Review of Recent Scholarship

Applications of word lists in second language learning and teaching

Thi Ngoc Yen Dang
University of Leeds, UK
Stuart Webb
University of Western Ontario, Canada

Abstract

Although *word lists* have generated a great deal of attention from researchers, there has been no comprehensive review of the applications of word lists in second language learning and teaching. This article reviews the development, validation, and applications of 50 word list studies that were published and discussed in major international peer-reviewed Applied Linguistics and TESOL journals from 2013 to 2023.

It shows that the methodology of word list development and validation has become more sophisticated and word list developers can see many potential applications of their lists in research and pedagogy. However, most applications of recently developed word lists have been restricted to Nation's (2016) BNC/COCA lists, and little is known about the degree to which most word lists have been used in pedagogical contexts. Our review indicates several directions for future research on word lists, including exploring the impact of published lists on pedagogy, replicating word list studies for learners in underrepresented contexts, and developing sustainable, low-cost methods of developing word lists to allow teachers and learners to create lists serving their own needs.

Keywords: word lists; corpus linguistics; vocabulary; academic vocabulary; technical vocabulary; general service vocabulary

1. Introduction

Word lists have a long history. Research on word lists began to generate attention in 1953 when Michael West created the General Service List (GSL) and has flourished since 2000 when Averil Coxhead developed the Academic Word List (AWL). In the last ten years alone, a large number of word lists have been created and published. Word lists is a key theme in books (e.g., Nation, 2016) and edited volumes (e.g., Webb, 2020) about teaching and learning vocabulary. However, there are no comprehensive reviews of the applications of word lists to language learning and teaching. This is surprising because researchers typically justify the

creation of new word lists by highlighting their value to learners. Moreover, the value and validity of lists may not be transparent to teachers, learners, and researchers, which hinders the chance for word lists to be implemented in pedagogical contexts.

This review is a timely response to these issues. Covering 50 publications from the last decade, it will identify current themes and issues in research on word lists for second language learning and teaching. It aims to raise awareness of (a) the value of word lists for language research and pedagogy and (b) the areas that need further attention in word list research. The review will focus on studies of word lists for learners of English as a second or foreign language. To provide a comprehensive overview of word list research, it will cover studies that are directly related to word lists (word list development and validation), as well as those using word lists to inform language learning, teaching, and assessment. To offer readers a clear overview of current trends in word list research, we will only review works published in the last 10 years (2013-2023). However, where necessary, we will link studies to seminal works such as West's (1953) General Service List and Coxhead's (2000) Academic Word List. Although a large number of word lists have been developed in the last decade, this review will focus on 50 word list studies that have been published and discussed in major international peer-reviewed Applied Linguistics and TESOL journals. This approach allows us to better identify the trends in word list studies that have generated interest over the last decade (2013-2023). References for these lists are presented in Appendix 1.

2. What kinds of word lists have been published in the last decade?

Figure 1 presents the publication dates of the 50 word list studies published in the last 10 years. After reaching a peak of eight studies published in 2018, slightly fewer studies have been published since then with only 3 published in 2022, and 2 lists published in 2019 and 2023. This suggests that there may now appear to be less of a need for new lists, or that *word lists* as a topic is not generating as much interest from researchers and publishers.

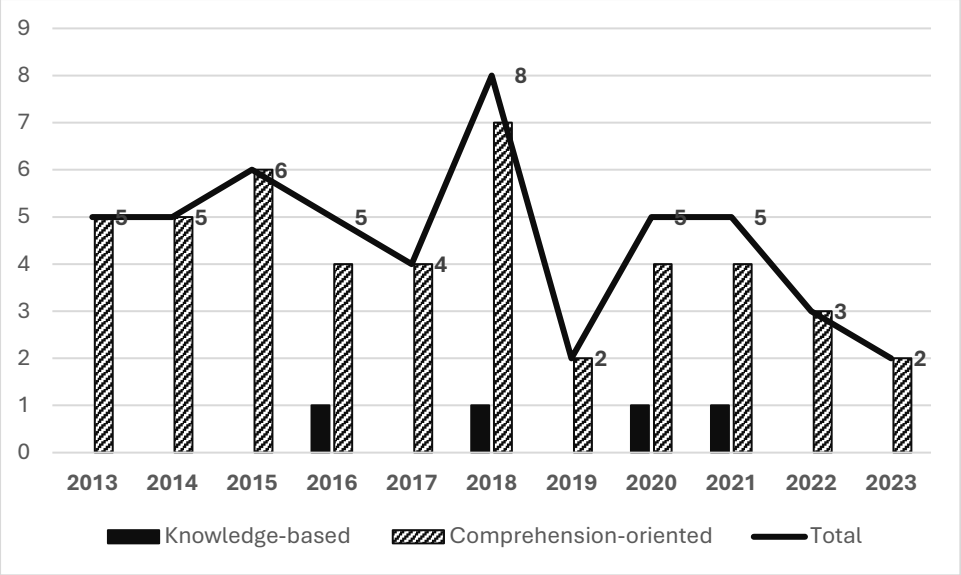

***Figure 1. Number of word list studies published in the last decade* (*N* = 50)**

Word lists can be classified into subtypes according to their purpose (general, academic, and technical), modality (written, spoken, and digital), and lexical components (single word and multiword). In terms of purpose, word lists can be categorized into general service lists (aiming to help learners to deal with the lexical challenges of different communicative situations in everyday language), academic vocabulary lists (aiming to help learners to cope with the lexical challenges of communication in a wide range of academic disciplines), and technical vocabulary lists (aiming to help learners to cope with the lexical challenges of communication in a specific discipline or professional area). Although general words and academic words still received attention from word list developers in the last decade, there was growing interest in creating lists of technical vocabulary (Figure 2). In fact, 66% of the word list studies published in the last decade focused on technical vocabulary whereas academic word lists and general word lists only accounted for 20% and 14%, respectively. The increasing number of studies developing technical lists could be the result of widespread use of English as the medium of communication in academic and professional settings.

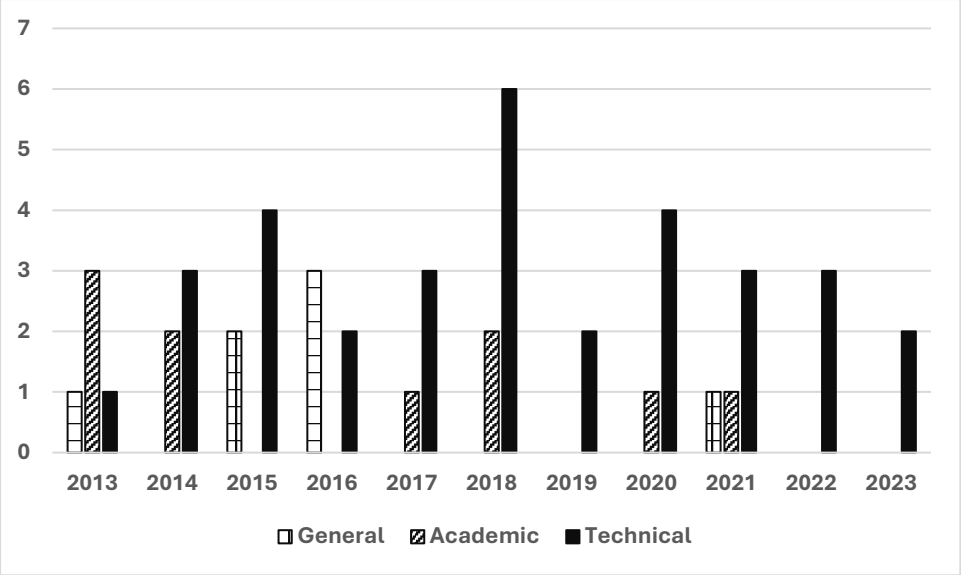

Figure 2. Word lists studies classified by list purpose (*N = 50)*

As for modality, written discourse is still dominant, being the focus of 58% of the word list studies (Figure 3). Yet increasing attention has been given to spoken and digital discourse. Since 2013, almost every year, at least one study has created lists to help learners to deal with the challenge of communication in either spoken discourse alone or both spoken and written discourse. Since 2022, word lists designed to support comprehension in digital discourse have appeared. In fact, a reasonable percentage of studies published in the last decade have developed lists from written plus spoken discourse (24%), spoken discourse alone (14%), and digital discourse (4%). The growing attention to spoken and digital discourse in word list research acknowledges the importance of comprehending spoken communication and recognition of the increasing need to understand digital discourse during and post the COVID pandemic period. This trend may be thanks to the findings of corpus-driven studies revealing that linguistic patterns in spoken discourse are different from those in written discourse to some degree (Biber, 2006) and that knowledge of items from written word lists may not always be sufficient for learners to deal with the lexical demands of spoken discourse (Dang et al., 2021). The growth in the number of spoken and digital word lists also reflects a growing number of second language acquisition (SLA) studies showing that vocabulary is learned from exposure to aural input and audiovisual input (see Reynolds, 2023). It is also the result of an increase in spoken and multimodal corpora, which makes it less challenging to develop word lists that are representative of these modes of communication.

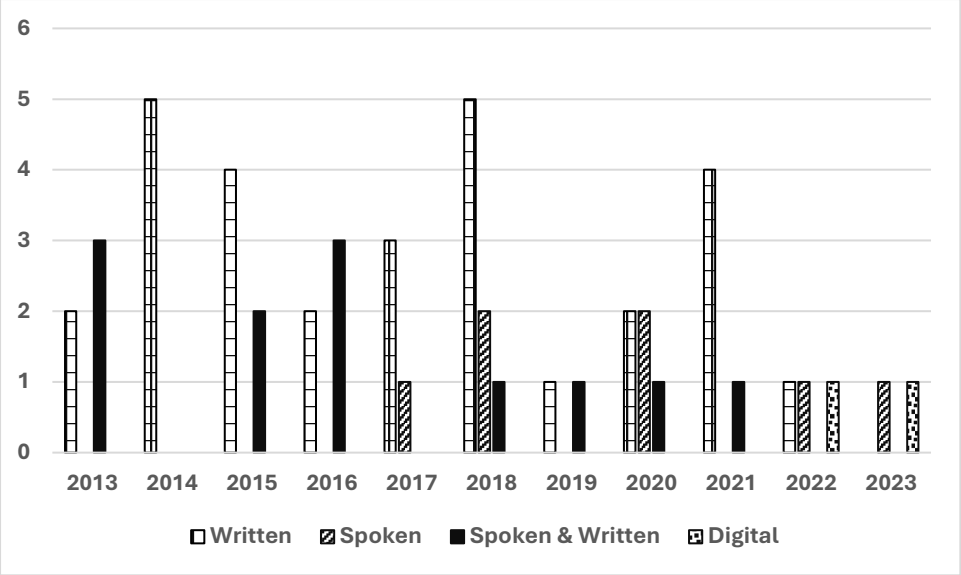

Figure 3. Word list studies classified by modality (N = 50)

In terms of lexical components, although most word lists published in the last decade (68%) examined single words, two major changes occurred. First, there was an increase in the number of lists of multiword sequences (24%) and lists made up of both multiword sequences and individual words (8%) (see Figure 4). This reflects a relatively recent trend in SLA research, which has emphasized the importance of knowledge of multiword sequences for second language (L2) communication (see Boers & Webb, 2018). Second, while a reasonable percentage of lists of single words (23.68%) still employed level 6 word families (e.g., Greene & Coxhead, 2015), none of the lists developed after 2020 chose this lexical unit. In fact, word lists using lemmas/flemmas (e.g., Dang & Webb, 2016) or word types (e.g., Bancroft-Billings, 2020) made up a slightly higher percentage of published lists (28.95% each). The level 6 word family counts a base form (*e.g., assure),* its inflected forms (e.g., *assures, assured, assuring*), its derived forms up to Bauer and Nation's (1993) affixes (e.g., *assurance, assuredly, reassurance, reassuring, reassuringly)* as one lexical unit. The lemma/flemma counts a base form (e.g., *assure*) and its inflected forms (e.g., *assures, assured, assuring*) as one lexical unit. The lemma counts word forms of different parts of speech as separate units but the flemma counts them as one unit. The word type counts unrepeated word forms as separate units. Other new lexical units were used in one study each: Level 3 partial word families in Nation (2016) and Nuclear word families in Cobb and Laufer (2021). The level 3 Partial word family includes Level 2 flemmas and four derivational affixes from Level 3 (*un-, -ly, -th,* and *-er*). The Nuclear word family counts the

base form of a Level 6 word family (*e.g., assure)* and the inflected forms and derived form made up of core affixes (e.g., *assurance, assured, reassuring*). Moreover, since Gardner and Davies's (2014) Academic Vocabulary List, a growing number of studies (13.16%) have developed parallel versions of word lists which employ different lexical units.

Also, 6.52% of the word lists (e.g., Coxhead & Demecheleer, 2018) used the word type as the primary lexical unit and then added any Level 6 family members of that word type which has technical meanings to the list. This variation in the lexical unit of word lists reflects a need to use different units of counting words for different target users (see Webb, 2021).

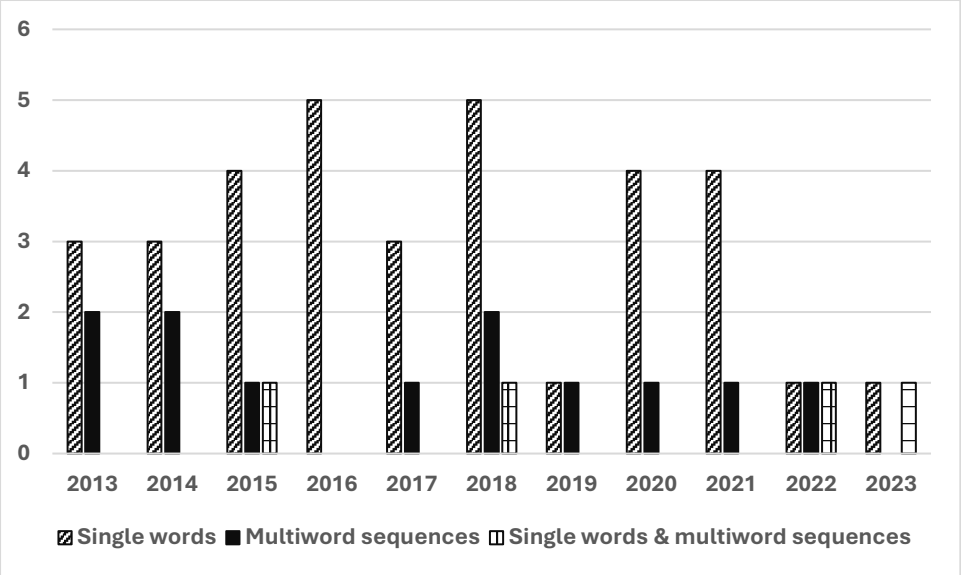

Figure 4. Word list studies classified by lexical component (N = 50)

Another trend revealed by Figure 1 is that most studies (92%) developed comprehension-oriented word lists (i.e., lists to enhance learners' comprehension of target communicative situations). However, since 2016, a small number of studies have developed knowledge-based word lists (8%) to represent the vocabulary that learners may already know. This indicates a pathway which may broaden word list research. While knowledge-based word lists may have value in revealing the words that are more likely to be known, the number of these lists is limited, and all of them have appeared recently. As a result, the impact of these lists is not clear yet. Given that most word lists published in the last decade are comprehension-oriented, the rest of this article will focus on the development, validation, and applications of comprehension-oriented lists.

3. How has the development and validation of comprehension-oriented word lists changed over the last decade?

3.1. Who are the target learners?

Identifying the target learners of word lists is important because it will influence the decision on all steps of word list construction, validation, and dissemination (Nation, 2016). Although all word lists developed in the last decade provided information about their targeted learners, this information is relatively general. Of the seven general service lists, six briefly mentioned their learners' language proficiency, educational levels, morphological knowledge, and/or first language (L1) backgrounds, and one (Browne, 2013) provided very vague information (L2 learners of English). Likewise, of the nine academic word lists, four (e.g., Chon & Shin, 2013) did not explicitly mention target learners and five (e.g., Ackerman & Chen, 2013) indicated that their learners were university English for Academic Purposes (EAP) students but did not provide further information.

Similarly, although all 30 technical word lists stated the disciplines of their target users, relatively few studies provided additional information. Studies mentioned learners' educational levels (60%), the country where learners took the language courses or used the language for specific purposes (33%), the kind of English courses (13%), and learners' prior vocabulary knowledge (3.33%). The modest information provided about target learners makes it difficult for teachers, learners, and policy makers to judge whether these lists are relevant to their context or not.

Our review also shows that little attention has been paid in word list research to learners from less privileged backgrounds, especially those who would like to use English for professional purposes. All academic word lists and 60% of the technical word lists aimed to support university students and scholars. Although some technical word lists have been developed for professionals (23.33%) and pre-university students (13.33%), they either focused on high-skilled professionals (business, trade, aviation, rugby) or pupils at international schools or middle/secondary schools in Global North countries (US, UK, Singapore, Germany).

Therefore, research on academic and technical word lists seems to be biased towards the needs of privileged groups. This tendency could be because in the past these groups were the most likely to use English for work and education. However, with globalization, a growing number of people from less privileged groups (e.g., street vendors at tourist attractions in the

Global South) need to use English at work and could benefit from word lists developed to serve their professional purposes.

3.2. How many items are in word lists?

While word lists should be long enough to cover key items that need to be known for comprehension, they should be short enough to fit learners' limited learning time. However, most general service lists and academic word lists published in the last decade are relatively long. A total of 71.43% of general service lists (e.g., Brezina & Gablasova, 2015) and 44.44% of academic word lists (e.g., Ackerman & Chen, 2013) had 2,000-9,000 items.

Although none of the technical word lists have more than 2,000 items, 20% of them still have from 1,000 to 2,000 items (e.g., Coxhead & Demecheleer, 2018; Green & Lambert, 2019). Several attempts have been made to break sizable lists into smaller sub-lists (e.g., Coxhead et al., 2019; Dang & Webb, 2016) or filter items in existing lists so that they have sizes that are more manageable to a specific learner population (e.g., Cobb & Laufer, 2021). Yet the large sizes of most existing lists may cause difficulty in implementing them into pedagogy.

3.3. What corpora have been used to develop word lists?

Word lists published in the last decade have benefited from the advances in corpus linguistics. In fact, all 46 comprehension-oriented lists were corpus-based. The availability of mega corpora (e.g., Davies's (2008) Corpus of Contemporary American English (COCA) Nesi and Thompson's (2006) British Academic Spoken English Corpus) and open resources (e.g., Massive Open Online Courses) have made developing large corpora for word list research less challenging than it used to be.

Corpora used to develop lists of general service vocabulary, academic vocabulary, and technical vocabulary have respectively reached 12 billion words (Brezina & Gablasova, 2015), 288 million words (Browne et al., 2013), and 18 million words (Greene & Coxhead, 2015). These corpora were much larger than those used to develop West's (1953) GSL (5 million words) and Coxhead's AWL (3.5 million words), and larger or comparable to the size needed for reliable lists of the most frequent 9,000 word types suggested by Sorell (2013).

Large corpora provide researchers with rich data so that they can better identify the lexical items that occur frequently in the target discourse. However, larger corpora do not necessarily ensure better word lists if the corpora do not represent the language that target learners are

likely to encounter. Corpus representativeness includes situational representativeness and linguistic representativeness (Biber, 1993; Egbert et al., 2022, cited in Miller, 2022). Situational representativeness refers to the extent to which corpora reflect the kinds of text and topics in the target domain whereas linguistic representativeness refers to the extent to which corpora reflect the distribution of lexical patterns in these domains.

Situational representativeness was considered by 95.65% of the reviewed studies. However, the situational representativeness of corpora in some studies is questionable. For example, the corpora used to create lists of general service vocabulary often underrepresents spoken language. Half of these lists (e.g., Brezina & Gablasova, 2015) were developed from corpora made up of largely written texts. Moreover, corpora for word lists to help learners deal with spoken communication in professional settings are very small in size, ranging from 4,157 words (Drayton & Coxhead, 2023) to 133,093 words (Coxhead & Demecheleer, 2018). The modest proportion of spoken texts in corpora for word lists, especially those in professional settings, might be due to the challenge of collecting and analyzing spoken data. Yet the imbalance between spoken and written texts means that word lists developed from these corpora are biased towards written language and may not represent the lexical items that learners are likely to encounter frequently in speech.

Miller (2022) pointed out that compared to situational representativeness, linguistic representativeness has rarely been taken into account by word list researchers. Our review supports Miller's point to some extent. Only 21.74% of the reviewed studies took linguistic representativeness into consideration. Except for Brezina and Gablasova (2015), these studies often compared items of a newly developed list with those of former lists of similar purposes. As these lists were developed based on different selection criteria, such comparison does not provide direct evidence of how well the newly developed list represents the target discourse domain.

While previous word list studies have considered corpus representativeness to some extent, further care could be taken to improve corpus representativeness to reduce differences across lists of similar purposes. Sorell's (2013) study with general words and Miller's (2022) study with specialized words provide some useful directions. Although focusing on different kinds of vocabulary, these studies employed similar methods; that is, they created lists from multiple similar-sized corpora which represented the same target discourse domains and then compared the overlaps among these lists to determine their stability.

Both Sorell (2013) and Miller (2022) found that the stability of word lists increased according to corpus size. This means that larger corpora are essential to increase the reliability of word lists. Sorell (2013) also suggested that in addition to size, careful sampling of texts is also important. Creating word lists from 50 million-word corpora which represented texts carefully sampled from four major text types, Sorell (2013) found that the variation for the most frequent 1,000 words of any of the text types was around 2% and for 3,000 words was less than 5%. Even with the most frequent 9,000 words, the variation was just 4%-7%. These findings indicate that a considerable level of reliability of word lists is possible with a large and well-designed corpus (see Nation and Sorell, 2016 for further information of how to design corpora for word list studies).

3.4. What criteria have been used to select items for word lists?

Frequency, range, and dispersion as selection criteria for lists of single words

In the last decade, frequency, range, and dispersion continue to be the most popular selection criteria for lists of single words, but they have been measured using more sophisticated indices. Frequency has been used in 61.76% of the lists of single words (21 lists). Figure 5 shows that while absolute frequency was still employed by 10 lists (e.g., Coxhead & Demecheleer, 2018), since 2015 relative frequency has been increasingly used (e.g., Drayton & Coxhead, 2023). Absolute frequency refers to the total number of occurrences of a word in a text or a corpus, and thus is a useful index if we only examine a single text or corpus. However, to examine the occurrence of a word across corpora or texts of different sizes, relative frequency is a more useful index. Relative frequency is calculated by dividing the absolute frequency of a word by the size of a text or a corpus and multiplying by the basis for normalization (e.g., 100 words or one million words). As relative frequency takes corpus sizes into consideration, it helps to minimize the bias toward corpus size in frequency counts. However, relative frequency may be misleading in small corpora because it can overestimate the occurrence of rare words in the corpus, and thus should be reported together with absolute frequency (Brezina, 2018).

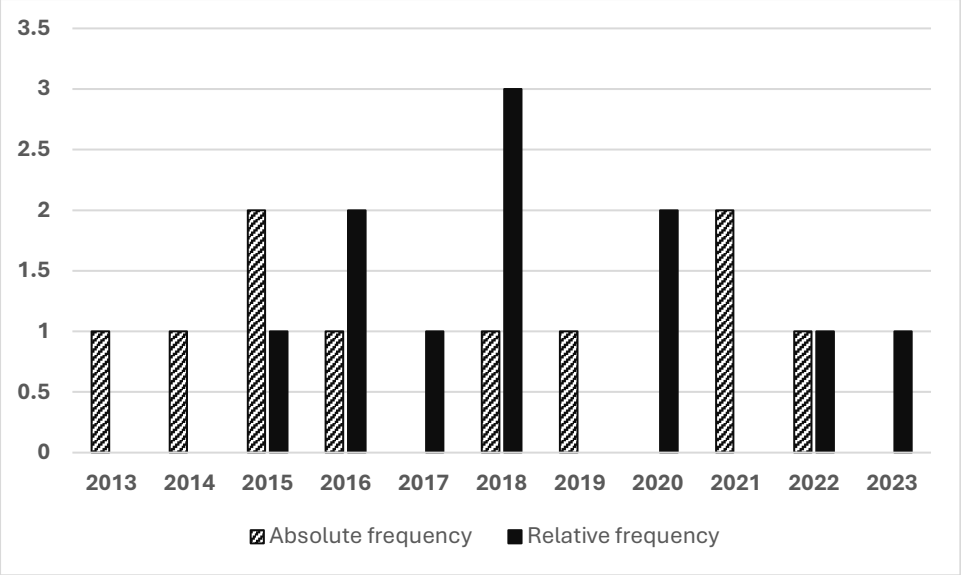

Figure 5. Indices of frequency in lists of single words

Frequency ensures the selected items represent the words that occur frequently in the corpus. However, relying solely on frequency may lead to the risk that the list will include items that are extremely frequent in a small number of texts, but are rare in most of the texts in the corpora. Therefore, range and dispersion are also important criteria, collectively employed in 52.94% of lists of single words. Range refers to the number of texts in the corpus in which a lexical item occurs while dispersion refers to the distribution of that item throughout the corpus. Using both range and dispersion ensures that selected items occur in a range of texts and the frequency of the items in each text are relatively similar. Range plus dispersion has been increasingly used as selection criteria in lists of single words (see Figure 6).

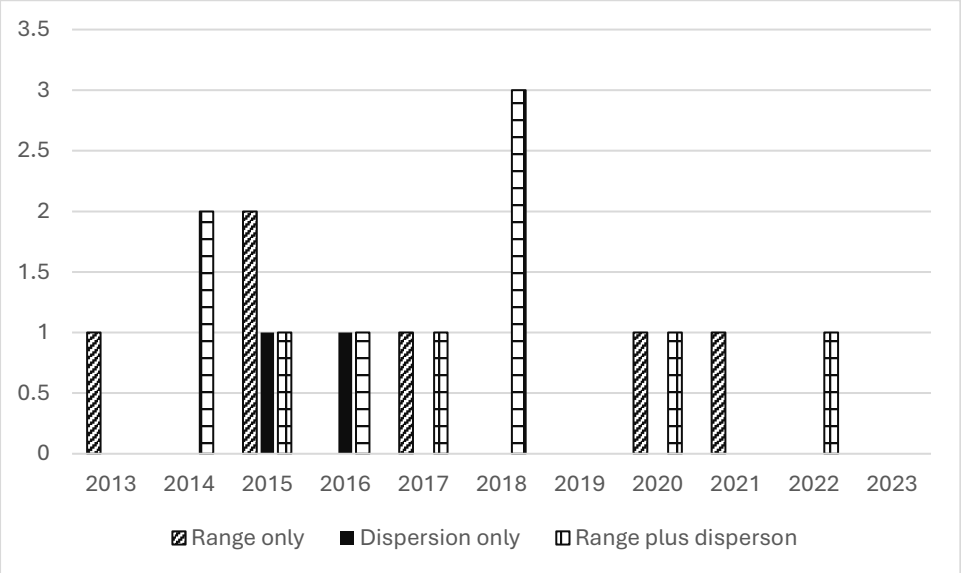

Figure 6. Number of lists of single words using range and dispersion (*N= 18*)

Specialization in lists of single words

Another change in the last decade is related to the approaches toward distinguishing specialized words from general service words in the development of lists of academic and technical words. Specialised word lists developed before 2013 often followed Coxhead's approach by assuming that L2 learners already know general service words, and thus excluded general service words, represented by West's GSL, from specialised lists. This approach was still adopted by 21.43% of the specialised word lists published post 2013 (e.g., Bancroft-Billings, 2020). Excluding general service words from specialized word lists takes learners' prior knowledge of general service vocabulary into account and minimizes the chance of teaching and learning known items. However, as the GSL was used as the general service word baseline, specialized word lists developed with this approach have been inevitably affected by the limitations of the GSL.

Several approaches have been taken to address this limitation. The first approach, employed by 10.71% of the specialised word lists, was to use updated lists of general service words. Yet the quality of specialized word lists still depends on the quality of lists used as the general service vocabulary baseline. For example, items in Browne's (2013) New Academic Word List (NAWL) are those outside Browne et al. (2013) New General Service List (NGSL). However, little validation evidence of the NGSL is provided, which makes the validity of the NAWL questionable.

A second approach, employed by 53.57% of the specialized word lists (e.g., Gardner & Davies, 2014), is to include general service words if they meet the selection criteria. This approach enables these lists to avoid the limitations of any general service vocabulary baseline. Yet it does not consider learners' prior knowledge of general vocabulary, which may result in inefficient use of learning time on known items.

Expanding on both approaches, recently a small number of word lists (14.29%) (e.g., Dang et al., 2017) included all general service words that met the selection criteria, but then classified them into levels based on their frequency in general English. This allows users to focus on learning academic and technical words from levels that are beyond their current knowledge of general vocabulary and enables specialized word lists to avoid limitations caused by general service baselines while still accommodating users' prior vocabulary knowledge.

Selection criteria for lists of multiword sequences

Studies developing lists of multiword sequences have widely employed frequency, a common selection criterion of single word lists, as their selection criterion. Frequency was used by 81.25% of lists of multiword sequences (13 lists), and most of these lists (11 lists) had relative frequency as the index (e.g., Ackerman & Chen, 2013).

However, other criteria which were commonly used in lists of single words were less common in the development of lists of multiword sequences. Range was employed in 50% of the lists (e.g., Lei & Liu, 2018) while dispersion in only 18.75% (Green & Lambert, 2019). This modest application of range and dispersion could be because these criteria may not always be relevant to multiword sequences. Combining range and dispersion enables researchers to identify items that occur frequently in a large number of texts. However, multiword sequences do not occur as frequently. If we use range and dispersion as additional criteria, corpora for the analysis should be much larger than those used to develop lists of single words. It may be challenging to develop such corpora, especially if the focus is on spoken and specialised contexts.

Apart from frequency, range, and dispersion, strength of association has been used in the selection of multiword sequences. Strength of association refers to the likelihood that two word forms re-occur together more than by chance, and thus is a useful criterion to select multiword sequences. Despite its value, strength of association was used as a criterion by only 37.5% of the studies.

Mutual information (MI) and/or t-score were used as indices of strength of association (e.g., Rogers et al., 2021) despite Gablasova et al.'s (2017) suggestions that apart from MI and t-scores, there are a range of indices of strength of association and the most appropriate index varies depending on research purposes.

Taken together, our review suggests that compared to lists of single words, lists of multiword sequences could benefit from more sophisticated selection criteria specifically developed for multiword sequences.

Supplements to corpus-driven criteria

Corpus-driven criteria (e.g., frequency, range, dispersion) offer word list research robust tools to identify the lexical items that students need to know for effective communication. They also enable researchers to identify items for their lists in an objective and quick way, which

allows other researchers to replicate their studies. However, word lists derived solely from corpora have limitations.

Due to practicability (e.g., financial constraints, copyright and ethical restrictions), researchers may not be able to access relevant texts. This can have negative impacts on the representativeness of a corpus and in turn, word lists themselves. Moreover, only using corpus-driven criteria often results in overly long lists, making it challenging to incorporate a list into curricula (Dang & Webb, 2016). Practitioners (e.g., Stein, 2017) raise the concern that word lists developed from corpora of language use at large may not represent the experience of a learner in a specific context, potentially discouraging implementation of word lists in pedagogy.

Recognizing this concern, 52.18% of the 46 word lists have used information from other sources (dictionaries, concordance lines, experts) (e.g., Dang, 2020; Lei & Liu, 2016) to further filter items from corpus-driven word lists. Despite these efforts, it is important to note that 32.61% (e.g., Brezina & Gablasova, 2015) relied solely on corpus-driven criteria. Meanwhile, 15.22% mentioned using supplements to corpora-driven criteria, but did not describe these criteria and/or procedure in detail making replication impossible (e.g., Browne, 2013). This lack of information makes it difficult for researchers and practitioners to judge the lists' quality.

3.5. How word lists have been validated?

Validation is an essential stage of word list development. However, 43.48% of the 46 word list studies (e.g., Chon & Shin, 2013) did not conduct validation while one study (Browne, 2013) mentioned having validated their list but did not report how the validation was done. This is problematic because the validity of the word lists is questionable and thus discourages use in research, language learning and teaching.

All remaining 25 studies, which included validation, used lexical coverage in corpora as a validation criterion (e.g., Brezina & Gablasova, 2015). Lexical coverage refers to the proportion of words in a corpus covered by items from a word list, and thus is a useful criterion to validate word lists. SLA studies (e.g., Schmitt et al., 2011; van Zeeland & Schmitt, 2013) have found that L2 learners' comprehension of texts increases according to the proportion of known words in the text.

Therefore, the higher the lexical coverage provided by a comprehension-oriented word list, the better the list may be. However, lexical coverage is only one factor determining the value of a pedagogical word list.

Other factors (e.g., target users, corpus quality, selection criteria) also matter. Therefore, two recent studies (Nation, 2016; Drayton & Coxhead, 2023) have used Nation's (2016) framework to evaluate their lists. This framework is presented as a series of aspects that need consideration to make a word list suitable to a specific pedagogical context (see Appendix 2). The first aspect (purpose) focuses on the learners and kinds of vocabulary knowledge that the list targets at. The next three aspects (unit of counting, main word lists, and other lists) cover the lexical units of the list and whether it is relevant to the target learners. The fifth and sixth aspects (corpus and making the lists) respectively evaluate the quality of the corpora used to develop the list and criteria used to select items for the list. The last two aspects (self-criticism and availability) consider whether list developers explicitly acknowledged the limitation of the lists and made it available for further evaluation.

Using Nation's framework makes the procedure of construction and validation of Nation's (2016) and Drayton and Coxhead's (2023) lists transparent to researchers and practitioners. However, because these researchers used their own judgement to evaluate their lists against Nation's framework, their evaluation may not always reflect what teachers, learners, and policy makers think. To make word lists relevant to learning and teaching, word list developers should involve these stakeholders in the validation. Yet none of the 46 comprehension-oriented word lists included this validation step.

He and Godfroid (2018) and Dang et al. (2022) have included teachers and learners in the validation of word lists. He and Godfroid (2018) asked English as a Second Language (ESL) teachers to rate the usefulness and difficulty of items from a word list used in an EAP course and used this information to sequence items in the list in terms of priority for learning. Dang et al. (2022) used ESL/English as a Foreign Language (EFL) teachers' judgment of word usefulness and EFL learners' knowledge to compare the BNC/COCA2000 (Nation, 2012, 2016) and the New General Service List (Brezina & Gablasova, 2015), the two general service lists which provided the highest lexical coverage among four established lists of general service words. He and Godfroid's (2018) and Dang et al.'s (2022) findings showcase the value of involving stakeholders in word list validation in making word lists better matched to pedagogical contexts. Yet the limited involvement of stakeholders in the

validation of published lists in the last decade may reduce the face validity of word lists in pedagogical contexts.

4. To what extent have published lists been applied in research on vocabulary learning, teaching, and assessment?

4.1. Applications of word lists for learning and teaching from word list developers' perspectives

Nation and Coxhead's (2021) research on native speaker vocabulary size and text coverage based on the BNC/COCA 25,000 word lists indicated that adult, educated native speakers of English tend to have a vocabulary of about 20,000 general word families and depending on their academic discipline, they may have an additional knowledge of a few hundred to several thousand technical words. Targeting this vocabulary size may be a daunting task for many EFL learners given that they may learn only 400 word families per year (Webb & Chang, 2012).

However, lexical profiling studies using corpus-based word lists such as the BNC/COCA lists and their earlier version (BNC lists) have indicated that knowledge of the most frequent 2,000-3,000 general word families might be sufficient for L2 learners to comprehend a range of discourse types and for a higher degree of comprehension, a knowledge of 8,000-9,000 word families is likely to be needed.

These findings indicate that the lexical demands of the tasks faced by L2 learners are more manageable than that suggested by the vocabulary size of native speakers. This finding is important because it demonstrates the value of corpus-based word lists in setting realistic learning goals for L2 learners. When combining this information with factors affecting learners' vocabulary growth (Keuleers et al., 2015), it can inform the decision on what vocabulary should be prioritized in L2 pedagogy. Therefore, Nation (2016) has pointed out that word lists have several applications in pedagogy, including setting learning goals, informing the selection of target words, the design of activities and materials across the Four Strands, and the design of tests to assess learners' vocabulary knowledge.

Of the 46 studies which developed comprehension-oriented word lists, 93.47% explicitly suggested possible applications of these lists in research and pedagogy while 6.52% did not. Table 2 shows that setting learning goals is the most frequently mentioned application (81.40%). It was followed by indicating target words for deliberate vocabulary learning activities and informing material design (62.79% each).

Next came reference tools for learners and teachers (37.21%) and informing vocabulary assessment (30.23%). These applications are aligned with Nation's (2016) suggestions about the applications of word lists for language learning and teaching.

Table 2. Applications mentioned by the developers of the 46 comprehension-oriented word lists

Applications	% of studies
Setting learning goals	81.40
Informing the selection of target words for deliberate vocabulary learning activities	62.79
Informing material design	62.79
Reference tools for learners/teachers	37.21
Informing vocabulary assessment	30.23

While word list researchers can see many potential applications of recently developed lists in research and pedagogy, few lists have been used in lexical profiling and vocabulary testing, the two major research lines of studies of word lists. Similarly, little evidence has been found about the applications of these lists in pedagogy.

4.2. Applications of word lists in lexical profiling research

Published word lists can be used in lexical profiling studies to indicate the number of words associated with unassisted comprehension of texts, to identify the words that are likely to be unknown to learners, and to determine the extent to which words that are useful for learners re-occur in the texts. Findings of such analyses indicate the extent to which useful words are likely to be learned from the materials and inform the selection and adaptation of the materials so that they can better facilitate learning. Results largely depend on the quality of the word lists used for analyses.

However, most lexical profiling studies published in the last decade (see Appendix 3) used either the BNC/COCA 25,000 lists (65.63%), its earlier version (BNC 14,000) (18.75%), or the AWL (9.38%). Only 12.5% have employed other lists, all of which were published before 2013 – EAP Science List (Coxhead & Hirsh, 2007), Academic Formulas List (Simpson-Vlach & Ellis, 2010), and Phrasal Expressions List (Martínez & Schmitt, 2012) – and those published from 2013 onward – Academic Collocation List (ACL) (Ackerman & Chen, 2013), Academic Vocabulary List (AVL) (Gardner & Davies, 2014), Academic Spoken Word List

(Dang et al., 2017), and Medical Spoken Word List (Dang, 2020). Importantly, most of these lists were used by no more than one study.

Another application of published lists in lexical profiling research is to examine L2 learners' free productive vocabulary knowledge in writing and speaking. Free productive vocabulary knowledge is often reported as lexical sophistication; that is, the percentage of advanced words produced by learners in spoken and written texts.

Depending on specific groups of learners, advanced vocabulary can be operationalized differently. For example, multiword sequences could be considered as advanced vocabulary for EFL learners because many EFL learners have insufficient knowledge of these lexical items (Nguyen & Webb, 2017).

Technical vocabulary could be regarded as advanced vocabulary for learners studying English for academic, specific, or professional purposes because it is different from vocabulary in general communication to some extent (Coxhead, 2018). Moreover, advanced vocabulary in spoken discourse may be different from that in written discourse (Dang et al., 2021). Despite the different ways of operationalizing advanced vocabulary, lexical profiling studies do not seem to make good use of available lists to expand on earlier studies. Studies in the last 10 years (see Appendix 4) still operationalize advanced vocabulary as items appearing in academic written word lists or those that are not general service word lists. Moreover, 62.5% of studies still used the AWL and the GSL to represent academic written words and general service words.

4.2. Applications of word lists in vocabulary testing

An essential step of constructing vocabulary tests is to identify a representative sample of words to be tested. The quality of these tests largely depends on the quality of the word lists from which the test items were sampled. Published word lists developed from large and representative samples of language in the target discourse are valuable sampling pools. If these lists are well-designed, it increases the validity of tests that sample items from the lists. If not, these tests cannot accurately measure knowledge of the targeted vocabulary.

Despite the large number of word lists published over the last decade, current vocabulary tests (see Appendix 5) only sample items from a very small number of word lists. The BNC/COCA 25,000 (Nation, 2016) is the most commonly sourced list, employed by 54.55%

of studies. It is followed by the AVL (Gardner & Davies, 2014) (adopted by 18.18% of the studies). Browne's (2013) NGSL, Ackerman and Chen's (2013) ACL, and Liu's (2011) phrasal verb list were used by 9.09% each. Although a large number of technical word lists have been published in the last 10 years, none have been used to develop vocabulary tests.

Except for the Updated Vocabulary Levels Test (Webb et al., 2017) and the Vocabulary Size Test (Coxhead et al., 2015; Nation & Coxhead, 2021), both of which were developed from the BNC/COCA lists, the use of other newly developed tests in vocabulary research is limited. Therefore, it is fair to say that the impact of word lists on vocabulary testing is limited to the application of the BNC/COCA lists.

Taken together, while some attempts have been made to use more recent lists, most research on lexical profiling and vocabulary testing is still based on a very small number of lists (the GSL, AWL, and the BNC/COCA 25,000). Little evidence is seen in the applications of other recently published lists. This suggests that despite what word list developers have suggested about the potential of recently developed lists, most have had limited practical value while one list (BNC/COCA25000) has had great value.

The predominant use of the BNC/COCA25,000 could be because it is the only list that has a large number of words at different frequency levels, and it is freely available on Paul Nation's website and open access software such as RANGE (Heatley et al., 2002), Lextutor (Cobb, n.d), or AntwordProfiler (Anthony, n.d). Meanwhile, the popularity of the AWL and GSL could be because these lists are popular among teachers and learners (see Section 5) and these lists are also available via various open access software.

5. To what extent have published lists been applied in pedagogical contexts?

Exploring the extent to which published lists have been applied in pedagogical contexts is important. If key stakeholders cannot see the value of word lists, they may be hesitant to implement them in teaching and learning. Consequently, word list research is unlikely to have an impact beyond the research community. Despite its importance, studies in this line are limited in number.

5.1. Key stakeholders' perception and practice

To the best of our knowledge, only five studies have investigated teachers' applications of word lists in real pedagogical contexts. Two are small in scale and only examined the

implementation of word lists briefly. Dang and Webb (2020) delivered a questionnaire to 16 Vietnamese EFL teachers whose teaching experience ranged from two to 24 years. Part of their questionnaire asked teachers to indicate resources for their vocabulary instruction. Word lists and their related applications were the three least popular: word lists (31.25%), research-based vocabulary tests (31.25%), and lexical profilers (12.5%).

In contrast, textbooks were most popular (87.5%). Coxhead, McLaughlin, and Reid (2019) interviewed a Fabrication tutor on how he used a research-informed technical word list in his course. The tutor reported that he used the list as a reference tool. When preparing materials, he looked through items in the lists, selected words that were relevant to the topic of each session, and developed a topic-based glossary for his students to use during the course. The list also helped him to design activities to enable students to actively learn words (e.g., writing words on board, drawing their attention to the words).

Three large-scale studies have been conducted to survey teachers' applications of word lists in pedagogical contexts: Burkett (2015), Banister (2016), and Thompson and Alzeer (2019) (see Appendix 6 for an overview of these studies). The studies used questionnaires and follow-up individual interviews to explore perceptions of published lists of teachers, learners, material and test designers. While Banister (2016) focused specifically on the AWL, the other studies examined perceptions of the AWL as well as other published lists. Participants in these studies were mainly working at tertiary institutions in Global North countries (US, UK, and Canada) and focused on teaching academic English.

All three studies found that teachers, learners, material and test designers had positive opinions about published lists. 88% of Banister's (2016) participants thought that the AWL was useful. Banister (2016) reported that the five most common reasons for positive attitudes were: '*the AWL contains relevant vocabulary'* (rated as either strongly agreed or agreed by 94.29% of the teachers), '*this type of general academic vocabulary will be useful for students*' (91.43%), '*the AWL is based on corpus research not teacher judgement* (85.71%), '*the AWL set clear vocabulary learning goals*'(65.71%), and '*the AWL is easy to incorporate into my lesson* (62.86%).

Likewise, a considerable percentage of participants in Burkett's (2015) study and Thompson and Alzeer's (2019) indicated that word lists were useful (84.21%, 60.81%, respectively). Unfortunately, Burkett's and Thompson and Alzeer's participants did not specify which word

lists they were referring to. Consequently, it is unclear whether their positive attitude toward word lists is toward a specific list or a range of lists.

While most participants had a positive attitude toward word lists, only 43% of Banister's participants reported using the AWL. Likewise, 50.53% of Burkett's (2015) participants and 79.72% of Thompson and Alzeer's (2019) participants reported using published lists. Burkett (2015) reported that among lists, the AWL was the most popular, used by 48% of participants, while Thompson and Alzeer (2019) reported that 35.85% of their participants used the AWL.

A very small number of other word lists were also reported to be used (GSL, Oxford3000, BNC/COCA, GSL, new General Service List, and NAWL) but the percentages of participants reported using these lists were small (2.11% to 15.09%). It is important to note that despite the large number of technical word lists and lists of multi-word units published in the last decade, all word lists reported to be used by these teachers were made up of general service or single academic words.

Regarding how word lists were applied in pedagogy, both Burkett (2015) and Thompson and Alzeer (2019) found that language teaching was the most popular application (53.68%, 40.54%, respectively), followed by language learning (46.32%, 32.43%), developing materials (25.26%, 21.62%), and developing tests (21.05%, 21.62%). In addition to these applications, Thompson and Alzeer's participants also reported using word lists in course design (16.22%) and for research purposes (16.22%). The applications reported by these teachers were consistent with what word list developers have proposed. However, in their surveys, Burkett (2015) and Thompson and Alzeer (2019) did not distinguish between published lists and in-house lists. Therefore, it is unclear from these findings to what extent published lists have been applied.

Banister's (2016) interviews showed that teachers had several ways of implementing the AWL into their teaching: (a) using the AWL Highlighter to focus students on words that are worth learning, (b) providing language-focused learning activities for students to use AWL words (e.g., gap filling, word building activities, pronunciation, collocation exercise), and (c) instructing students to look at the percentage of AWL words in passages to select relevant texts for reading activities.

The way that Banister's interviewees implemented the AWL in their teaching was aligned with what word list researchers have suggested. However, it does not mean that all teachers followed such a principled approach. Two-thirds of Banister's participants reported that they only introduced the AWL briefly and then let students use the list as a self-study tool. Similarly, half of Burkett's (2015) participants who reported using wordlists for self-study indicated that they used lists discreetly rather than with any supplementary materials. While giving learners the freedom to learn items from published lists in their own time helps learners to develop autonomy, to ensure effectiveness, teachers should provide careful training to students. Yet it is unclear from the survey results if the teachers trained their students and whether the students followed their instruction.

Burkett (2015), Banister (2016), and Thompson and Alzeer (2019) also revealed a number of concerns from teachers, learners, course and material developers about word lists. Answers to open-ended questions and interviews indicated that most of these concerns were related to features of word lists themselves. Concerns included the following:

- Having lists which are separate from what was taught in prescribed textbooks makes it difficult to incorporate the lists in meaning-focused activities.
- Before implementing published lists, teachers and learners, course and material developers need to know where they can access lists and how lists were developed and validated.
- Some word lists have not gone through a rigorous validation process, which may negatively impact learning.
- Teachers, learners, course and material developers would welcome supplementary materials to come with published lists.
- Most published lists lack rich information about the words (e.g., context in which the words are used and their definition).
- Most lists are too long to fit in a language course.

In addition to concerns related to the lists themselves, several participants in Banister's (2016) study reported that word lists may be misused due to a lack of knowledge about how to incorporate them into learning programs. In fact, Burkett (2015), Banister (2016), and Thompson and Alzeer (2019) also showed that other common reasons why word lists were not used is that potential users were unaware of the existence of published lists, and they lacked experience and training in using word lists. This suggests that for published lists to be

implemented in pedagogical contexts, it is important to increase awareness of lists that are freely available and provide greater information about lists so that they can be effectively implemented in teaching and learning.

Taken together, the three studies provide useful insights into stakeholders' perspectives on word lists. However, although results suggested generally positive attitudes towards word lists, these findings may be representative of teachers who were interested in word list studies.

Most participants in Burkett's (2015) and Banister's (2016) studies were working in the U.S and the U.K while Thompson and Alzeer's participants were researchers. Thus, they may be more aware of published word lists than teachers, learners, and course and material developers working in low-resourced contexts. Moreover, given that the AWL was the most frequently used list, most opinions may relate primarily to the AWL. Perspectives on other published lists are less transparent.

5.2. Impact of local prescribed word lists on learning, teaching, and assessment

While surveys with teachers show that published lists appear to be implemented in pedagogical contexts to some extent, in many EFL contexts, what is covered in textbooks and tests is often affected by word lists prescribed by ministries of education. Two studies have examined the pedagogical impact of prescribed lists.

Jin et al. (2016) examined the extent to which items from the List of Basic English cover the vocabulary in the reading texts of the Senior high school entrance exam, a compulsory exam students need to take at the end of their basic education stage. The List of Basic English was prescribed by the Ministry of Education of China to guide the design of materials in the basic education curriculum in China. Results showed that items from this list plus marginal words and proper nouns accounted for 92.82% of the total number of words in the reading texts. Jin et al. pointed out that this coverage is below 95%, the coverage cut-off points suggested by vocabulary research for reasonable comprehension of written texts. Therefore, they noted that there are still gaps between the vocabulary covered in the curriculum and those assessed in the test.

Reynolds et al. (2018) examined the effect of learning items from a reference word list developed by the College Entrance Examination Center in Taiwan on vocabulary acquisition

and retention of first and second-year university students in this context. In Taiwan, before entering university, students studied General English courses at the secondary school level.

The design of materials and tests at secondary schools in Taiwan is guided by this reference list. Analysis of the participants' scores on items in the Vocabulary Size Test showed that the words being tested in the VST and also occurring in the reference word list were more likely to be learned than those that did not occur in the list. This finding indicated that this prescribed list may have a significant impact on L2 learners' vocabulary learning and retention in Taiwan.

Together, Jin et al.'s (2016) and Reynolds et al.'s (2018) findings suggest that prescribed lists developed by local authorities have been incorporated into the curriculum more effectively and have greater impacts on learning, teaching, and assessment than word lists developed by researchers.

However, little information is known about how these lists were developed and to what extent they reflect the items needed for communication. It is important to investigate this overlap. Given the significant impact of prescribed lists on learning, if items in these lists were not carefully selected, it means that teaching and learning time will be wasted on words that are not helpful for future language use.

6. How to maximize the impact of word list research in foreign language learning and teaching?

In the last decade, a large number of word lists have been published that can potentially inform language learning, teaching, and assessment. However, Coxhead's (2000) AWL appears to be the only list to date that is recognized beyond the research community and has had a significant impact on language learning pedagogy. Therefore, we believe that instead of putting efforts into developing new lists, researchers should focus more on exploring ways to promote existing word list research findings among key stakeholders and develop greater support for the implementation of word list use inside and outside language learning classrooms.

6.1. Further exploring the perceptions of key stakeholders

As a first step, it is essential to collect more evidence on the extent to which teachers and learners are aware of available word lists and their related research findings. It would be

particularly useful to know the extent to which different word lists have been used in various contexts. Although a small number of studies have touched on these areas to some degree, they are based on self-reported data from a sample of convenient stakeholders, most of whom are teachers working at universities in the Global North and/or are interested in word lists, which may bias results.

Moreover, no studies include the opinions of policy makers (e.g., those working at ministries of education) which may have the greatest impact on what is covered in curriculum, textbooks, and tests. This in turn affects teachers and learners' selection of vocabulary for teaching and learning.

As revealed from Burkett's (2015) surveys, one reason why his participants were hesitant to implement published lists in their contexts was the inconsistency between what was presented in these lists and what was covered in the textbooks. Therefore, if policy makers can see the value of published lists for students, they may incorporate these lists to inform the design of materials and tests for students. This will then create a positive test washback to encourage teachers and learners to teach and learn items from published lists.

6.2. **Addressing the concerns of key stakeholders**

While waiting for further evidence to be collected, several actions should be taken to address issues that were reported in earlier studies (Banister, 2015; Burkett, 2015; Thompson & Alzeer, 2019) that discourage the implementation of published lists in pedagogy.

Accessibility

Teachers and learners need to not only be aware of existing lists and how these lists were developed and validated, but they also need to know where they can access these lists. Of the 46 comprehension-oriented word lists, 4.35% were not available (e.g., Liu & Han, 2015) and 63.04% are only available as article appendices (e.g., Chon & Shin, 2013), which makes access for teachers and learners challenging.

Word list developers could make their lists and relevant guidelines available in open platforms such as the OSF www.osf.io and IRIS www.iris-database.org to improve accessibility. Furthermore, researchers can make word list research findings more accessible to practitioners via Open Accessible Summaries in Language Studies (OASIS) https://oasis-database.org/. To save key stakeholders time from searching for information, it is also worth

having a uniformed online archive for teachers and learners where word lists can be uploaded together with their guidelines.

Developing ready-made materials and providing CPD training to teachers

One reason for the popularity of the AWL is the availability of a large number of ready-made materials for learning AWL items. Of the 46 comprehension-oriented word lists, 65.22% are only available as a list of word forms. To use these lists for text analysis, teachers will need to convert them into text files and use them with specialised software such as RANGE (Heatley et al., 2002) or Antword Profiler (Anthony, n.d). This complex process may discourage teachers and learners from using lists.

While certain efforts have been made to create user-friendly platforms to analyze texts (e.g., Cobb's (n.d) Lextutor and Smith's (n.d) *EAP* Foundation) and check the frequency level of a BNC/COCA word and its family member (Anthony's (2013) Word Family Finder), there would be value in having researchers work closely with publishers to co-design ready-made learning materials that help students to learn word list items (e.g., textbooks). There is also a need to offer teachers Continuing Professional Development (CPD) training opportunities that follow research-informed principles for using word lists in text analysis software to set learning goals, assess learning progress, and design learning materials and activities. Teachers can then reflect and explore how to implement these ideas in their own contexts and evaluate the success of such interventions.

Developing word lists to accommodate needs

One reason why word lists may not be used is that teachers and learners may not believe most lists are relevant to their context. Therefore, word list researchers should be mindful when thinking about developing new lists. Researchers could develop close partnerships with key stakeholders to inform all steps of designing and implementing word lists into pedagogy. Such co-production can help lists match the needs of target users and thus maximize impact.

If there are no clear needs, it is not worth putting effort into developing new lists. A more sustainable way to tackle concerns about the relevance of published lists to specific contexts is to provide key stakeholders with training so that they can develop Do-It-Yourself word lists to serve their own needs. Nation's (2016) book is among the very first attempts to do so, but more actions should be taken.

Promoting learners' autonomy in vocabulary learning

Most published lists present word forms without any contextual information. This makes it difficult to implement the use of lists into pedagogical contexts, because knowing a word involves not only learning its form. One useful way to address this concern is to give learners more autonomy in exploring the features of items in published lists. This can be done by engaging them in data-driven learning (DDL) activities to gain information about the selected words (e.g., meaning, collocations, part of speech). For example, concordancers can be effective for learning collocations, and different senses of words.

7. Directions for future research

Several strands of research require further investigation. The first is exploring the impact of published lists on pedagogy. Research investigating how word lists are implemented in pedagogical contexts is clearly warranted. Such studies should include the participation of teachers, learners, materials designers, and ministry of education officials and should employ more sophisticated research methods than questionnaires and individual interviews. Moreover, interventions that examine the effects of implementation of word lists in language programs are needed. Such studies may provide evidence on whether the claims of researchers about the value of word lists for learning and teaching hold true. Another avenue for further research is to examine the effects of CPD training activities and DDL activities on the implementation of word lists in teaching and learning. This might highlight optimal learning outcomes of word list informed pedagogy. Together, findings of studies in this strand may provide evidence of the value of published lists for L2 learning and teaching, which in turn may increase awareness of the use of word lists among the language learning community.

In addition to promoting findings, replicating word list studies for learners in under-represented contexts is another useful direction for future research. Most published lists, especially academic and technical word lists, are in English and based on analysis of language produced by privileged groups of people. Developing lists for learners from disadvantaged backgrounds and in other languages would make word list research more inclusive. Moreover, a great deal has been learned from the many studies of English word lists that could be applied to the development of lists in other languages. While there have been some initiatives (e.g., Coxhead et al, 2020; Coxhead & Tu'amoheloa, 2019a, b, c, d; Finlayson et al., 2024; Jakobsen et al., 2018), for more word lists in other languages to be

developed, it is crucial for researchers to develop large and representative corpora of other foreign languages and relevant corpus tools for researchers to analyze the vocabulary in these corpora.

Another important goal for word list developers is to critically think of more sustainable, low-cost ways to create word lists, which allow teachers and learners to create word lists to serve their own needs. While corpus linguistics offers an innovative way of creating word lists, creating corpora is time- and resource-consuming and developing corpus-based word lists requires certain levels of corpus literacy. Generative Artificial Intelligence (GenAI) from Large Language Models replicates natural language use at large, and thus could be a potential way to overcome these challenges.

Questions arising

1. What are key stakeholders' perceptions about published lists?
2. To what extent have published lists been implemented in pedagogical contexts?
3. To what extent do published lists overlap with lists prescribed by local authorities?
4. What are the effects of a principled vocabulary learning program which implements word lists on L2 learners' vocabulary knowledge and language proficiency?
5. What are the effects of CPD vocabulary training activities on teachers' cognition and practice related to implementing published lists in pedagogy?
6. What are the effects of DDL/CALL activities which incorporate published lists on L2 learners' vocabulary knowledge and language proficiency?
7. How can the findings of English word lists be applied to research on other foreign languages and under-represented contexts?
8. How well can GenAI generate word lists serve needs in a specific context?

Acknowledgements

We would like to thank Graeme Porte and the six anonymous reviewers for their constructive and valuable feedback, which has significantly enhanced the quality of our manuscript. Our thanks to Judith Hanks for sharing her experience in preparing a proposal for a state-of-the-art review.

Author bios

Thi Ngoc Yen Dang is an Associate Professor in Language Education at the University of Leeds. She has taught English and Applied Linguistics/TESOL in Vietnam, New Zealand, and the UK. Her research interests include vocabulary studies and corpus linguistics, and she supervises MA and PhD students in these areas. She has developed a range of word lists for second/foreign language learners and examined the applications of these lists in various aspects of language learning, teaching, and assessment. Her articles have appeared in various journals (e.g., *Applied Linguistics, Language Learning, TESOL Quarterly, Language Teaching Research, Studies in Second Language Acquisition*).

Stuart Webb is a Professor of Applied Linguistics at the University of Western Ontario. He currently teaches on the Masters in TESOL program and supervises students at the MA, PhD, and post-doctorate levels. His research interests include vocabulary studies, second language acquisition, and extensive reading, listening, and viewing. His articles have been published in journals such as *Applied Linguistics* and *Language Learning*. His latest books are *The Routledge Handbook of Vocabulary Studies*, and *How Vocabulary is Learned* (with Paul Nation).

References

Ackermann, K., & Chen, Y. H. (2013). Developing the Academic Collocation List (ACL) – A corpus driven and expert-judged approach. *Journal of English for Academic Purposes*, *12*, 235–247.doi: 10.1016/j.jeap.2013.08.002

Anthony, L. (n.d). AntWordProfiler. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software

Anthony, L. (2013). Word Family Finder. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software/wordfamilyfinder/

Bancroft-Billings, S. (2020). Identifying spoken technical legal vocabulary in a law school classroom. *English for Specific Purposes*, *60*, 9–25.doi: 10.1016/j.esp.2020.04.003

Banister, C. (2016). The academic word list: Exploring teacher practices, attitudes, and beliefs through a web-based survey and interviews. *The Journal of Teaching English for Specific and Academic Purposes*, *4*(2), 309–325. https://espeap.junis.ni.ac.rs/index.php/espeap/article/view/380/238

Bauer, L., & Nation, I. S. P. (1993). Word families. *International Journal of Lexicography*, *6*(4), 253–279.doi: 10.1093/ijl/6.4.253

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. John Benjamins.

Boers, F., & Webb, S. (2018). Teaching and learning collocation in adult second and foreign language learning. *Language Teaching*, *51*(1), 77–89.doi: 10.1017/S0261444817000301

Brezina, V. (2018). *Statistics in corpus linguistics: A practical guide*. Cambridge University Press.

Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the New General Service List. *Applied Linguistics*, *36*(1), 1–22. doi: 10.1093/applin/amt018

Browne, C. (2013). The New General Service List: Celebrating 60 years of vocabulary learning. *The Language Teacher*, *4*(37), 13–16. https://jalt-publications.org/files/pdf-article/37.4tlt_featureds.pdf

Browne, C., Culligan, B., & Phillips, J. (2013). *A new academic word list*. http://www.newacademicwordlist.org/

Burkett, T. (2015). An investigation into the use of frequency vocabulary lists in university intensive English programs. *International Journal of Bilingual & Multilingual Teachers of English*, *3*(2), 71–83.

Chon, Y. V., & Shin, D. (2013). A corpus-driven analysis of spoken and written academic collocations. *Multimedia-Assisted Language Learning*, *16*(3), 11–38.

Cobb, T. (n.d). *Compleat Lexical Tutor*. https://www.lextutor.ca/cgi-bin/range/texts/index.pl

Cobb, T., & Laufer, B. (2021). The Nuclear Word Family List: A list of the most frequent family members, including base and affixed words. *Language Learning*, *71*(3), 834–871.doi: 10.1111/lang.12452

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*(2), 213–238.doi: 10.2307/3587951

Coxhead, A. (2018). *Vocabulary and English for Specific Purposes Research: Quantitative and qualitative perspectives*. Routledge.

Coxhead, A., & Demecheleer, M. (2018). Investigating the technical vocabulary of Plumbing. *English for Specific Purposes*, *51*, 84–97.doi: 10.1016/j.esp.2018.03.006

Coxhead, A., & Hirsh, D. (2007). A pilot science-specific word list. *Revue Francaise de Linguistique Appliquee*, *XII*(2), 65–78.
Coxhead, A., McLaughlin, E., & Reid, A. (2019). The development and application of a specialised word list: The case of Fabrication. *Journal of Vocational Education and Training*, *71*(2), 175–200.doi: 10.1080/13636820.2018.1471094

Coxhead, A., Nation, P., & Sim, D. (2015). Measuring the vocabulary size of native speakers of English in New Zealand secondary schools. *New Zealand Journal of Educational Studies*, *50*(1), 121–135.

Coxhead, A., Parkinson, J., & Tu'amoheloa, F. (2020). Using Talanoa to develop bilingual word lists of technical vocabulary in the trades. *International Journal of Bilingual Education and Bilingualism*, *23*(5), 513–533.doi: 10.1080/13670050.2017.1374329

Coxhead, A. & Tu'amoheloa, F. (2019a). Bilingual English-Tongan Technical Word List for Plumbing. Occasional Publication No. 30. Wellington: School of Linguistics and Applied Language Studies, Victoria University of Wellington, New Zealand.
Coxhead, A. & Tu'amoheloa, F. (2019b). Bilingual English-Tongan Technical Word List for Automotive Technology. Occasional Publication No. 31. Wellington: School of Linguistics and Applied Language Studies, Victoria University of Wellington, New Zealand.
Coxhead, A. & Tu'amoheloa, F. (2019c). Bilingual English-Tongan Technical Word List for Carpentry. Occasional Publication No. 32. Wellington: School of Linguistics and Applied Language Studies, Victoria University of Wellington, New Zealand.
Coxhead, A. & Tu'amoheloa, F. (2019d). Bilingual English-Tongan Technical Word List for Fabrication. Occasional Publication No.33. Wellington: School of Linguistics and Applied Language Studies, Victoria University of Wellington, New Zealand.
Dang, T. N. Y. (2020). The potential for learning specialized vocabulary of university lectures and seminars through watching discipline-related TV programs: Insights from medical corpora. *TESOL Quarterly*, *54*(2), 436–459.doi: 10.1002/tesq.552

Dang, T. N. Y., Coxhead, A., & Webb, S. (2017). The academic spoken word list. *Language Learning*, *67*(4), 959–997.doi: 10.1111/lang.12253

Dang, T. N. Y., Coxhead, A., & Webb, S. (2021). Vocabulary in academic spoken English. *New Zealand Studies in Applied Linguistics*, *26*(2).
https://www.alanz.org.nz/journal/vocabulary-in-academic-spoken-english/

Dang, T. N. Y., & Webb, S. (2016). Making an essential word list for beginners. In *Making and Using Word Lists for Language Learning and Testing* (pp. 153–167). John Benjamins.

Dang, T. N. Y., & Webb, S. (2020). Vocabulary and good language teachers. In C. Griffiths & Z. Tajeddin (Eds.), *Lessons from good language teachers* (pp. 203–218). Cambridge University Press.

Dang, T. N. Y., Webb, S., & Coxhead, A. (2022). Evaluating lists of high-frequency words: Teachers' and learners' perspectives. *Language Teaching Research*, *26*(4), 617–641. doi: 10.1177/13621688209111

Davies, M. (2008) *The Corpus of Contemporary American English (COCA)*. Available online at https://www.english-corpora.org/coca/.

Drayton, J., & Coxhead, A. (2023). The development, evaluation and application of an aviation radiotelephony specialised technical vocabulary list. *English for Specific Purposes*, *69*, 51–66.doi: 10.1016/j.esp.2022.10.001.

Finlayson, N., Marsden, E., & Hawkes, R. (2024). Creating and evaluating corpus-informed word lists for adolescent, beginner-to low-intermediate learners of French, German, and Spanish. *Language Teaching Research*, 1–57. doi: 10.1177/13621688241288

Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning*, *67*(S1), 155–179.doi: 10.1111/lang.12225

Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, *35*(3), 305–327.doi: 10.1093/applin/amt015

Green, C., & Lambert, J. (2019). Position vectors, homologous chromosomes and gamma rays: Promoting disciplinary literacy through Secondary Phrase Lists. *English for Specific Purposes*, *53*, 1–12.doi: 10.1016/j.esp.2018.08.004

Greene, J. W., & Coxhead, A. (2015). *Academic vocabulary for middle school students: Research-based lists and strategies for key content areas*. Brookes Publishing.

He, X., & Godfroid, A. (2018). Choosing words to teach: A novel method for vocabulary selection and its practical application. *TESOL Quarterly*, *53*(2), 348–371.doi: 10.1002/tesq.483

Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). *Range: A program for the analysis of vocabulary in texts* [Computer software]. http://www.vuw.ac.nz/lals/staff/paul-nation/nation.aspx

Jakobsen, A. S., Coxhead, A. & Henriksen, B. (2018). General and academic high frequency vocabulary in Danish. *Nordand,* 13 (1), 64–89.doi: 10.18261/issn.2535-3381-2018-01-04

Jin, T., Li, Y., & Li, B. (2016). Vocabulary coverage of reading tests: Gaps between teaching and testing. *TESOL Quarterly*, *50*(4), 955–964.doi: 10.1002/tesq.324

Keuleers, E., Stevens, M., Mandera, P., & Brysbaert, M. (2015). Wordknowledge in the crowd: Measuring vocabulary size and word prevalence in a massive online experiment. *The Quarterly Journal Journal of Experimental Psychology*, *68*(8), 1665–1692.doi: 10.1080/17470218.2015.1022560

Lei, L., & Liu, D. (2016). A new medical academic word list: A corpus-based study with enhanced methodology. *Journal of English for Academic Purposes, 22,* 42–53.doi: 10.1016/j.jeap.2016.01.008

Lei, L., & Liu, D. (2018). The academic English collocation list. *International Journal of Corpus Linguistics*, *23*(2), 216–242.doi: 10.1075/ijcl.16135.lei

Liu, D. (2011). The most frequently used English phrasal verbs in American and British English: A multicorpus examination. *TESOL Quarterly*, *45*(4), 661–688.doi: 10.5054/tq.2011.247707

Liu, J., & Han, L. (2015). A corpus-based environmental academic word list building and its validity test. *English for Specific Purposes*, *39*, 1–11.doi: 10.1016/j.esp.2015.03.001

Martínez, R., & Schmitt, N. (2012). A Phrasal Expressions List. *Applied Linguistics, 33*(3), 299–320.doi: 10.1093/applin/ams010

Miller, D. (2022). Replication as a means of assessing corpus representativeness and the generalizability of specialized word lists. *Applied Corpus Linguistics*, *2*(3), 100027. doi: 10.1016/j.acorp.2022.100027

Nation, I. S. P. (2016). *Making and using word lists for language learning and testing*. John Benjamins.

Nation, I. S. P., & Coxhead, A. (2021). *Measuring native-speaker vocabulary size*. John Benjamins.

Nation, I. S. P., & Sorell, C. J. (2016). Corpus selection and design. In *Making and using word lists for language learning and testing* (pp. 96–105). John Benjamins.
Nesi, H. & Thompson, P. (2006). *British Academic Spoken English Corpus.* Literary and Linguistics Data Service. http://hdl.handle.net/20.500.14106/2525.
Reynolds, B. L. (2023). *Vocabulary Learning in the Wild*. Springer.

Nguyen, T. M. H., & Webb, S. (2017). Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research*, *31*(3), 98–320.doi: 10.1177/1362168816639619

Reynolds, B. L., Shih, Y.-C., & Wu, W.-H. (2018). Modeling Taiwanese adolescent learners' English vocabulary acquisition and retention: The washback effect of the College Entrance Examination Center's reference word list. *English for Specific Purposes, 52,* 47–59.doi: 10.1016/j.esp.2018.08.001

Rogers, J., Müller, A., Daulton, F. E., Dickinson, P., Florescu, C., Reid, G., & Stoeckel, T. (2021). The creation and application of a large-scale corpus-based academic multi-word unit list. *Journal of English for Academic Purposes*, *62*. doi: 10.1016/j.esp.2021.01.001

Sorell, C. J. (2013). *A study of issues and techniques for creating core vocabulary lists for English as an International Language* [PhD thesis]. Victoria University of Wellington.

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, *95*, 26–43.doi: 10.1111/j.1540-4781.2011.01146.x

Simpson-Vlach, R., & Ellis, R. (2010). An Academic Formulas List: New methods in phraseology research. *Applied Linguistics*, *31*(4), 487–512.doi: 10.1093/applin/amp058

Smith. S. (n.d). EAP Foundation https://www.eapfoundation.com/
Stein, G. (2017). Some thoughts on the issue of core vocabularies: A response to Vaclav Brezina and Dana Gablasova: 'Is there a core general vocabulary?' Introducing the New General Service List. *Applied Linguistics*, *38*(5), 759–763.doi: 10.1093/applin/amw027

Sorell, C. J. (2013). *A study of issues and techniques for creating core vocabulary lists for English as an International Language* [PhD thesis]. Victoria University of Wellington.

Thompson, P., & Alzeer, S. N. (2019). A survey of issues, practices and views related to corpus-based word lists for English language teaching and learning. *International Journal of Applied Linguistics & English Literature, 8*(6), 43–53.doi: 10.7575/aiac.ijalel.v.8n.6p.43

van Zeeland, H., & Schmitt, N. (2013). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics*, *34*(4), 457–479.doi: 10.1093/applin/ams074

Webb, S. (2020). *The Routledge Handbook of Vocabulary Studies* (S. Webb, Ed.; pp. 225–239). Routledge.

Webb, S. (2021). The lemma dilemma: How should words be operationalized in research and pedagogy? *Studies in Second Language Acquisition*, *43*(5), 941–949.doi: 10.1017/S0272263121000784

Webb, S. A., & Chang, A. C.-S. (2012). Second language vocabulary growth. *RELC Journal*, *43*(1), 113–126. doi: 10.1177/0033688212439367

Webb, S., Sasao, Y., & Balance, O. (2017). The Updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL-International Journal of Applied Linguistics*, *168*(1), 34–70.doi: 10.1075/itl.168.1.02web

West, M. (1953). *A General Service List of English words*. Longman, Green & Co.

ONLINE SUPPLEMENTARY MATERIALS

Appendix 1. Fifty word list studies being reviewed in the present study

2013

1. Ackermann, K., & Chen, Y. H. (2013). Developing the Academic Collocation List (ACL) – A corpus driven and expert-judged approach. *Journal of English for Academic Purposes*, *12*, 235–247.doi: 10.1016/j.jeap.2013.08.002
2. Browne, C. (2013). The New General Service List: Celebrating 60 years of vocabulary learning. *The Language Teacher*, *4*(37), 13–16. https://jalt-publications.org/files/pdf-article/37.4tlt_featureds.pdf
3. Browne, C., Culligan, B., & Phillips, J. (2013). *A new academic word list*. http://www.newacademicwordlist.org/
4. Chon, Y. V., & Shin, D. (2013). A corpus-driven analysis of spoken and written academic collocations. *Multimedia-Assisted Language Learning*, *16*(3), 11–38.
5. Khani, R., & Tazik, K. (2013). Towards the development of an academic word list for applied linguistics research articles. *System*, *44*(2), 209–232.doi: 10.1177/00336882134884

2014

1. Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, *35*(3), 305–327.doi: 10.1093/applin/amt015
2. Hsu, W. (2014a). Measuring the vocabulary load of engineering textbooks for EFL undergraduates. *English for Specific Purposes*, *33*, 54–65.doi: 10.1016/j.esp.2013.07.001
3. Hsu, W. (2014b). The most frequent opaque formulaic sequences in English-medium college textbooks. *System*, *47*, 146–161.doi: 10.1016/j.system.2014.10.001
4. Tangpijaikul, M. (2014). Preparing business vocabulary for the ESP classroom. *RELC Journal*, *45*(1), 51–65.doi: 10.1177/0033688214522

5. Wood, D., & Appel, R. (2014). Multiword constructions in first year business and engineering university textbooks and EAP textbooks. *Journal of English for Academic Purposes*, *15*, 1–13.doi: 10.1016/j.jeap.2014.03.002

2015

1. Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the New General Service List. *Applied Linguistics*, *36*(1), 1–22. doi: 10.1093/applin/amt018
2. Garnier, M., & Schmitt, N. (2015). The PHaVE List: A pedagogical list of phrasal verbs and their most frequent meaning senses. *Language Teaching Research*, *19*(6), 645–666.doi: 10.1177/13621688145597
3. Grabowski, L. (2015). Keywords and lexical bundles within English pharmaceutical discourse: A corpus-driven description. *English for Specific Purposes*, *38*, 23–33.doi: 10.1016/j.esp.2014.10.004
4. Greene, J. W., & Coxhead, A. (2015). *Academic vocabulary for middle school students: Research-based lists and strategies for key content areas*. Brookes Publishing.
5. Liu, J., & Han, L. (2015). A corpus-based environmental academic word list building and its validity test. *English for Specific Purposes*, *39*, 1–11.doi: 10.1016/j.esp.2015.03.001
6. Yang, M.-N. (2015). A nursing academic word list. *English for Specific Purposes*, *37*, 27–38.doi: 10.1016/j.esp.2014.05.003

2016

1. Dang, T. N. Y., & Webb, S. (2016). Making an essential word list for beginners. In *Making and Using Word Lists for Language Learning and Testing* (pp. 153–167). John Benjamins.
2. Lei, L., & Liu, D. (2016). A new medical academic word list: A corpus-based study with enhanced methodology. *Journal of English for Academic Purposes*, *22*, 42–53.doi: 10.1016/j.jeap.2016.01.008

3. Maher, P. (2016). The use of semi-technical vocabulary to understand the epistemology of a disciplinary field. *Journal of English for Academic Purposes*, *22*, 92–108.doi: 10.1016/j.jeap.2016.01.010

4. Nation, I. S. P. (2012). *The BNC/COCA word family lists* [Computer software]. http://www.victoria.ac.nz/lals/about/staff/paul-nation

5. Nation, I. S. P. (2016). *Making and using word lists for language learning and testing*. John Benjamins.

2017

1. Cunningham, K. J. (2017). A phraseological exploration of recent mathematics research articles through key phrase frames. *Journal of English for Academic Purposes*, *25*, 71–83.doi: 10.1016/j.jeap.2016.11.005

2. Dang, T. N. Y., Coxhead, A., & Webb, S. (2017). The academic spoken word list. *Language Learning*, *67*(4), 959–997.doi: 10.1111/lang.12253

3. Ha, Y. H. H., & Hyland, K. (2017). What is technicality? A Technicality analysis model for EAP vocabulary. *Journal of English for Academic Purposes*, *28*, 35–49.doi: 10.1016/j.jeap.2017.06.003

4. Watson-Todd, R. (2017). An opaque engineering word list: Which words should a teacher focus on? *English for Specific Purposes*, *45*, 31–39.doi: 10.1016/j.esp.2016.08.003

2018

1. Coxhead, A., & Demecheleer, M. (2018). Investigating the technical vocabulary of Plumbing. *English for Specific Purposes*, *51*, 84–97.doi: 10.1016/j.esp.2018.03.006

2. Dang, T. N. Y. (2018a). A Hard Science Spoken Word List. *ITL-International Journal of Applied Linguistics*, *169*(1), 44–71.doi: 10.1075/bct.109.itl.00006.dan

3. Dang, T. N. Y. (2018b). The nature of vocabulary in academic speech of hard and soft sciences. *English for Specific Purposes*, *51*, 69–83.doi: 10.1016/j.esp.2018.03.004

4. Green, C., & Lambert, J. (2018). Advancing disciplinary literacy through English for academic purposes: Discipline-specific wordlists, collocations and word families for

eight secondary subjects. *Journal of English for Academic Purposes*, *35*, 105–115.doi: 10.1016/j.jeap.2018.07.004

5. Lei, L., & Liu, D. (2018). The academic English collocation list. *International Journal of Corpus Linguistics*, *23*(2), 216–242.doi: 10.1075/ijcl.16135.lei

6. Lu, X., Yoon, J., & Kisselev, O. (2018). A phrase-frame list for social science research article introductions. *Journal of English for Academic Purposes, 36,* 76–85.doi: 10.1016/j.jeap.2018.09.004

7. Malmström, H., Pecorari, D., & Shaw, P. (2018). Words for what? Contrasting university students' receptive and productive academic vocabulary needs. *English for Specific Purposes*, *50*, 28–39.doi: 10.1016/j.esp.2017.11.002

8. Tongpoon-Patanasorn, A. (2018). Developing a frequent technical words list for finance: A hybrid approach. *English for Specific Purposes*, *51*, 45–54.doi: 10.1016/j.esp.2018.03.002

2019

1. Coxhead, A., McLaughlin, E., & Reid, A. (2019). The development and application of a specialised word list: The case of Fabrication. *Journal of Vocational Education and Training*, *71*(2), 175–200.doi: 10.1080/13636820.2018.1471094

2. Green, C., & Lambert, J. (2019). Position vectors, homologous chromosomes and gamma rays: Promoting disciplinary literacy through Secondary Phrase Lists. *English for Specific Purposes*, *53*, 1–12.doi: 10.1016/j.esp.2018.08.004

2020

1. Bancroft-Billings, S. (2020). Identifying spoken technical legal vocabulary in a law school classroom. *English for Specific Purposes*, *60*, 9–25.doi:10.1016/j.esp.2020.04.003

2. Brysbaert, M., Keuleers, E., & Mandera, P. (2020). Which words do English non-native speakers know? New supernational levels based on yes/no decision. *Second Language Research*. doi: 10.1177/0267658320934526

3. Bi, J. (2020). How large a vocabulary do Chinese computer science undergraduates need to read English-medium specialist textbooks? *English for Specific Purposes*, *58*, 77–89.doi: 10.1016/j.esp.2020.01.001

4. Dang, T. N. Y. (2020). The potential for learning specialized vocabulary of university lectures and seminars through watching discipline-related TV programs: Insights from medical corpora. *TESOL Quarterly*, *54*(2), 436–459.doi: 10.1002/tesq.552

5. Liu, D., & Myers, D. (2020). The most-common phrasal verbs with their key meanings for spoken and academic written English: A corpus analysis. *Language Teaching Research*, *24*(3), 403–424.doi: 10.1177/13621688187983

2021

1. Cobb, T., & Laufer, B. (2021). The Nuclear Word Family List: A list of the most frequent family members, including base and affixed words. *Language Learning*, *71*(3), 834–871.doi: 10.1111/lang.12452

2. Otto, P. (2021). Choosing specialized vocabulary to teach with data-driven learning: An example from civil engineering. *English for Specific Purposes*, *61*, 32–46.doi: 10.1016/j.esp.2020.08.003

3. Roesler, D. (2021). When a bug is not a bug: An introduction to the computer science academic vocabulary list. *Journal of English for Academic Purposes*, *54*, 101044.doi: 10.1016/j.jeap.2021.101044

4. Rogers, J., Müller, A., Daulton, F. E., Dickinson, P., Florescu, C., Reid, G., & Stoeckel, T. (2021). The creation and application of a large-scale corpus-based academic multi-word unit list. *Journal of English for Academic Purposes*, *62*. doi: 10.1016/j.esp.2021.01.001

5. Schmitt, N., Dunn, K., O'Sullivan, B., Anthony, L., & Kremmel, B. (2021). Introducing knowledge-based vocabulary lists (KVL). *TESOL Journal*. doi: 10.1002/tesj.622

2022

1. Alasmary, A. (2022). Academic lexical bundles in graduate-level math texts: A corpus-based expert-approved list. *Language Teaching Research*, *26*(1), 99–123.doi: 10.1177/13621688198773

2. Arndt, R. (2022). A specialized vocabulary list from an original corpus of digital science resources for middle school learners. *Journal of English for Academic Purposes*. doi: 10.1016/j.jeap.2022.101187.

3. Benson, S., & Coxhead, A. (2022). Technical single and multiword unit vocabulary in spoken rugby discourse. *English for Specific Purposes*, *66*, 111–130.doi: 10.1016/j.esp.2022.02.001

2023

1. Drayton, J., & Coxhead, A. (2023). The development, evaluation and application of an aviation radiotelephony specialised technical vocabulary list. *English for Specific Purposes*, *69*, 51–66.doi: 10.1016/j.esp.2022.10.001

2. Heidt, J., Pinchbeck, G. G., & Rodgers, M. P. H. (2023). The Good Gaming (GG) List: Key vocabulary in videogames. In B. L. Reynolds (Ed.), *Vocabulary Learning in the Wild* (pp. 143–161). Springer.

Appendix 2. Nation's (2016, p.131-132) framework of evaluating word lists

Focus	Questions
Purpose	Was the target population for the word list clearly describes?
	Was the purpose of the list clearly described?
Unit of couniting	Was the unit of counting suited to the purpose?
	Was the unit of counting clearly defined, including issues such as UK vs US spelling, alternative spelling, part of speech, abbreviation, and number?
	Was the unit of counting explicitly well-justified?
Main word lists	Was there an explicit description of what would be counted as words and what should not be included?
	Were homoforms dealt with?
	Were proper names dealt with, including proper name homoforms?
	Were content bearing proper names distinguished?
	Were hyphenated word dealt with?
	Were transparent compounds deals with in a way consisten with hyphenated words?
	Were acronyms dealt with, including acronym homoforms?
	Were the proper name lits and other lists revised on the basis of initial output?
Other lists	were marginal words dealt with?
	where any other supplementary lists used?
Corpus	Was the content of the corpus suited to the purpose of the list

	Was the corpus large enough to get reliable results?
	Was the corpus divided into sub-corpora so range and dispersion could be measured?
	Were the sub-corpora large enough, of equal size, and coherent?
	Was the corpus checked for errors?
Making the lists	Were the criteria for including and ordering in the list (frequency, range, dispersion or some composite measure) clearly described and justified?
	Were the criteria for making sub-lists clearly described and justified?
	Were any subjective criteria used? Were they described and justified?
	Were the list checked against competing list not just for coverage but also for overlapping and non-overlapping words?
Self-criticism	Are the weaknesses of the lists clearly acknowledged
Availability	Are the list readily available in electronic form for evaluation?

Appendix 3. Word lists used in lexical profiling studies published in 2013-2023

Studies	Specific kind	Word lists
Rolls & Rodgers (2018)	Science fiction fantasy texts	the EAP Science List (Coxhead & Hirsh, 2007)
Coxhead, Rahmat, & Yang (2020)	EFL textbooks in Indonesia and China	AWL (Coxhead, 2000), AVL (Gardner & Davies, 2014), ACL (Ackerman & Chen, 2013), AFL (Simpson-Vlach & Ellis, 2010)
Dang (2020)	Medical television programs	Medical Spoken Word List (Dang, 2020)
McQuillan (2020)	Novels	AWL (Coxhead, 2000)
Dang & Long (2023)	Online news	AWL (Coxhead, 2000), AVL (Gardner & Davies, 2014), ASWL (Dang et al., 2017), AFL (Simpson-Vlach & Ellis, 2010), PHRASE list (Martínez & Schmitt, 2012)

Webb & Macalister (2013)	Graded readers, text written for L1 children, text written for L1 older readers	BNC 14,000
Dang & Webb (2014)	Academic lectures & seminars	BNC 14,000
Kaneko (2014)	TOELF iBT reading texts	BNC 14,000
Kaneko (2015)	TOELF iBT listening texts	BNC 14,000
Webb & Paribakht (2015)	Reading & listening texts in standardized language tests	BNC 14,000
Nurmukhamedov (2017)	TED Talk	BNC 14,000
Nation (2014)	Novels	BNC/COCA 25,000
Hsu (2014)	Engineering textbooks	BNC/COCA 25,000
McQuillan (2016)	Novels	BNC/COCA 25,000
Tegge (2017)	Teacher-selected songs & chart songs	BNC/COCA 25,000
Coxhead (2017)	Teacher talk in EAL, Math, and Science courses at international secondary schools	BNC/COCA 25,000
Coxhead et al. (2017)	University labs and tutorials	BNC/COCA 25,000
Coxhead & Boutorwick (2018)	EAL, Math, and Science materials in an international school	BNC/COCA 25,000
Coxhead & Demecheleer (2018)	Learning materials (spoken & written) in vocational college	BNC/COCA 25,000
Hsu (2019)	Online news	BNC/COCA 25,000
Nguyen (2020)	Reading texts in high school EFL textbooks in Vietnam	BNC/COCA 25,000
Sun & Dang (2020)	High school EFL textbooks in China	BNC/COCA 25,000
Yang & Coxhead (2020)	High school EFL textbooks in China	BNC/COCA 25,000
Rodgers & Heidt (2020)	Video games	BNC/COCA 25,000
Dang (2022a)	Academic lectures	BNC/COCA 25,000
Dang (2022b)	Conference presentations	BNC/COCA 25,000

Arndt (2022b)	Digital science resources for middle school students	BNC/COCA 25,000
Benson & Coxhead (2022)	authentic spoken interactions in a rugby setting & ruby TV commentary	BNC/COCA 25,000
Lu & Dang (2022)	EAP learning materials in China	BNC/COCA 25,000
Bergström et al. (2022)	Secondary school EFL textbooks in Sweden	BNC/COCA 25,000
Nguyen (2023)	TED ED video	BNC/COCA 25,000
Candarli (2023)	Youtube video	BNC/COCA 25,000

References

Arndt, R. (2022). Vocabulary in digital science resources for middle school learners. *Applied Corpus Linguistics*, *2*, 100023.doi: 10.1016/j.acorp.2022.100023

Benson, S., & Coxhead, A. (2022). Technical single and multiword unit vocabulary in spoken rugby discourse. *English for Specific Purposes*, *66*, 111–130.doi: 10.1016/j.esp.2022.02.001

Bergström, D., Norberg, C., & Nordlund, M. (2022). Do textbooks support incidental vocabulary learning? – A corpus-based study of Swedish intermediate EFL materials. *Education Inquiry*. doi:10.1080/20004508.2022.2163050

Candarli, D. (2023). YouTube for Incidental Vocabulary Learning. In B. L. Reynolds (Ed.), *Vocabulary learning in the wild* (pp. 221–240). Springer.

Coxhead, A. (2017). Academic vocabulary in teacher talk: Challenges and opportunities for pedagogy. *Oslo Studies in Language*, *9*(3), 29–44.doi: 10.5617/osla.5845

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., Dang, T. N. Y., & Mukai, S. (2017). Single and multi-word unit vocabulary in university tutorials and laboratories: Evidence from corpora and textbooks. *Journal of English for Academic Purposes, 30,* 66–78.doi: 10.1016/j.jeap.2017.11.001

Coxhead, A., Rahmat, Y., & Yang, L. (2020). Academic single and multiword vocabulary in EFL textbooks: Case studies from Indonesia and China. *TESOLANZ Journal*, *28*, 75–88.

Dang, T. N. Y. (2020). The potential for learning specialized vocabulary of university lectures and seminars through watching discipline-related TV programs: Insights from medical corpora. *TESOL Quarterly*, *54*(2), 436–459.doi: 10.1002/tesq.552

Dang, T. N. Y. (2022a). A corpus-based study of vocabulary in conference presentations. *Journal of English for Academic Purposes*, *59*. doi: 10.1016/j.jeap.2022.101144

Dang, T. N. Y. (2022b). Vocabulary in academic lectures. *Journal of English for Academic Purposes*, *58*. doi:10.1016/j.jeap.2022.101123

Dang, T. N. Y., & Long, X. (2024). Online news as a resource for incidental learning of core academic words, academic formulas, and general formulas. *TESOL Quarterly*, *58*(1), 32–62. doi:10.1002/tesq.3208

Dang, T. N. Y., & Webb, S. (2014). The lexical profile of academic spoken English. *English for Specific Purposes*, *33*, 66–76.doi: 10.1016/j.esp.2013.08.001

Hsu, W. (2014). Measuring the vocabulary load of engineering textbooks for EFL undergraduates. *English for Specific Purposes, 33,* 54–65.doi: 10.1016/j.esp.2013.07.001

Hsu, W. (2019). Voice of America News as voluminous reading material for mid-frequency vocabulary learning. *RELC Journal, 50*(3), 408–421.doi: 10.1002/tesq.3208

Kaneko, M. (2014). Is the vocabulary level of the reading section of the TOEFL Internet-based test beyond the lexical level of Japanese senior high school students? *Vocabulary Learning and Instruction., 3*(1), 44–50.doi: 10.7820/vli.v03.1.kaneko

Kaneko, M. (2015). Vocabulary size required for the TOEFL iBT Listening Section. *The Language Teacher*, *39*(1), 9–14.

Lu, C., & Dang, T. N. Y. (2022). Vocabulary in EAP learning materials: What can we learn from teachers, learners, and corpora? *System*, *106*. doi:10.1016/j.system.2022.102791

McQuillan, J. (2020). Harry Potter and the prisoners of vocabulary instruction: Acquiring academic language at Hogwarts. *Reading in a Foreign Language*, *32*(2), 122–142.doi: 10125/67377

McQuillan, J. (2016). What can readers read after graded readers? *Reading in a Foreign Language*, *28*(1), 63–78.

Nation, P. (2014). How much input do you need to learn the most frequent 9,000 words? *Reading in a Foreign Language*, *26*(2), 1–16.

Nguyen, C. D. (2021). Lexical features of reading passages in English-language textbooks for Vietnamese high-school students: Do they foster both content and vocabulary gain? *RELC Journal*, *52*(3), 509–522.doi: 10.1177/0033688219895

Nguyen, C. D. (2023). TED Ed for Incidental L2 Academic Vocabulary Learning: A Corpus-Driven Study. In B. L. Reynolds (Ed.), *Vocabulary learning in the wild* (pp. 241–262). Springer.

Nurmukhamedov, U. (2017). Lexical coverage of TED Talks: Implications for vocabulary instruction. *TESOL Journal*, *8*(4), 768–790.doi: 10.1002/tesj.323

Rodgers, M. P. H., & Heidt, J. (2021). Levelling up comprehensible input and vocabulary learning: The lexical profile of video games. In V. Werner & F. Tegge (Eds.), *Pop culture in language education* (pp. 215–227). Routledge.

Rolls, H., & Rodgers, M. P. H. (2018). Science-specific technical vocabulary in science fiction-fantasy texts: A case for 'language through literature'. *English for Specific Purposes*, *48*, 44–56.doi: 10.1016/j.esp.2017.07.002

Sun, Y., & Dang, T. N. Y. (2020). Vocabulary in high-school EFL textbooks: Texts and learner knowledge. *System*, *93*. doi: 10.1016/j.system.2020.102279

Tegge, F. (2017). The lexical coverage of popular songs in English language teaching. *System*, *67*, 87–98.doi: 10.1016/j.system.2017.04.016

Webb, S., & Macalister, J. (2013). Is text written for children useful for L2 extensive reading? *TESOL Quarterly*, *47*(2), 300–322.doi: 10.1002/tesq.70

Webb, S., & Paribakht, T. S. (2015). What is the relationship between the lexical profile of test items and performance on a standardized English proficiency test? *English for Specific Purposes*, *38*, 34–43.doi: 10.1016/j.esp.2014.11.001

Yang, L., & Coxhead, A. (2020). A corpus-based study of vocabulary in the New Concept English textbook series. *RELC Journal*. doi:10.1177/0033688220964162

Appendix 4. Published word lists in recent studies on lexical profile of student writing and speaking (2013-2023)

Word lists	Purposes/Studies
AWL (Coxhead, 2000)	The use of academic vocabulary (as an index of lexical sophistication) (Mazgutova & Kormos, 2015)

AWL (Coxhead, 2000) & GSL (West, 1953)	The use of academic vocabulary and words that are not general service words (as indices of lexical sophistication) (Zheng, 2016)
AVL (Gardner & Davies, 2014)	The use of academic vocabulary (Csomay et al., 2018; Durrant, 2016; Olsson, 2021)
AWL (Coxhead, 2000), NAWL (Brown et al., n.d), GSL (West, 1953), & NGSL (Brown et al., 2013),	The use of academic vocabulary and words that are not general service words (as indices of lexical sophistication) (Higginbotham & Reid, 2019)
the AFL (Simpson-Vlach & Ellis, 2010), the ASWL (Dang et al., 2017), AWL and AVL, GSL and BNC/COCA 5000 lists	The use of academic vocabulary and words that are not general service words (Smith et al., 2020)
AWL (Coxhead, 2000) & GSL (West, 1953)	The use of academic vocabulary and words that are not general service words (as indices of lexical sophistication) (Zaytseva et al., 2021

References

Csomay, E., & Prades, A. (2018). Academic vocabulary in ESL student papers: A corpus-based study. *Journal of English for Academic Purposes*, *33*, 100–118.doi: /10.1016/j.jeap.2018.02.003

Durrant, P. (2016). To what extent is the Academic Vocabulary List relevant to university student writing? *English for Specific Purposes*, *43*, 49–61.doi: 10.1016/j.esp.2016.01.004

Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for Academic Purposes programme. *Journal of Second Language Writing*, *29*, 3–15.doi: 10.1016/j.jslw.2015.06.004

Olsson, E. (2021). A comparative study of CLIL implementation in upper secondary school in Sweden and students' development of L2 English academic vocabulary. *Language Teaching Research*, 1–26.doi: 10.1177/13621688211045

Smith, G. F., Kyle, K., & Crossley, A. S. (2020). Word lists and the role of academic vocabulary use in high stakes speaking assessments. *International Journal of Learner Corpus Research*, *6*(2), 194–220.doi: 10.1075/ijlcr.20008.smi

Zaytseva, V., Miralpeix, I., & Pérez-Vidal, C. (2021). Because words matter: Investigating vocabulary development across contexts and modalities. *Language Teaching Research*, *25*(2), 162–184.doi: 10.1177/1362168819852

Zheng, Y. (2016). The complex, dynamic development of L2 lexical use: A longitudinal study on Chinese learners of English. *System*, *56*, 40–53.doi: 10.1016/j.system.2015.11.007

Appendix 5. Word lists used in vocabulary test developed in the last decade

Word lists	Test
BNC/COCA 5000 (Nation, 2012, 2016)	Updated Vocabulary Levels Test (Webb et al., 2017)
BNC/COCA 5000 (Nation, 2012, 2016)	Listening Vocabulary Levels Test (McLean et al., 2015)
BNC/COCA 5000 (Nation, 2012, 2016)	receptive/orthographic (RecOrth) vocabulary knowledge test, the productive/orthographic (ProOrth) vocabulary knowledge test and the productive/phonological (ProPhon) vocabulary knowledge test (Cheng and Matthew, 2016)
BNC/COCA 25,000 (Nation, 2012, 2016)	Vocabulary Size Test (Coxhead et al., 2015; Nation & Coxhead, 2021)
BNC/COCA2000, 3000, 5000, and 10,000 (Nation, 2012, 2016)	CATSS (Aviad-Levitzky & Laufer, 2019)
BNC/COCA 3000 (Nation, 2012, 2016)	Receptive Collocation Test (Nguyen & Webb, 2017)
AVL (Gardner & Davies, 2014)	New Academic Vocabulary Levels Test (Pecorari et al., 2019)
AVL (Gardner & Davies, 2014)	Depth test of academic vocabulary (Read & Dang, 2022)
New GSL (Brown et al., 2014)	New General Service List Test (Stoeckel & Bennett, 2015)
ACL (Ackerman & Chen, 2013)	Academic Collocation Tests (Nguyen et al., 2023)
Phrasal verb list (Liu, 2011)	Aural phrasal verb meaning recall test (Cheng et al., 2023)

References

Aviad-Levitzky, T., & Laufer, B. (2019). The New Computer Adaptive Test of Size and Strength (CATSS): Development and validation. *Language Assessment Quarterly*, *16*(3), 345–368.doi: 10.1080/15434303.2019.1649409

Cheng, J., & Matthew, J. (2016). The relationship between three measures of L2 vocabulary knowledge and L2 listening and reading. *Language Testing*, *35*(1), 3–25.doi: 10.1177/026553221667

Cheng, J., Matthews, J., Lange, K., & McLean, S. (2023). Aural Single-Word and Aural Phrasal Verb Knowledge and Their Relationships to L2 Listening Comprehension. *TESOL Quarterly*, *57*(1), 213–241.doi: 10.1002/tesq.3137

Coxhead, A., Nation, P., & Sim, D. (2015). Measuring the vocabulary size of native speakers of English in New Zealand secondary schools. *New Zealand Journal of Educational Studies*, *50*(1), 121–135.

McLean, S., Kramer, B., & Beglar, D. (2015). The creation and validation of a listening vocabulary levels test. *Language Teaching Research, 19*(6), 741–760.doi: 10.1177/1362168814567

Nation, I. S. P., & Coxhead, A. (2021). *Measuring native-speaker vocabulary size*. John Benjamins.

Nguyen, T. M. H., Coxhead, A., & Gu, P. (2023). Argument-based validation of Academic Collocation Tests. *Language Testing*. doi:10.1177/026553222311984

Nguyen, T. M. H., & Webb, S. (2017). Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research*, *31*(3), 98–320.doi: 10.1177/1362168816639

Pecorari, D., Shaw, P., & Malmstrom, H. (2019). Developing a new academic vocabulary test. *Journal of English for Academic Purposes*, *39*, 59–71.doi: 10.1016/j.jeap.2019.02.004

Read, J., & Dang, T. N. Y. (2022). Measuring depth of academic vocabulary knowledge. *Language Teaching Research*. doi:10.1177/13621688221105913

Stoeckel, T., & Bennett, P. (2015). A test of the New General Service List. *Vocabulary Learning and Instruction*, *4*(1), 1–8.

Webb, S., Sasao, Y., & Balance, O. (2017). The Updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL-International Journal of Applied Linguistics*, *168*(1), 34–70.doi: 10.1075/itl.168.1.02web

Appendix 6. Overview of Burkett (2015), Banister (2016), and Thompson and Alzeer (2019)

Categories	Burkett (2015)	Banister (2016)	Thompson and Alzeer (2019)
Methods	Questionnaires + follow-up interviews	Questionnaires + follow-up interviews	Questionnaires + follow-up interviews
Investigated lists	AWL and a range of other lists	AWL	AWL and a range of other lists
Participants	95 teachers, administrators, curriculum/assessment developers, coordinators, and researchers in 657 university intensive English programs. Most of these participants were working in institutions in the US (60%), Canada (10.53%), UK (10.53%), and Australia (3.16%). The percentage of respondents from other countries is very small and limited to only 6 countries: United Arab Emirates (5.26%), Turkey (4.21%), Kuwait (3.16%), Bahrain (1.05%), China (1.05%), and the Czech Repulic (1.05%).	193 EAP teachers. Of the 135 teachers who indicated their place of employment, 118 indicated working in higher education institutions and 97 were UK based. Banister did not provide further information about the backgrounds of those who did not work in higher education institutions nor outside the UK.	74 language teachers, learners, researchers, and test and material developers who had experienced working and/or studying in ESL/EFL contexts. However, Thompson and Alzeer did not report the countries from which these participants taught and learned English. More than 60% of the participants worked in EAP/ESP programs, about 40% worked in General English, and nearly 3% worked in professional settings.

Suitability/usefulness/ importance of these lists	* 84.21% of the participants indicated that these lists were either very suitable or somewhat suitable. * 85.26% indicated that these lists were either very important or somewhat important.	Of the 83 teachers reporting, 81 out of 83 teachers who reported using the AWL with their students provided further evaluation of the list. 88% of them found the AWL either extremely useful or quite useful, 11% did not find the list that useful, and 1% did not find the list useful at all. ***Reasons for positive attitude*** When teachers who have used the AWL and had positive opinions were asked about the reasons for their positive attitude toward the AWL,70 responded. The five most common reasons were 'the AWL contains relevant vocabulary' (rated as either strongly agreed or agreed by 94.29%), 'this type of general academic vocabulary will be useful for students' (rated as either strongly agreed or agreed by 91.43%),	60.81% of the participants indicated that word lists in general are useful, 18.92% had negative ideas about word lists, and 20.27% did not responses. ***Problems of word lists*** When being asked about the problem of word lists, size was considered the most problematic (mentioned by 43.24% of the participants). Next is a lack of explanation (mentioned by 29.73%), format (mentioned by 8.11%) and access (mentioned by 6.76%). The open-ended questions also raised some other problems with word list, including: (a) Items presented without context, (b) issues with consistency and validity in the development of word lists, (c) lack of supplementary materials used

		and ‘the AWL is based on corpus research not teacher judgement (rated as either strongly agreed or agreed by 85.71%), the AWL set a clear vocabulary learning goals (rated as either strongly agreed or agreed by 65.71%), ‘the AWL is easy to incorporate into my lesson’ (rated as either strongly agreed or agreed by 62.86%). ***Reasons for the negative attitude*** When teachers who have used the AWL and had negative opinions were asked about the reasons for their negative attitude toward the AWL, the top reasons for their negative opinion were ‘using a word list is not communicative’ (either agreed or strongly agreed by more than 70%), ‘dullness of the list’ (either agreed or	with the lists, and (d) rote learning items from word lists demotivating learning.

		strongly agreed by 52%), 'content of the list' (either agreed or strongly agreed by 55%).	
Familiarity with published lists	76.84% of the participants indicated either very familiar or somewhat familiar with available frequency-based word lists.	72.54% of the participants confirmed recognizing the AWL.	N/A
Formally used vocabulary lists in their courses	50.53% of the participants reported using a vocabulary list, 33.68% reported not using, and 14.74% did not respond to the question.	43% reported having used the AWL with their students directly, 16.06% reported having not used, 1.04% reported unsure, and 39.90% did not answer the questions.	79.72% of the participants used word lists in their context while 20.27% did not.
Word lists used in their courses	1. AWL (indicated being used by 48.42% of the participants) 2. GSL (11.58%) 3. Oxford 3000 (6.32%) 4. an institutionally developed list (6.32%) 5. Nations's list based on the BNC[1] (2.11%) 6. The Word Frequency List of American English[2] (2.11%).	N/A	Among the 59 participants who indicated using word lists, 54.24% combined published lists and self-made lists, 32.20% used ready-made published lists, and 20.34% used self-made lists. When asked to specify the name of the word lists that they used, 53 participants responded to the question, the AWL was the most popular (used by 35.85% of the participants). It was followed by the

			AVL (15.09%), the BNC/COCA and the GSL (11.32% each), the new general service list[3], the new-GSL (have 26.09%), and Browne's new-AWL (5.66%).
How word lists were used in their courses	Teaching was the most popular applications (indicated by 53.68% of the participants). It was followed by learning (46.32%), developing materials (25.26%), and developing tests (21.05%). The majority of the participants indicated providing word lists with practice materials for self-study (27.37%), providing them as discrete list for self study (26.32%), or not explicitly providing but incorporating word lists into materials (17.94%). A small number of them (3.68%) indicated using word lists in the classroom.	Two-thirds of the survey respondents reported that they only introduced the AWL briefly and let students to use then as self-study tools. Banister did not report precisely what the remaining participants did.	Language teaching and language learning were the most popular applications, indicated by 40.54% and 32.43% of the participants, respectively. They were followed by language testing (21.62%), material development (21.62%), course design (16.22%), and research purposes (16.22%).

Teacher perception of good word lists	N/A	N/A	When asked about the features of good word lists, 51.32% indicated 'being topic-related'. Other features were 'based on corpus evidence' (40.54%), based on subjective judgement for selection and raking (29.73%), and based on statistical measures for selection and raking (28.38%), and small in size (25.68%).

[1,2] Burkett did not mention which lists are but could be Nation's BNC lists and Nation's BNC/COCA lists

[3] It is not clear from the participants's response whether the Browne et al.'s or Brezina and Gablasova lists.