech tagged, which means that grammatical information is encoded in the files. The corpus tool Sketch Engine<sup>4</sup> (Kilgariff, Rychlý, Smrž, & Tugwell, 2004), which automatically tags the files used (with a modified version of the Penn TreeTagger tagset [Taylor, Marcus, & Santorini, 2003]), was used in the analysis. Therefore, no further preparation of the material was necessary.

## 2.2 Analysis

The analysis focused on the frequency distribution and recycling of words. This means that the texts in the materials are illuminated from a lexical perspective. It should thus be noted that the analysis does not capture the learning opportunities of the texts or textbooks as a whole, rather, the incidental learning opportunities from reading is in focus. In order to answer the research questions, the corpus of reading texts was analysed using Sketch Engine and Compleat Web VP (Cobb (n.d.). To characterise the lexical input (RQ1), the frequency distribution of the texts in the sub-corpora was mapped with the vocabulary profiler (VP), where the input is compared to the BNC-COCA 1–25K frequency lists (Nation, 2017). As a first step, the lexical coverage of the texts was established to assess whether the texts are suitable for the intended learners. Here, each individual token (i.e. orthographic word) was counted and the percentage of words from each frequency band was provided. The underlying idea of the approach is that incidental acquisition of vocabulary from reading is only likely to occur if learners



know about 95% of the words already (cf. Laufer, 2020; Nation & Anthony, 2013). Moreover, as the textbooks are aimed at intermediate learners of English who are expected to be on at least A2 level at the beginning of secondary school (Skolverket, 2012), it was decided that the high-frequency words of English can be expected to already be known to the learners (see also Nordlund, Kokkola, & Rydström, *in press*; Snoder & Laufer, 2022, where it is indicated that Swedish students are likely to have even larger vocabularies than this). The analysis of the textbook content in relation to frequency bands is not reported band by band. Instead, the analysis followed Schmitt's and Schmitt (2014) division into high- (first 3,000 word families), mid- (3,001–9,000 word families) and low-frequency words (9,001–).<sup>5</sup> As the texts encompass a considerably higher proportion of proper nouns than general English (see results reported in Nation, 2006), these are presented in a category of their own in order not to skew the frequency distribution figures. The texts were also analysed on a lemma level, which means that a word and its inflectional forms were considered one item (e.g. *jump*, *jumps*, *jumping* and *jumped* are all represented by the lemma JUMP), using Sketch Engine. This analysis was included to map how many different words can be acquired in a text.

To investigate the extent of recycling in the textbooks, the raw frequency of lemmas was analysed using the word list function in Sketch Engine. Lemmas was used as the counting unit as it is likely that recycling of *sing* will occur also when encountering an inflected form, while it is less likely that encountering a derivational form (i.e. a word-family approach) would function as consolidation of the word (Reynolds & Wible, 2014; Reynolds, 2015). Following Webb's (2007), it was decided that a lemma had to occur ten times to be considered frequently recycled and thus presumably possible to learn from the exposure in the materials. Further, to ensure that words occurring frequently in only one or a few texts would not be considered, it was decided that the lemma had to occur in at least three different texts. For example, the lemma PEANUT occurs 66 times in the *Awesome* subcorpus, but all occurrences are from the same text. In the results section, the category 9< recurrences means a total of ten or more occurrences in three or more texts. The recycling analysis thus entailed first establishing the total number of lemmas in the subcorpora and then the number of lemmas with a raw frequency of ten or more, to be able to establish the extent of recycling in the materials. When investigating recycling, the analysis focused on the open word classes nouns, lexical verbs, adjectives and adverbs as these are primary carriers of meaning and important for vocabulary acquisition. The nine modal auxiliary verbs and the primary verbs *be*, *do* and *have* were thus excluded. Focusing on lexical word classes was also opted for to reach valid recycling measures as function words are few and very likely to recur in texts (cf. Biber, Johansson, Leech, Conrad, & Finegan, 2004). The list of lemmas occurring ten or more times were then analysed using Compleat Web VP (Cobb, n.d.) to map their frequency in general English. In this way, the analysis sought to establish how useful the recycled words are to students and whether they are suitable for learning at their level. Given the previously mentioned assumption about Swedish students' vocabulary sizes, the contention was that the textbooks should recycle mid-frequency words, if they are to support continued progression at this level.

As can be seen above, the lexical unit used in the analysis was primarily lemmas (except for the frequency distribution on a token level to assess 95% text coverage).

Within the vocabulary field, there is an ongoing debate concerning what counting unit is most suitable for measuring learners' vocabulary sizes, where both word families (e.g. Dang, 2021; Snoder & Laufer, 2022) and lemmas (e.g. Gablasova & Brezina, 2021; Kremmel, 2016) have been proposed. As mentioned above, lemma was chosen as the counting unit as this was judged to be more reliable than word families when assessing recycling, as inflectional variation does not impact repetition effects (Reynolds, 2015). However, to assess the results in a language learning context (i.e. whether the words are suitable for the intended learners), the lemmas were categorised in comparison to the BNC/COCA 1–25K list (Nation, 2017), which is a word-family based list. Moreover, Schmitt's and Schmitt (2014) division into high-, mid- and low-frequency vocabulary is also based on word families. As these lists are word-family based, words that in themselves would be of lower frequency can potentially receive a higher frequency value. For instance, *command* occurs in the second frequency band, together with *commandment*, which in a lemma-based word list occurs in the twelfth band. This means that comparing the results of this study with a lemma-based word list would have rendered somewhat different results. However, to enable an understanding of the results of the present study as concerns what words students should learn, a comparison with word family-based lists was necessary, as there are no corresponding lemma-based guidelines, to the best of the authors' knowledge. This means that the assessment of the suitability of the lexical profile of the texts is based on a word family-approach, although the counting unit used is lemmas. As going from lemmas to word families is a many-to-one mapping (Gablasova & Brezina, 2021), this should not be a serious methodological issue. It should be noted that this means that the study evaluates the results in relation to usefulness for learners, rather than in comparison to lemma frequency in general English.

### 3. Results

#### 3.1 The characteristics of the input

This section presents the analysis focused on characterising the lexical input (RQ1) and evaluating its suitability for the target students. To illuminate and give a better understanding of the results, the figures presented are compared to general language use and previous textbook studies. The first step in the analysis was to investigate the frequency distribution on a token level (see Table 2), in order to characterise the entirety of the language input that students encounter in the materials. As can be seen in the table, the proportion of high-frequency items is quite similar across the textbook series. *Gleerups* has the lowest proportion of high-frequency items (87.4%) and *Good Stuff Gold* the highest (91.4%). The proportion of high-frequency vocabulary is similar to general

**Table 2.** Percentage of total number of tokens per frequency level (and cumulative coverage).

	Awesome	Digilär	Gleerups	GS Gold	Wings
<i>High (1-3K)</i>	90.7	88.6	87.4	91.4	89.9
<i>Proper nouns</i>	4.8 (95.5)	5.5 (94.1)	5.3 (92.7)	4.2 (95.6)	4.8 (94.7)
<i>Mid (4-9K)</i>	2.9 (98.4)	3.4 (97.5)	2.7 (95.4)	2.5 (98.1)	3.5 (98.2)
<i>Low (10K-)</i>	1.5 (≈100)	2.6 (≈100)	4.6 (100)	1.8 (≈100)	1.8 (100)

English, where studies have indicated a coverage of 89–95% (Nation, 2006) or 90% (Nation & Anthony, 2013). In terms of mid-frequency vocabulary, the proportion ranges from 2.5% (*Good Stuff Gold*) to 3.5% (*Wings*), which is lower than for common language use, where 5% has been found to be a regular coverage level (Nation & Anthony, 2013; Nation, 2006). Moreover, as can be seen below, proper nouns constitute quite a substantial proportion of the tokens, both in comparison to other textbooks (1.69%; Yang & Coxhead, 2020) and general English (2–4%; Nation, 2006).

An important aspect of vocabulary learning from reading is that the input is adapted so that most words are already known to the learner (at least 95%; Nation & Anthony, 2013), as this facilitates acquisition of the remaining words. Considering that the target students are expected to be on A2 level when starting secondary school (Skolverket, 2012), it could be assumed that they already know the high-frequency words of English. Accordingly, texts aimed at these students should have at least 95% high-frequency tokens. As can be seen in Table 2, the proportion of the texts covered by high-frequency vocabulary and proper nouns is close to 95% for all materials except for *Gleerups* (92.7%). With the assumption that the students can comprehend proper nouns in context,<sup>6</sup> this means that the target students are likely to be able to read and comprehend the texts and develop vocabulary by reading them for content. Considering the estimated vocabulary size of the target students, the results indicate that the materials are suitable for the intended learners because they encompass an appropriate proportion of unknown vocabulary (i.e. target vocabulary).

While Table 2 illustrates whether the texts have a suitable vocabulary load on a token level, it does not show the number of different words that students can learn by reading them. Nation and Anthony (2013) point out that when adapting texts, token percentage is not always a good figure to consider, rather, the raw number of new vocabulary has to be examined as well to illuminate whether texts are suitable for the intended learners. Therefore, the frequency distribution was also calculated on a lemma level (see Table 3), where each lemma is counted once, highlighting how many different lemmas from high-, mid- and low- frequency bands that students are exposed to in the textbooks.

To evaluate whether the number of mid-frequency lemmas is sufficient for the target audience, an approximation of the vocabulary learning objective was considered. As pointed out by Nation and Coxhead (2021), an increase of about 1,000 word families per year is desirable for EFL learners, although this may be difficult (see Webb & Chang, 2012). It could regardless be expected that teaching materials provide opportunities for a desirable vocabulary growth. Moreover, as the students can be expected to know the high-frequency words, these lemmas should come from the mid-frequency bands. As can be seen in Table 3, the teaching materials encompass between 777 and 1,441 mid-frequency lemmas, which means that students receive input to this number of lemmas over a period of three years. Given that students should learn over 1,000

**Table 3.** Number of lemmas (and proportion of all lemmas) in frequency bands.

	Awesome	Digilär	Gleerups	GS Gold	Wings
<i>High (1-3K)</i>	3,612 (56.8)	2,767 (59.0)	3,317 (58.1)	3,213 (59.2)	3,629 (54.2)
<i>Proper nouns</i>	917 (14.4)	732 (15.6)	869 (15.2)	716 (13.2)	947 (14.1)
<i>Mid (4-9K)</i>	1,239 (19.5)	777 (16.6)	1,029 (18.0)	1,060 (19.5)	1,441 (21.5)
<i>Low (10K-)</i>	596 (9.4)	415 (8.8)	493 (8.6)	444 (8.2)	684 (10.2)

words per year to keep up in their vocabulary learning, these figures appear quite low. It could be expected that the amount of novel vocabulary from mid-frequency bands would be higher, considering the intended group of learners.

### 3.2 Recycling

While it is important that educational texts are adapted to the intended students' needs in terms of exposure, it is, as mentioned, also essential that they provide good learning opportunities via recycling. The following section presents the results relating to the recycling of words (RQ2) and the nature of the vocabulary recycled (RQ3). Ten or more encounters were set as the cut-off point for a lemma to be considered often recycled (cf. Webb, 2007) and hence possible to be learned incidentally from exposure in the materials. As can be seen in Table 4 below, between 10.7 and 15.6% of all lemmas are recycled to this extent. As mentioned, there is no general figure for what should be considered adequate recycling in teaching materials as this may differ depending on, for example, text length. Yet, these figures indicate that only a limited proportion of the vocabulary in the materials can potentially be acquired from reading the texts included. Further, over 70% of the lemmas occur four times or fewer in the materials, which is only about one time per textbook and year. This is a noteworthy finding because it means that acquisition of a very large proportion of the vocabulary input is not supported via recycling.

In relation to the different word classes, the lowest figures of sufficient recycling were found for nouns, where less than or about 10% of the lemmas recur ten times or more. Adjectives show similar patterns, with between 8.2 and 12.1% recycled lemmas (see Table 5). For lexical verbs and adverbs, on the other hand, more lemmas are recycled frequently. This is especially the case for adverbs where all series recycle about 20% of the lemmas ten or more times. The learning conditions between word classes thus differ

**Table 4.** Recycling of all lemmas in raw frequency and percentage.

Times <sup>7</sup>	All lemmas	
	1–4	9<
<i>Awesome</i>	4,812 (73.4)	889 (13.6)
<i>Digilär</i>	3,751 (78.0)	508 (10.7)
<i>Gleerups</i>	4,168 (72.0)	841 (14.5)
<i>GS Gold</i>	3,909 (70.5)	865 (15.6)
<i>Wings</i>	5,183 (74.3)	902 (12.9)

**Table 5.** Recycling of lemmas from lexical word classes in raw frequency and percentage.

Times	Nouns		Adjectives		Lexical verbs		Adverbs	
	1–4	9<	1–4	9<	1–4	9<	1–4	9<
<i>Awesome</i>	3,491 (77.8)	386 (8.6)	912 (80.4)	106 (9.3)	898 (72.6)	187 (15.1)	229 (67.8)	74 (21.9)
<i>Digilär</i>	2,742 (82.9)	200 (6.0)	645 (82.2)	64 (8.2)	563 (75.4)	85 (13.1)	154 (65.5)	50 (21.3)
<i>Gleerups</i>	3,020 (76.3)	362 (9.2)	759 (78.7)	105 (10.9)	719 (71.4)	174 (17.3)	225 (67.6)	72 (21.6)
<i>GS Gold</i>	2,829 (74.6)	395 (10.4)	730 (78.0)	113 (12.1)	776 (71.7)	176 (16.3)	216 (64.5)	76 (22.7)
<i>Wings</i>	3,739 (78.5)	403 (8.5)	978 (81.6)	113 (9.4)	918 (73.1)	176 (14.0)	279 (69.6)	78 (19.5)

substantially and students have more opportunities to acquire verbs and adverbs than the other lexical word classes.

It should, however, be noted that part of these figures could be attributed to the fact the number of lemmas in the word classes differs in the materials and it is therefore reasonable that word classes comprising fewer words recur more frequently. This may explain the high recycling figures for adverbs. It is, however, notable that lexical verbs show higher proportions of recycling than adjectives, although there are more verbs than adjective lemmas in the materials. Also, especially nouns are likely to be more dependent on the thematic content than the other word classes, which could explain the more varied use. A similar pattern was found by Nordlund (2016) in her study of EFL materials for young learners, where the materials frequently recycled a considerably higher proportion of lexical verbs than nouns and adjectives. This suggests that the word class differences should be interpreted with caution as they may, in part, be a result of the nature of English.

A finding that a certain percentage of the words in a text occur a particular number of times in a text does not in itself illuminate the vocabulary learning conditions of the material. It is also of relevance to understand the nature of the recurring vocabulary. When texts are adapted to learners' needs, it is not only important that a suitable proportion of the input is unknown vocabulary, but also that the vocabulary is recycled. Considering that the investigated teaching materials are aimed at intermediate students, it would be facilitative for vocabulary development if a substantial proportion of the recycled items is from the mid-frequency bands and not, as shown in Table 6 below, words from high-frequency bands, that is, words the students using the books are already likely to know.

Since high-frequency items are the most common words in English, it is not surprising that many items recurring ten or more times come from these bands. But figures over 90% show that the recycling of vocabulary is not structurally planned, as almost no frequently recycled vocabulary is beyond the target students' estimated level. As can be seen in Table 7 below, mid-frequency words only constitute between 0.5 and 1.4% of the commonly recycled items.

**Table 6.** Raw frequency and proportion of frequently recycled<sup>8</sup> lemmas in high-frequency bands.

	Awesome	Digilär	Gleerups	GS Gold	Wings
<i>All lemmas</i>	844 (93.9)	482 (94.9)	787 (93.6)	815 (94.6)	844 (93.6)
<i>Nouns</i>	346(89.9)	177 (88.5)	314 (86.7)	347 (88.3)	350 (86.9)
<i>Lexical verbs</i>	187 (100)	94 (98.9)	174 (100)	176 (100)	175 (99.5)
<i>Adjectives</i>	105 (99.1)	62 (96.9)	103 (98.1)	113 (100)	111 (98.3)
<i>Adverbs</i>	74 (100)	50 (100)	72 (100)	75 (98.7)	78 (100)

**Table 7.** Raw frequency and proportion of frequently recycled lemmas<sup>8</sup> in mid-frequency bands.

	Awesome	Digilär	Gleerups	GS Gold	Wings
<i>All lemmas</i>	4 (0.5)	7 (1.4)	8 (0.9)	9 (1.0)	5 (0.6)
<i>Nouns</i>	-	4 (2.0)	4 (1.1)	9 (2.3)	2 (0.4)
<i>Lexical verbs</i>	-	1 (1.1)	-	-	1 (0.6)
<i>Adjectives</i>	-	1 (1.6)	1 (1.0)	-	-
<i>Adverbs</i>	-	-	-	-	-

The number of mid-frequency words recycled ten or more times never exceeds ten in any material. This means that the materials do not provide opportunities for acquiring almost any mid-frequency words through reading the texts. Worth noticing is also that every teaching material has at least one word class where no mid-frequency words are recycled ten or more times. The figures could be compared to the results found in Sun and Dang's (2020) analysis of Chinese EFL textbooks, where 18.06% of the mid- and low-frequency words in the textbooks were recycled ten or more times. Their findings show that higher proportions of mid-frequency recycling in educational texts can occur. The extent to which mid-frequency words are recycled in the teaching materials thus appears surprisingly low considering the target students.

#### 4. Discussion

As vocabulary learning is a central aspect of language learning that requires structured support (Nation, 2006; Schmitt, 2008), the present study has used corpus-based methods to investigate the vocabulary component of texts in EFL teaching materials to find out whether they support learning in this respect. As mentioned, textbooks are widely used in the foreign language classroom and have the potential to enhance students' learning if they are constructed with a focus on how vocabulary is learned. The way textbooks are constructed is especially important as many EFL teachers do not consider vocabulary as a distinct learning objective, and therefore do not plan their teaching with vocabulary in mind (Bergström, Norberg, & Nordlund, 2022; Gao & Ma, 2011). The specific focus of the study presented here was on the incidental vocabulary learning opportunities in Swedish intermediate EFL materials, in terms of the nature of the lexical input and the recycling of mid-frequency words. The results indicate an ambiguous situation where, on the one hand, the texts appear adapted to the target students' expected vocabulary level, while, on the other hand, vocabulary on the target level is not recycled in the materials.

The lexical nature of the texts investigated suggests that the teaching materials do indeed facilitate vocabulary learning from reading, as the proportion of known vocabulary largely corresponds to proposed measures for graded readers (Nation & Anthony, 2013). The students using the materials are thus likely to comprehend the texts and acquire new words by reading them. In this respect, the materials appear appropriate for the target group. These findings are incongruent with previous studies where a main deficit of EFL materials identified that the input is too difficult for the students (e.g. Norberg & Nordlund, 2018; Sun & Dang, 2020), which, as suggested, can hamper and even have a negative effect on learning. It is important to note, however, that the lexical input in this study was compared to the minimum vocabulary level expected. There are most likely students using the material with a considerably larger vocabulary than the formal requirements for intermediate students aiming for B1, that is, the level to pass English in secondary school (see, for example, Snoder & Laufer, 2022). It is uncertain to what extent the materials challenge these students. Nevertheless, in terms of the lexical nature of the texts, the teaching materials appear quite suitable as support for vocabulary development at intermediate level.

However, for learning to take place, new words need to be encountered many times. The analysis showed that the mid-frequency words in the materials are not recycled



sufficiently for vocabulary development to take place from reading the texts. It was found that almost only high-frequency vocabulary, that is, words that students are likely to know already, is recycled sufficiently in the material. (e.g. Webb, 2007). As it is almost impossible for a teacher to ensure recycling of all target vocabulary in the classroom (Schmitt, 2019), the finding that mid-frequency vocabulary is not given principled attention is noteworthy, as it means that these words are not likely to stay in the students' long-term memory. Similar results have been reported by Matsuoka and Hirsh (2010), who go as far as to conclude that the material they studied is unsuitable as support for vocabulary development (see also Yang & Coxhead, 2020). An absence of systematicity in recycling of mid-frequency vocabulary has also been identified in previous Swedish studies (e.g. Nordlund, 2015a, b). Taken together, this is an indication that low levels of recycling is a widespread phenomenon in EFL materials globally and regardless of the intended age group. It should be noted, however, that the analysis focuses on recycling in the texts in the materials and does not take additional recycling in, for instance, word lists and exercises into consideration. Nevertheless, the incidental vocabulary conditions appear insufficient for the student group.

The results of this study undoubtedly show that recycling is not paid attention to in the construction of texts for EFL materials, which means that books intended for intermediate learners are not in line with the vocabulary learning objectives for that particular level of learning. It could, however, be argued that many students engage in extramural English activities (cf. Sundqvist, 2009), where it is likely that vocabulary at target level is repeatedly encountered. For these students, the teaching material is not the main source of input, which could decrease the negative impact. Yet, as school has a responsibility for all students' learning, it appears reasonable to expect teaching materials to provide adequate vocabulary learning opportunities, especially as aspects such as recycling are difficult to ensure otherwise (cf. Schmitt, 2019). For students not involved in extramural activities, it is of utmost importance that the input provided in the classroom suffice for vocabulary development. As the results indicate that this is not the case, a possible implication is that the gap between learners with additional English exposure (e.g. in their spare time activities) and the ones without increases over time, especially as a bigger vocabulary facilitates further learning (see e.g. Elgort, Perfetti, Rickles, & Stafura, 2015; Stanovich, 2009).

The study is based on a limited data set and illuminates incidental learning opportunities from reading the texts in materials, that is, the analysis focuses on some aspects of the textbooks but does not take all vocabulary learning opportunities into consideration, such as exercises. The findings are yet indicative and suggest a lack of systematicity regarding the vocabulary component of EFL textbooks. As mentioned, the results show, like previous research (e.g. Nordlund, 2016; Yang & Coxhead, 2020), that teaching materials are not constructed in accordance with research concerning how students' vocabulary development can be facilitated. While vocabulary researchers have presented principles and guidelines for foreign language education (e.g. Nation, 2011), these appear to have limited impact on materials development. For this reason, it is important that authors are made aware of and learn to use tools to analyse their texts from a vocabulary perspective, such as online vocabulary profilers, and that vocabulary is given prominence in, for instance, explicit guidelines from publishers. Moreover, the results also indicate that teachers need to pay attention to mid-frequency words and



recycling and not rely on the exposure in teaching materials. Teachers can, for example, seek to ensure additional repetitions of important words that they identify in the texts by bringing them up in class or assigning extra exercises.

The study has shown the value of using corpus-based methods when examining language learning materials. As the textual aspects illuminated by these methods, such as recycling, comparison between texts and overarching patterns, cannot be seen by the individual teacher using a material, studies of this kind contribute with additional perspectives on the language classroom. Further corpus-based studies of teaching materials are hence encouraged and they could, for example, examine materials from other educational contexts or the progression over several stages of learning. Moreover, the intentional vocabulary learning opportunities could also be focused on by, for instance, studying the explicit vocabulary exercises or word lists, which is an area yet largely unstudied (see, however, Brown, 2011; Nordlund & Norberg, 2020). Finally, to understand the role of the textbook for students' language learning, more studies investigating how teachers use teaching materials for vocabulary purposes would be useful.

## 5. Conclusion

The present study has shown that the vocabulary component of texts in Swedish EFL materials is not structured in a way that supports incidental vocabulary learning from reading. It is true that the materials studied have some beneficial aspects for vocabulary learning, such as having an appropriate proportion of known words, but the positive effects are likely to be reduced as recycling is not dealt with systematically. Accordingly, the support provided from the teaching materials for students' incidental vocabulary development is likely to be very limited. This may have negative consequences for the students' language learning, especially if they do not engage in English activities outside of school. It can be concluded that if teaching materials are to function as tools in the language classroom, vocabulary needs to be given a more prominent place during the materials design process. Moreover, it also appears necessary that constructing materials based on vocabulary research becomes a prioritised matter for publishers.

## Notes

1. A word family encompasses a base word and its inflectional and derivational forms (e.g. the verb *paint* and its inflectional forms *painting*, *painted*, *paints* as well as derivations such as *painter* belong to the same word family).
2. The digital teaching material *Digilär* provides two versions of each text, one called Standard and one called Advanced. Here, the Standard version of the texts was analysed.
3. *Digilär* was accessed between 9 December 2020 and 8 January 2021. *Gleerups Engelska* was accessed between 25 January and 4 February 2021.
4. <http://www.sketchengine.eu>.
5. Words that are not covered by the first 25,000 word families in the BNC and COCA, so-called off-list words, were grouped with the low-frequency words.
6. It should be noted, however, that although vocabulary researchers generally work under this assumption, this has been questioned more recently, by, for example, Klassen (2022) who argues that the assumption lacks empirical support and that some proper nouns are likely to pose problems for L2 readers.

7. 5–9 occurrences have not been included in the results presentation as the primary focus is on words that occur frequently in the materials.
8. 10 or more occurrences in the corpus.

## Notes on contributors

*Denise Bergström* is PhD student in English and Education at Luleå University of Technology, Sweden. Her main research interest has a focus on EFL vocabulary learning and teaching.

*Cathrine Norberg* is Professor of English at Luleå University of Technology, Sweden. Her main research interest is corpus linguistics with a focus on gender, and educationally related issues.

*Marie Nordlund* is Senior Lecturer in English at Luleå University of Technology, Sweden. Her main research interest lies within corpus linguistics and applied linguistics.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## ORCID

Denise Bergström  <http://orcid.org/0000-0002-7098-213X>

Cathrine Norberg  <http://orcid.org/0000-0001-9170-1459>

Marie Nordlund  <http://orcid.org/0000-0002-6004-4115>

## Primary sources

- Childs-Cutler, A., & Gentili Cronholm, A. (2017). *Awesome 8 textbook*. Stockholm: Sanoma utbildning.
- Childs-Cutler, A., & Gentili Cronholm, A. (2018). *Awesome 9 textbook*. Stockholm: Sanoma utbildning.
- Childs-Cutler, A., Gentili Cronholm, A., & Niemann, J. (2016). *Awesome 7 textbook*. Stockholm: Sanoma utbildning.
- Coombs, A., Bayard, A., Hagvörn, R., & Johansson, K. (2012). *Good stuff gold B textbook*. Stockholm: Liber.
- Coombs, A., Bayard, A., Hagvörn, R., & Johansson, K. (2013). *Good stuff gold C textbook*. Stockholm: Liber.
- Coombs, A., Bayard, A., Hagvörn, R., & Johansson, K. (2014). *Good stuff gold D textbook*. Stockholm: Liber.
- Frato, K., Cederwall, A., Rinnesjö, S., Glover, M., Glover, R., Hedberg, B., & Malmberg, P. (2016). *Wings 8 textbook* (2nd ed.). Stockholm: Natur & Kultur.
- Frato, K., Cederwall, A., Rinnesjö, S., Glover, M., Glover, R., Hedberg, B., & Malmberg, P. (2017). *Wings 9 textbook* (2nd ed.). Stockholm: Natur & Kultur.
- Frato, K., Mellerby, A., Rinnesjö, S., Glover, M., Glover, R., Hedberg, B., & Malmberg, P. (2015). *Wings 7 textbook* (2nd ed.). Stockholm: Natur & Kultur.
- Gode, L. (2014). *Digilär engelska 7–9* [Digilär English 7–9] (Stockholm: Natur & Kultur). Last accessed 8 January 2021.
- Taylor, J., Carlsson, M., & Smith, M. (n.d.). *Gleerups engelska 7–9* [Gleerups English 7–9] (Malmö: Gleerups). Last accessed 4 February 2021.

## References

- Alsaif, A., & Milton, J. (2012). Vocabulary input from school textbooks as a potential contributor to the small vocabulary uptake gained by English as a foreign language learners in Saudi Arabia. *The Language Learning Journal*, 40(1), 21–33. doi:10.1080/09571736.2012.658221
- Bergström, D., Norberg, C., & Nordlund, M. (2021). “The Text Comes First” – principles guiding EFL materials developers’ vocabulary content decisions. *Scandinavian Journal of Educational Research*, 1–15. Advance online publication. doi:10.1080/00313831.2021.1990122
- Bergström, D., Norberg, C., & Nordlund, M. (2022). “Words are picked up along the way” – Swedish EFL teachers’ conceptualizations of vocabulary knowledge and learning. *Language Awareness*, 31(4), 393–409. doi:10.1080/09658416.2021.1893326
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (2004). *Longman grammar of spoken and written English* (4<sup>th</sup> ed.). Harlow: Longman.
- Brown, D. (2011). What aspects of vocabulary knowledge do textbooks give attention to? *Language Teaching Research*, 15(1), 83–97. doi:10.1177/1362168810383345
- Brown, D. (2021). Incidental vocabulary learning in a Japanese university L2-English language classroom over a semester. *TESOL Journal*, 12(4), e595. doi:10.1002/tesj.595
- Chen, C., & Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics*, 31(5), 693–713. doi:10.1093/applin/amq031
- Cobb, T. (n.d.) Compleat Web VP (Version 2.5) [Computer software]. <https://www.lexutor.ca/vp/comp/>
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Strasbourg: Council of Europe, Modern Languages Division.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213–238. doi:10.2307/3587951
- Coxhead, A., & Boutorwick, T. J. (2018). Longitudinal vocabulary development in an EMI international school context: Learners and texts in EAL, maths, and science. *TESOL Quarterly*, 52(3), 588–610. doi:10.1002/tesq.450
- Coxhead, A., Stevens, L., & Tinkle, J. (2010). Why might secondary science textbooks be difficult to read? *New Zealand Studies in Applied Linguistics*, 16(2), 37–52.
- Dang, T. N. Y. (2021). Selecting lexical units in wordlists for EFL learners. *Studies in Second Language Acquisition*, 43(5), 954–957. doi:10.1017/S0272263121000681
- Elgort, I., Perfetti, C. A., Rickles, B., & Stafura, J. Z. (2015). Contextual learning of L2 word meanings: Second language proficiency modulates behavioural and event-related brain potential (ERP) indicators of learning. *Language, Cognition and Neuroscience*, 30(5), 506–528. doi:10.1080/23273798.2014.942673
- Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2), 143–188. doi:10.1017/S0272263102002024
- Ellis, N. C., & Wulff, S. (2015). Usage-based approaches to SLA. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (2<sup>nd</sup> ed., pp. 75–93). New York, NY: Routledge.
- Gablasova, D., & Brezina, V. (2021). Words that matter in L2 research and pedagogy: A corpus-linguistics perspective. *Studies in Second Language Acquisition*, 43(5), 958–961. doi:10.1017/S027226312100070X
- Gao, X., & Ma, Q. (2011). Vocabulary learning and teaching beliefs of pre-service and in-service teachers in Hong Kong and mainland China. *Language Awareness*, 20(4), 327–342. doi:10.1080/09658416.2011.579977
- Gray, J. (2016). ELT materials: Claims, critiques and controversies. In G. Hall (Ed.), *The Routledge handbook of English language teaching* (pp. 95–108). London: Routledge.
- Hermagustiana, I., Hamra, A., Rahman, Q., & Salija, K. (2017). Reflection of teacher cognition in EFL vocabulary instructional practices in Indonesia. *International Journal of English Linguistics*, 7(5), 34–45. doi:10.5539/ijel.v7n5p34
- Hughes, S. H. (2019). Coursebooks: Is there more than meets the eye? *ELT Journal*, 73(4), 447–455. doi:10.1093/elt/ccz040

- Hulstijn, J. H. (2003). Incidental and intentional learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 349–381). Malden, MA: Blackwell.
- Hu, M., & Nation, I. S. P. (2000). Vocabulary density and reading comprehension. *Reading in a Foreign Language*, 23(1), 403–430.
- Kilgariff, A., Rychlý, P., Smrž, P., & Tugwell, D. (2004). The sketch engine. In G. Williams & S. Vessier (Eds.), *Proceedings of the eleventh Euralex congress* Lorient, France (pp. 105–116). Université de Bretagne-Sud.
- Klassen, K. (2022). Proper name theory and implications for second language reading. *Language Teaching*, 55(2), 149–155. doi:10.1017/S026144482100015X
- Kremmel, B. (2016). Word families and frequency bands in vocabulary tests: Challenging conventions. *TESOL Quarterly*, 50(4), 976–987. doi:10.1002/tesq.329
- Laufer, B. (2020). Lexical coverages, inferencing unknown words and reading comprehension: How are they related? *TESOL Quarterly*, 54(4), 1076–1085. doi:10.1002/tesq.3004
- Ljung, M. (1990). *A study of TEFL vocabulary*. Stockholm: Almqvist & Wiksell.
- Matsuoka, W., & Hirsh, D. (2010). Vocabulary learning through reading: Does an ELT course book provide good opportunities? *Reading in a Foreign Language*, 22(1), 56–70.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63(1), 59–82. doi:10.3138/cmlr.63.1.59
- Nation, I. S. P. (2011). Research into practice: Vocabulary. *Language Teaching*, 44(4), 529–539. doi:10.1017/S0261444811000267
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2<sup>nd</sup> ed.). Cambridge: Cambridge University Press.
- Nation, I. S. P. (2016). Word lists. In I. S. P. Nation (Ed.), *Making and using word lists for language learning and testing* (pp. 3–22). Amsterdam: John Benjamins.
- Nation, I. S. P. (2017). The BNC/COCA Level 6 word family lists (Version 1.0.0) [Data file]. Available from <https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/vocabulary-lists>
- Nation, I. S. P., & Anthony, L. (2013). Mid-frequency readers. *Journal of Extensive Reading*, 1(1), 5–16.
- Nation, I. S. P., & Coxhead, A. (2021). *Measuring native-speaker vocabulary size*. Amsterdam: John Benjamins.
- Newton, J. (2020). Approaches to learning vocabulary inside the classroom. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 255–270). London: Routledge.
- Norberg, C., & Nordlund, M. (2018). A corpus-based study of lexis in L2 English textbooks. *Journal of Language Teaching and Research*, 9(3), 463–473. doi:10.17507/jltr.0903.03
- Nordlund, M. (2015a). Vocabulary acquisition and the textbook: What opportunities are there for young learners? *ITL – International Journal of Applied Linguistics*, 166(2), 199–228. doi:10.1075/itl.166.2.01nor
- Nordlund, M. (2015b, 1). Vocabulary in textbooks for young learners. *Educare*, 2015, 73–92.
- Nordlund, M. (2016). EFL textbooks for young learners: A comparative analysis of vocabulary. *Education Inquiry*, 7(1), 47–68. doi:10.3402/edui.v7.27764
- Nordlund, M., Kokkola, L., & Rydström, U. (in press). Teenagers' vocabulary knowledge and self-efficacy as elements in effective L2 online searches. *Nordic Journal of English Studies*, 21 (2).
- Nordlund, M., & Norberg, C. (2020). Vocabulary in EFL teaching materials for young learners. *International Journal of Language Studies*, 14(1), 89–116.
- Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do *Things Fall Apart*? *Reading in a Foreign Language*, 22(1), 31–55.
- Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, 18(1), 1–28.
- Prowse, P. (2011). How writers write. In B. Tomlinson (Ed.), *Materials development in language teaching* (pp. 155–173). Cambridge: Cambridge University Press.
- Qian, D. D., & Lin, L. H. F. (2020). The relationship between vocabulary knowledge and language proficiency. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 66–80). London: Routledge.

- Reynolds, B. L. (2015). The effects of word form variation and frequency on second language incidental vocabulary acquisition through reading. *Applied Linguistics Review*, 6(4), 467–497. doi:10.1515/applirev-2015-0021
- Reynolds, B. L., & Wible, D. (2014). Frequency in incidental vocabulary acquisition research: An undefined concept and some consequences. *TESOL Quarterly*, 48(4), 843–861. doi:10.1002/tesq.197
- Sakata, N. (2019). Profiling vocabulary for proficiency development: Effects of input and general frequencies on L2 learning. *System*, 87, Article 102167. doi:10.1016/j.system.2019.102167
- Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329–363. doi:10.1177/1362168808089921
- Schmitt, N. (2019). Understanding vocabulary acquisition, instruction, and assessment: A research agenda. *Language Teaching*, 52(2), 261–274. doi:10.1017/S0261444819000053
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, 95(1), 26–43. doi:10.1111/j.1540-4781.2011.01146.x
- Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 47(4), 484–503. doi:10.1017/S0261444812000018
- Schmitt, N., & Schmitt, D. (2020). *Vocabulary in language teaching*. New York, NY: Cambridge University Press.
- Skolverket. (2012). *Bedömning av språklig kompetens. En studie av samstämmigheten mellan Internationella språkstudien 2011 och svenska styrdokument [Assessment of linguistic competence. A study on the correspondence between the European Survey on Language Competences 2011 and Swedish steering documents]*. Stockholm: Skolverket.
- Snoder, P., & Laufer, B. (2022). EFL learners' receptive knowledge of derived words: The case of Swedish adolescents. Advance online publication *TESOL Quarterly*, 56(4), 1242–1265. doi:10.1002/tesq.3101
- Stanovich, K. E. (2009). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Journal of Education*, 189(1–2), 23–55. doi:10.1177/0022057409189001-204
- Sun, Y., & Dang, T. N. Y. (2020). Vocabulary in high-school EFL textbooks: Texts and learner knowledge. *System*, 93, Article 102279. doi:10.1016/j.system.2020.102279
- Sundqvist, P. (2009). *Extramural English matters: Out-of-school English and its impact on Swedish ninth graders' oral proficiency and vocabulary* [Doctoral dissertation, Karlstad University].
- Taylor, A., Marcus, M., & Santorini, B. (2003). The Penn Treebank: An overview. In A. Abeillé (Ed.), *Treebanks* (pp. 5–22). Heidelberg: Springer.
- Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language Learning*, 69(3), 559–599. doi:10.1111/lang.12343
- Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, 15(2), 130–163.
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28(1), 46–65. doi:10.1093/applin/aml048
- Webb, S. (2014). Repetition in incidental vocabulary learning. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1–6). Wiley. doi:10.1002/9781405198431.wbeal1425
- Webb, S. (2020). Incidental vocabulary learning. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 225–239). London: Routledge.
- Webb, S. A., & Chang, A. C. S. (2012). Second language vocabulary growth. *RELC Journal*, 43(1), 113–126. doi:10.1177/0033688212439367
- Webb, S., & Nation, I. S. P. (2017). *How vocabulary is learned*. Oxford: Oxford University Press.
- Webb, S., & Rodgers, M. P. H. (2009). The lexical coverage of movies. *Applied Linguistics*, 30(3), 407–427. doi:10.1093/applin/amp010
- Yang, L., & Coxhead, A. (2020). A corpus-based study of vocabulary in the New Concept English textbook series. Advance online publication *RELC Journal*, 53(3), 597–611. doi:10.1177/0033688220964162