Received: 26 June 2023 | Revised: 20 May 2024 | Accepted: 28 May 2024

DOI: 10.1111/ijal.12581

ORIGINAL ARTICLE

WILEY

Technical vocabulary in government spoken communications: The team of five million in bubbles, PPE and CBACs

Timothy Rossiter[1] | **Averil Coxhead**[2]

[1]Faculty of Foreign Languages, Van Hien University, Ho Chi Minh City, Vietnam

[2]School of Linguistics and Applied Language Studies, Victoria University of Wellington, Wellington, New Zealand

Correspondence
Timothy Rossiter, Van Hien University, Faculty of Foreign Languages, 613 Au Co Street, Phu Trung Ward, Tan Phu District, Ho Chi Minh City, Vietnam.
Email: timrossiter_@hotmail.com and timrossiter@vhu.edu.vn
Averil Coxhead, School of Linguistics and Applied Language Studies, Victoria University of Wellington, PO Box 600, Wellington 6140, New Zealand.
Email: Averil.coxhead@vuw.ac.nz

[Correction added on July 31, 2024, after the first online publication: The corresponding author's email ID has been updated in this version.]

Abstract

The New Zealand government delivered regular 1 p.m. televised COVID-19 briefings from March 2020. These events had a crucial communicative function and were usually headed by top government and medical officials. This study focuses on technical vocabulary in a corpus made up of these briefings, including single words (grouped into technical word families) and acronyms (e.g., *bubble* and *PPE*) as well as the most frequent two to five-word multiword units (MWUs; e.g., *case numbers*, *genomic sequencing*, and *chains of transmission*) containing at least one technical single-word family member. The corpus consists of 20 prepared speeches: 10 each in 2020 and 2021 by Prime Minister Jacinda Ardern and Director-General of Health Dr. Ashley Bloomfield (50,782 tokens). The results showed that 6.02% of the single-word families (e.g., *outbreak(-s)*, *contact(-s/-less)*) in the texts were technical, which may present a challenge for comprehension. Unsurprisingly, the Director-General of Health used more technical vocabulary than the Prime Minister. The top 20 MWUs containing technical vocabulary were identified in the corpus. Most were two-word collocations (e.g., *negative test*, *testing centre/s*, and *number of tests*). Implications for identifying and dealing with technical vocabulary in both government communications and language education are discussed.

KEYWORDS
Ashley Bloomfield, COVID-19, Jacinda Ardern, multiword units, New Zealand, technical vocabulary

1 | INTRODUCTION

In New Zealand and around the world, governments have attempted to inform the wider public on issues relating to the Coronavirus. People needed to understand a range of technical concepts related to the pandemic, including, for example, viral spread, restrictions on gatherings and vaccine mandates. In New Zealand, regular televised COVID-19 press briefings, at 1 p.m., were watched by many. Then Prime Minister Jacinda Ardern and Director-General of Health Dr. Ashley Bloomfield were two of the most common speakers in this context. The briefings provide an excellent example of direct communication between government and the general public. Beattie and Priestly (2021) found three key themes in these briefings: open communication, deliberate use of language to encourage changes in behaviour and an emphasis on care. The focus of the current paper is on vocabulary as a part of language use, including terms such as *breaking the chain*, *going early*, and *team of five million* (Beattie & Priestly, 2021, p. 4). The questions are what vocabulary was used in these briefings, by which speaker and how technical is this language use compared to other kinds of spoken language. These questions are important for both clear government communication and for language teachers and learners who used the briefings for information and input during the pandemic. Two major concerns about language teaching and learning in this research, which took place in a country where English is the dominant language are: (1) the New Zealand population is linguistically diverse and a large number of overseas visitors were in the country for the first lockdown, and (2) the pandemic created challenges for encountering and understanding new vocabulary (e.g., medical) in everyday language use and dealing with newly coined words.

The language of government communication during the pandemic is challenging, not just because of the potential difficulties inherent in communicating to a diverse public but because the subject matter includes specialised medical communication. Medical specialists such as Dr. Bloomfield are required to relay often highly specialised information. The current corpus-based study focuses on the regular COVID-19 briefings in New Zealand and the identification of technical single and multiword units (MWUs). It also looks at how vocabulary use differs by speaker and whether it changes over time. Technical vocabulary can be a barrier to general comprehension without pre-existing subject knowledge or support such as definitions for unfamiliar vocabulary.

1.1 | The vocabulary of a pandemic

The emergence of COVID-19 has caused words, phrases and acronyms to be introduced into everyday language use, or for everyday words to take on extended technical meanings (e.g., *bubble*). Some items are pandemic neologisms and have been created, for example, through clipping as in *coronavirus*, *the 'rona* and *coronials* (Bateup & Henderson, 2020) and *WFH* (working from home) (Asif et al., 2021). Others come from specialist domains such as medicine (e.g., *outbreak*). Bateup and Henderson (2020, p. 43) categorised pandemic lexis into areas such as *health*, *symptoms of illness* (e.g., *a/symptomatic*, *infectious*) and *social relationships* (e.g., *social bubble*, *self-isolating*). Beattie and Priestly (2021) found that the 1 p.m. briefings contained military metaphors, including *give our healthcare system a fighting chance* (Ardern, 2020, cited in Beattie and Priestly (2021, p. 4)). New words and unfamiliar medical vocabulary being presented through spoken English create challenges for language learners.

It is relatively easy to find online collections of the vocabulary of COVID-19, which have been compiled for teaching and learning purposes, such as the EnglishClub Coronavirus COVID-19 glossary (n.d.), which provides a definition and example of a word or expression related to the pandemic in context. Here is an example of *flatten the curve*:

> flatten the curve (verb - figurative): change the steep upward curve on a graph of new disease cases to a flatter, shallower upward curve over a longer time period through measures such as social distancing - Authorities hope that by introducing social distancing they will be able to flatten the curve and avoid hospitals being rapidly overwhelmed with new cases (EnglishClub, n.d.).

Morrison (2021) notes the addition of a large number of words to Te Reo Māori, which he included in the first edition of the *Raupō Phrasebook of Modern Māori*, such as *mate korona—coronavirus*.

The language of coronavirus has been the focus of research using mostly newspapers drawn together into a corpus (body of texts) for analysis. Davies' (2021) coronavirus corpus contains more than one billion words and is an international collection of thousands of English-language newspaper texts containing the keywords *coronavirus*, *COVID* or *COVID-19* from 20 countries. The collection began in 2020. The corpus interface allows users to track the emergence and disappearance of single words and MWUs across years, months, and even days. For example, the use of COVID-related MWUs such as *flatten the curve* and *social distancing* increased dramatically after March 2020 (Davies, 2021). Jiang and Hyland (2022) drew on Davies' (2020) corpus to create a 2020 coronavirus corpus consisting of 120,000 news texts from each month, which totalled 12.3 million words. The most common words in their corpus were *symptom*, *guideline*, and *restriction* in that year. They also identified many new or previously infrequently used collocations including *contact tracing* and *herd immunity*. Hyland and Jiang (2021) developed a corpus of COVID-19 scientific journal articles in medicine and biology, which were highly cited in 2020, to investigate the use, amount and kinds of hyperbole and "glamorisation" by the researchers/writers in attempts to persuade readers. They found significantly more hyping in the COVID corpus compared to a reference corpus of medicine and biology journal articles. These studies present part of the picture of vocabulary in COVID times based on newspapers. The current study aims to investigate the technical lexis of coronavirus in spoken government public communications and whether or how it varies by speaker over 2020 and 2021. Note that the focus is not the evaluation of learning outcomes with this vocabulary nor the comprehension of it by receivers of the specific spoken texts. These areas are outside the scope of this article.

1.2 | Characteristics of technical vocabulary

A key characteristic of technical vocabulary is that it is closely related to specialist subjects and domains (Chung & Nation, 2004). It is difficult to read, write, speak, or listen to a specialist subject without some knowledge of its specialised vocabulary. This means people inside the field will have a greater understanding of this lexis than those outside the field (Chung & Nation, 2004; Coxhead, 2018). Technical vocabulary in such cases can be easy to identify because we do not recognise them as a part of everyday language use. In fields such as medicine, technical vocabulary may be readily identifiable for several reasons: (1) the words may not be very frequent in everyday English (e.g., *incision* and *febrile*), (2) they can be made up of Graeco-Latin words (e.g., *cardiopulmonary resuscitation* or *CPR*) or (3) they tend to only occur in medical settings (e.g., *Antiphospholipid syndrome* and *ablation*). We may learn some of this technical vocabulary through interest, discussions with family or friends, medical treatment, or through watching medical dramas, for example.

A second characteristic of technical vocabulary is that it might occur in everyday language use and specialized contexts. In such cases, these words may have the same meaning in both contexts (e.g., *fatigue*, *forehead*), with an extended meaning (e.g., *host* and *intolerance*), or with a new meaning (e.g., *accommodation*, *stroke*, *pupil*). Quero and Coxhead (2018) found that 15% of the first 1000 (high frequency) word families of Nation's (2013) frequency-based British National Corpus/Corpus of Contemporary American English (BNC/COCA), 13% of the second 1000, and 17% of the third 1000 had a medical meaning. In other cases, everyday words such as *itching* can have a medical equivalent (*pruritus*). Technical vocabulary can include proper nouns (e.g., *Parkinson's*), transparent compounds (e.g., *superspreader*),

or acronyms (e.g., *PPE*). Abbreviations can be difficult to guess and require content knowledge to unpack (Coxhead et al., 2020). The pandemic language contains many acronyms, for example, *personal protective equipment* (*PPE*), so it is important to take these items into consideration in the present study.

Third, technical vocabulary can include single words (e.g., *epidemic*) and MWUs (e.g., *stress fracture*, *islets of Langerhans*). A MWU is an umbrella term for a range of word combinations from a pair to a group of words that commonly co-occur (Nation, 2013). It is likely a large proportion of languages are made up of MWUs (Nation & Webb, 2011), including collocations (*according to*), binomials (*fish and chips*), and idioms (*piece of cake*) (Siyanova-Chanturia et al., 2011). The meaning of MWUs may be deduced from their constituent parts (e.g., *take a chance*) or the sum of their parts (e.g., *by and large*) (Nation, 2013). A MWU may be made up of words that combine to make a new meaning (e.g., *superspreader*) or can pivot on a technical word (e.g., *test + positive/negative; positive/negative test*).

Technical MWUs, like technical single words, can require a level of specialised knowledge to understand (Coxhead, 2018). Without content knowledge of Carpentry, for example, people may not know a *hip rafter* from a *jack rafter*. Some MWU items are used in specific areas, such as *bio-secure bubbles* as part of isolation protocols for athletes (Hassan Jamaal & Friedman, 2021) while others may be used more generally with a technical word combined with a non-technical content-carrying word (e.g., *support bubble*, *family bubble*). In the COVID context, a bubble is a "self-isolating household unit" (Kearns et al., 2021). Like single words, some MWUs are more frequent than others. Bigrams or two-word collocations are the most common. Coxhead et al. (2020) found 1,802 MWUs in written pedagogical texts in Fabrication including *cutting edge* (47 occurrences) and *shielding gas* (44 occurrences). In Carpentry, some MWUs were very frequent, such as *plaster board* and *on site*, which occurred over 100 times in a 300,000-word corpus (Coxhead et al., 2020). In this study, we focus on single, two, three, and four-word technical MWUs. For the two-word collocations, we included only lexical collocations, that is, those made up of two content-carrying words (e.g., *physical distancing*) rather than grammatical collocations, which can be made up of a content-carrying word and a grammatical word (e.g., verb + that = *suggest that*). Grammatical collocations are more frequent than lexical collocations (Durrant, 2009; Liu, 2012).

A final key point is that technical vocabulary can make up a substantial proportion of words in written texts. Chung and Nation (2003) compared technical vocabulary in an Anatomy and Applied Linguistics corpus. They found almost one in three words occurring as a technical term in the Anatomy corpus and one in five words in the Applied Linguistics corpus (Chung & Nation, 2003). A similar proportion was found by Quero (2015): 29% of tokens in a medical textbook corpus were specialized medical words. Coxhead and Demecheleer (2018) developed a Plumbing Word List based on written and spoken pedagogical texts. The list contained technical word families that met selection criteria (e.g., *pipe(s/-ing/-ed*; *drain(s/-ing/-ed*). It covered 32.17% of the written corpus. Similar coverage (between 34% and 38%) was found in studies by Coxhead et al. (2020) in Fabrication, Automotive Technology, Plumbing, and Carpentry over written corpora. It appears that spoken technical texts may contain substantially fewer technical words than written texts. The Plumbing Word List covered only 11.14% of a spoken corpus of Plumbing tutor talk (Coxhead & Demecheleer, 2018) and similar findings arose from Fabrication (see Coxhead et al., 2019), Automotive Technology (see Coxhead et al., 2021), and Carpentry (see Coxhead et al., 2016 for the written text analysis) lists with between 9% and 12% coverage of spoken corpora of the respective trades. That said, Drayton and Coxhead (2023) developed a Technical Aviation Radiotelephony Vocabulary List, which contained 274 items. Its coverage reached nearly 89% over a corpus of standard phraseology. Such high coverage in contrast to the other studies in this area reflects the focused nature of communication between pilots and air traffic control. In this study, we focus on prepared speeches from COVID-19 briefings. These texts are written to be spoken and for a general audience. One issue with such texts is that their vocabulary profile may be more similar to written than spoken texts. Coxhead and Walls (2012) investigated the vocabulary in transcripts of TED Talks and found their vocabulary load is closer to written than spoken texts. Further, research into technical vocabulary in trades education found that roughly one in ten words in the spoken texts were technical compared to three to four words in ten in the written texts (Coxhead et al., 2020).

TABLE 1 Chung and Nation's (2004, p. 105) technical word scale, adapted.

Step	Examples
Step 1—Words (e.g., function words) with no specific connection to a field.	*do*, *the*, *so* and *it*.
Step 2—Words that are connected to a field such as anatomy, for example, related to aspects of the body including movement.	*structures* and *protects*.
Step 3—Words that are used in the specific field as well as in everyday English; they could also be technical in another field or not.	*ribs*, *cavity*, *lungs* and *breathing*.
Step 4—Words with a specific meaning in a field; unlikely to be in everyday English; use is restricted by the field.	*vertebrae*, *pectoralis*, *viscera* and *pedicle*.

1.3 | Identifying technical vocabulary

There are various ways in which we might identify technical vocabulary in written or spoken texts. Chung and Nation (2004) compared four methods of technical word identification in an anatomy corpus: using a scale, medical dictionaries, corpus comparison and textual analysis of diagrams, labels and definitions. Their four-point scale was the most reliable (Chung & Nation, 2004) (see Table 1). The scale was designed for the identification of vocabulary in anatomy, with the least technical rating at the top and the most technical at the bottom. This scale was used for the identification of technical language in the present study.

Coxhead et al. (2016) used corpus-based frequency measures as a first step in the development of their Carpentry Word List. They then drew on Chung and Nation's (2004) scale as a second step and third, consulted Carpentry tutors, corpus concordances, technical dictionaries, and websites in cases of ambiguity (Coxhead et al., 2016). Some disagreement between experts about the level of a word's technicality arose due to personal and/or professional interpretation (Coxhead et al., 2016). The present research has four main aims. The first is to identify technical single words in the COVID-19 briefings corpus as a whole, by speaker, and by year. The second is to identify the top 20 technical two, three and four-word MWUs in the corpus again as a whole, by speaker, and by year. The third aim is to provide suggestions on what to do with technical vocabulary in spoken communication for both learning and teaching and government communication purposes, based on their level of technicality. The fourth aim is related to the challenging area of technical language identification, and analysis. To effectively fulfil the first two analyses, we needed to develop a methodologically sound approach to the identification of both technical single and multiword units.

Research questions

1. What technical single tokens are used (a) in the whole corpus, (b) by speaker and (c) by year?
2. What are the most frequent 20 technical MWUs used (a) in the whole corpus, (b) by speaker and (c) by year?

2 | METHODOLOGY

This section outlines the corpus development and preparation for analysis and the steps for identifying technical single words and MWUs.

2.1 | Developing the corpus

The Government COVID-19 press briefings corpus was developed using several principles:

1. Only briefings where both Prime Minister Jacinda Ardern and Dr. Ashley Bloomfield presented were included. They presented most often at the briefings, particularly when there were major announcements to make, such as the

national lockdown starting in March 2020. Briefings where one party spoke alone or alongside other parties were excluded.
2. Media question sessions were excluded.
3. The minimum length of the briefings from both speakers was set at 700 words.
4. Ten briefings were collected each from 2020 and 2021.

Seventeen out of the 20 briefing transcripts are official records from the Beehive website (n.d.) and the three others (February 14, February 17 and March 1 2021) were sourced from the Ministry of Health—Manatū Hauora (n.d.) YouTube page. They were transcribed by the first author. The COVID-19 briefings corpus contains 50,782 tokens. It breaks down into two roughly equal parts by year: 2020 (25,956 tokens) and 2021 (24,826 tokens) and by speaker: Bloomfield (21,521 tokens) and Ardern (29,261 tokens).

2.2 | Corpus preparation

Several steps were followed to prepare the corpus. The first step involved categorising all the words in the corpus so they could be accounted for in Nation's (2017) BNC/COCA 25 1000 frequency-based word lists and supplementary lists of proper nouns; marginal words, i.e., *um*, *ah*; acronyms and transparent compounds. The Range program (Heatley et al., 2002) was used for this process. Any words in the corpus that were not in these lists were either added to existing lists (e.g., *Bloomfield* was added to the proper noun list, *frontline* to the transparent compound lists, and *Managed Isolation and Quarantine* [*MIQ*] to the acronyms list) or a new word list was created to enable these items to be counted, such as *motu* from Te Reo Māori, a local favourite—*zoohuis* (Zoom plus *huis;* meaning meetings—thank you, Rawinia Higgins) and miscellaneous new words such as *Bluetooth* and *yoyoing*. Six website addresses were removed from the corpus.

The next step dealt with hyphenated words such as *self-isolating*, *front-line* and *all-of-government*, depending on their constitution. The first principle for dealing with hyphenated lexis was to break them apart, in line with Nation (2016), so *all-of-government* became *all of government*. These items were then re-hyphenated in the MWU analysis. Transparent compounds (e.g. *front-line*) were joined and added to the transparent compound list.

Finally, marginal words (e.g., *um*, *er*) had been treated differently in the sources for the corpus. The three YouTube texts contained marginal words but the 17 official transcripts did not. The YouTube texts were edited for marginal words to make them consistent with the rest of the corpus.

2.3 | Data analysis

In this section, we look first at identifying single technical words and then at technical MWUs.

2.3.1 | Identifying technical single words

The process of identifying single technical words (and their technical word family members) involved multiple steps. First, Chung and Nation's (2004; see Table 1) four-step semantic scale was used to identify technical vocabulary. Items classified at Steps 3 or 4 are considered technical words. Any items that might fall between Step 2 and Step 3 were checked using corpus consultation. This meant their frequency in the online general English COCA corpus (Davies, 2009; more than 1 billion words) was compared to their frequency in the Corona Corpus (Davies, 2021; n.d.) of 1335 million words (as of January 16, 2022). If the token occurred more frequently in the Coronavirus corpus than the general COCA corpus, it was considered a candidate for being a technical word. Merriam-Webster's Medical Dictionary (n.d.) was also consulted for final analysis. Some technical words were difficult to place in Chung and Nation's (2004)

scale because they appeared more everyday words than technical, for example, *fever* and *sneeze*. They fall between Step 2 and Step 3 of Chung and Nation's (2004) scale. Corpus comparison showed *sneeze* had a frequency of 1367 in the COCA corpus (Davies, 2009) and 6666 in the Corona Corpus (Davies, 2021; n.d.) and *fever* had a frequency of 13,121 in the COCA corpus versus 73,535 in the Corona corpus. Both words appear in the Merriam-Webster Medical Dictionary. The case of these two technical words shows the benefit of the corpus comparison and dictionary-checking steps. Word family relationships were considered during the analysis. For example, *vaccine*, *vaccination* and *vaccinations* were grouped together because each of these words met the technicality criteria. The most frequent technical word in the family is presented as the headword and all family words follow, as in this example: *vaccine (vaccination/-s/-vaccinate/-ed)*. At each stage, the work was double-checked by the first author who then consulted with the second author, an experienced scholar in technical vocabulary.

2.3.2 | Identifying technical multiword units

It took three attempts to identify technical MWUs in the corpus. The first approach was to analyse the whole corpus using Antconc (Anthony, 2019) to produce a complete list of two to five-word Ngrams. However, this approach was untenable because it yielded more than 13,235 results in 260 pages of data. We then decided to use the already identified single technical words as the starting point because these words were closely related to the pandemic and were essential in many cases to understand the spread, treatment and response in the 1 p.m. briefings. However, even going through these one by one proved unworkable.

Finally, we narrowed the dataset to the 20 most frequent single technical word families (e.g., *test/-s/-ing-ed*) for efficiency (Benson & Coxhead, 2022). We used the Antconc Ngram tool (Anthony, 2019) and the most frequent technical words as head/node to be queried to locate any MWUs emerging from the single technical words. Note that a change in head/node has an impact on corpus results, as a helpful reviewer pointed out.

Analysis of the corpus for each of the top 20 technical words and word family counterparts involved several steps. When all word family relationships were included, a total of 191 single technical terms were identified from Range results; that is, the single technical word *test* was considered the headword of further technical words: *tests/testing/tested*. We based the analysis on types because not all word family members have a technical meaning (Coxhead, 2018). Only lexical collocations and not grammatical collocations (e.g., *the vaccine*) were taken into account to limit the amount of data. Research into MWUs is growing at pace. For an overview of formulaic sequences in language, see Siyanova-Chanturia and Pellicer-Sánchez (2019).

The frequency setting was a minimum of two occurrences for all two- to five-word Ngrams, and we looked to the left and right of the search word for any MWUs produced. The collocation window was set at two- to five-word Ngrams because longer word strings are not likely, and they usually contain smaller word strings (Byrd & Coxhead, 2010). A total of 69 two to five-word Ngrams were generated using these settings. It is interesting to note that some of the more frequent technical words, for example, *contact*, were part of the most frequent MWUs. The full corpus analysis was carried out first, involving every text in the study. This was followed by an analysis of each sub-corpora individually after delineation into the following four areas: the 2020 texts, 2021 texts, Prime Minister Ardern texts and Dr. Bloomfield texts.

3 | RESULTS AND DISCUSSION

3.1 | Research question 1: What technical single tokens are used (a) in the whole corpus, (b) by speaker and (c) by year?

The single technical words account for 6.02% of the corpus, meaning that approximately one in 17 words in the briefings is technical in the whole corpus. Many of the technical words identified in the corpus are related to medicine and the pandemic, for example, *swab (-s/-bing/-bed)* (26 occurrences) and *genome/genomic* (22 occurrences).

Excerpt 1 from Dr. Bloomfield's speech on 20/08/2021 contains nine examples of technical language in *italics*, such as *cases* and *bubbles*:

Excerpt 1. *Of the 31 community* cases, *19 are now confirmed to be* epidemiologically linked *to that Auckland* outbreak, *and the other 12 are being investigated but there is every sign that they are* linked, *at first glance. Of today's* cases, *three are in family* bubbles *with previously reported* cases, *and two are in the same family* bubble.

Excerpt 2 from Prime Minister Ardern's speech on 01/11/2021 contains six examples of technical vocabulary. Note that *vaccine levels* in this excerpt co-occur.

Excerpt 2. *While we've continued to see* cases *here, which represents a likely long tail, these have remained* linked. *We've had no unexpected* wastewater *detections, and* testing *rates are high.* Vaccine levels *are also providing an extra layer of protection.*

Bloomfield used more technical words in his speeches (7.66%) (e.g., *case/cases*—291 occurrences and *test/tests/testing/tested/retested*—287 occurrences) compared to Ardern (4.81%). He used words like *case* and *testing* more than Ardern. This difference is important considering Bloomfield's corpus is somewhat smaller (21,521 tokens) than Ardern's (29,261 tokens). That the Director-General of Health uses one technical word every 13 words should not be surprising given his role of providing science-based information (Table 2).

Both the Prime Minister and Dr. Bloomfield made use of technical acronyms, for example, *ESR* (7), *MIQ* (7), *PPE* (5) and *CBACs* (4) (Appendix A). Using abbreviations may save time for speakers but listeners may not be familiar with these lexical items. Listeners may not know, for example, that *ESR* stands for the *Institute of Environmental Science & Research*, which specialises in science research in New Zealand and carries out wastewater testing for COVID-19. Another important point is that knowledge of the everyday meaning of a word does not guarantee knowledge of a technical meaning (Coxhead et al., 2020); for example, *bubble(-s)* is a technical word that occurs 30 times in the corpus. The general definition from the Merriam-Webster online dictionary (n.d.) defines *bubble* as *a thin film of liquid inflated with air or gas*, which is quite different from the metaphorical meaning in a pandemic. Another example is *modelling*. The Merriam-Webster dictionary (n.d.) gives three meanings for *model*, the first is *a usually small copy of something* and the third is *a set of ideas and numbers that describe the past, present, or future state of*

TABLE 2 Examples of technical single words by speaker.

Technical terms	Ardern freq.	Bloomfield freq.	Occurrences
case (-s)	145	291	436
test (-s/-ing/-ed/retested)	120	287	407
contact (-s/-less)	85	103	188
COVID/COVID-19	97	80	177
isolate (-ing/-ed/-ion)	38	77	115
symptom (-s/-atic/-asymptomatic	26	68	94
positive (-ity)	13	52	65
vaccine (-s/vaccination/-s/-vaccinate(-ing/-ed/-unvaccinated)	108	51	159
link (-s/-ed/-age)/unlinked	11	51	62
level/-s	213	47	260
alert (-s)	89	44	133
cluster (-s/subclusters)	27	43	70

TABLE 3 The 20 most frequent technical word families from the COVID-19 briefings corpus by year.

Items	Frequency 2020	Frequency 2021	Total
case (-s)	193	243	436
test (-s/-ing/-ed/retested)	183	224	407
level (-s)	156	104	260
contact (-s/-less)	83	105	188
COVID/COVID-19	96	81	177
vaccine (-s/vaccination/-s/-vaccinate(-ing/-ed/-unvaccinated)	7	152	159
alert (-s)	70	63	133
isolate (-ing/-ed/-ion)	53	62	115
symptom (-s/-atic/-asymptomatic	46	48	94
outbreak (-s)	28	57	85
Virus	50	23	73
cluster (-s/subclusters)	64	6	70
positive (-ity)	36	29	65
link (-s/-ed/-age)/unlinked	26	36	62
infect (-ed/-ive/-ious/-iousness/-ion/-ions)	20	21	41
transmit (-ting/-ted/-ssion/-table)	17	22	39
spread (-ing)	24	13	37
lockdown (-s)	22	13	35
distanced (-ing)	21	10	31
sequence (-s/-ing/-ed/)	17	14	31
Total	**1212**	**1326**	**2538**

something—i.e., *a mathematical model*. The latter definition suggests the technical use of the corpus word rather than general use.

The 2021 corpus contained a higher proportion of technical single words (6.41%) compared to 2020 (5.65%) (Table 3). Technical acronyms occurred in the briefings but mostly with a frequency of under 30 (Appendix A shows the next most frequent 30 technical words in the 1 p.m. briefings by year), for example, *ICU* (12 occurrences) and *MIQ* (7 occurrences). Note that the frequency of some technical words rose from 2020 to 2021 (e.g., *vaccine* and family members) while others dropped from 2020 to 2021 (e.g., *cluster* and *lockdown* and their family members).

The changes in technical language frequency were an interesting aspect of this study, though perhaps not surprising for several reasons. By 2021, the New Zealand pandemic response was moving at speed, with the emergence of new technology like vaccines entering the discourse. It is also possible that government communications became more technical as the public gained familiarity with the specialized terms, though this is beyond the scope of this research. Readers interested in more detail about New Zealand's pandemic response can find information on the Ministry of Health *COVID-19: Protecting Aotearoa New Zealand* website (n.d.).

We must consider if this proportion of technical language represents a burden to listener comprehension. When technical vocabulary is used with general audiences, a lack of content knowledge could prevent understanding of the terms (Chung & Nation, 2004; Coxhead, 2018). The majority of research on technical language has been on written corpora. Both Chung and Nation's (2003) study on anatomy textbooks and Coxhead et al.'s (2020) studies in the trades found technical words covered more than 30% of each corpus. Coxhead and Demecheleer's (2018) research on

TABLE 4 Top 20 most frequent multiword units containing one technical word in the 1 p.m. briefings corpus.

Technical single word	Two-word MWU	Three-word MWU	Occurrences
level (-s) (*N* = 121)			
level	alert level		98
levels	alert levels		23
alert (*N* = 121) alert level/s			
contact (-s) (*N* = 73)			
contact	contact tracing		41
contacts	close contacts		32
test (-s/-ing/-ed) (*N* = 54)			
test	negative test		6
tests		Number of tests	4
testing	testing centre/s		25
tested	tested positive		19
isolate (-ing/-ed/-ion) (*N* = 46)			
isolate	self-isolate		9
isolating	self-isolating		4
isolation	managed isolation		33

Abbreviation: MWU, multiword unit.

plumbing and Coxhead et al.'s (2020) trades education research both showed the analysis of spoken corpora had more than 9–12% technical coverage. These spoken corpora returned higher coverage than the present study. This could be due to differences in purpose and scope. The COVID-19 briefings corpus had an informative focus with the intention to advise the entire New Zealand community whereas the trades corpora were based on recordings of polytechnic lecturers with a pedagogical focus (Coxhead, 2018; Coxhead et al., 2020).

3.2 | Research question 2: What are the most frequent 20 technical multiword units: (a) in the whole corpus, (b) by speaker and (c) by year?

The three most common MWUs were *alert level/s* (121), *contact tracing* (41) and *managed isolation* (33). Table 4 shows examples of technical MWUs (two, three and four-word MWUs) by frequency. Presenting results in this way means we can see if frequent single technical terms also make frequent MWUs. The shaded rows show word family relationships for each word. The most frequent technical MWUs are two-word lexical collocations (27), eight three-word MWUs and five four-word MWUs (see also Appendix B). No five-word MWUs met the frequency threshold. There are only a relatively small number of occurrences for each MWU, perhaps, because grammatical collocations were excluded. For example, the grammatical collocation *of COVID* occurred 50 times, but the most frequent lexical collocation with the word COVID, *cases of COVID*, occurred just five times. We can see in Appendix B that some technical words (e.g., *test* with 108 MWUs in total) were more productive in MWUs than others (e.g., *link* with 14 MWUs).

Table 5 below breaks down the technical MWU data by year. *Alert level/s* remains consistently high in both 2020 (63 occurrences) and 2021 (58 occurrences) whereas *vaccination centre/s* does not occur in the 2020 corpus but appears 24 times in 2021. This change signals a shift toward the vaccination program as the pandemic developed.

Table 6 below shows that the six most common MWUs in the Ardern corpus are *alert level/s* (83), *contact tracing* (24) and *testing centre/s* (22). Appendix B contains the top 20 list. The most common MWUs in the Bloomfield corpus are

TABLE 5 The most frequent multiword units (total) in 2020 and 2021.

Two-word MWUs	2020	2021	Total
alert level	51	47	98
contact tracing	25	16	41
managed isolation	20	13	33
close contacts	12	20	32
testing centre/s	7	18	25
vaccination centre/s	0	24	24
alert levels	12	11	23
tested positive	16	3	19

Abbreviation: MWUs, multiword units.

TABLE 6 Six most frequent two-, three- and four-word technical multiword units by speaker.

Technical term	Two-word MWU	Three-word MWU	Ardern	Bloomfield	Total
level (-s) (*N* = 121)					
level	alert level		60	38	98
levels	alert levels		23	0	23
contact (-s/-less) (*N* = 73)					
contact	contact tracing		24	17	41
Contacts	close contacts		6	26	32
test (-s/-ing/-ed) (*N* = 54)					
Test	negative test		1	5	6
Tests	number of tests		0	4	4
Testing	testing centre/s		3	22	25
Tested	tested positive		9	10	19
trace (-s/-ing/-ed) (*N* = 49)					
trace	contact trace		7	1	8
Tracing	contact tracing		24	17	41
isolate (-ing/-ed/-ion) (*N* = 46)					
isolate	self isolate		5	4	9
Isolating	self isolating		1	3	4
Isolation	managed isolation		8	25	33
case (-s) (*N* = 43)					
Case	case numbers		8	5	13
Cases	new cases		4	26	30

Abbreviation: MWUs, multiword units.

alert level/s (38), *close contacts* (26), and *managed isolation* (25). Some points of difference are that *chain/s of transmission* occurs eight times in Ardern's texts compared to twice in Bloomfield's and Ardern uses *genomic sequencing* only twice compared to Bloomfield's 13. This pattern supports the earlier finding that Bloomfield uses more technical medical MWUs compared to the Prime Minister. She appears to use more figurative than medical MWUs, such as *team of*

five million, the virus being *a dangerous enemy* and the vaccine being *our armour*. A larger corpus would provide more examples of language in use.

One issue with reporting MWU data was the possible duplication of results. It is noticeable that some single technical words are combined with other single technical words to create MWUs. This created a dilemma. Should *alert level* be categorised under *alert* or *level*? We decided to display all MWU results from the top 20 single technical words, which risks duplication but offers transparency. Another issue was deciding which approach to take for MWU analysis. Three methods were trialled and all have their advantages and disadvantages. The final approach focuses on the technical MWUs of just the top 20 single technical words and their technical family members. This approach risks the exclusion of technical MWUs that are not made up of at least one of the top 20 single technical words and their family members.

4 | IMPLICATIONS

While we do not wish this time to ever come again, these suggestions are both specific to the 1 p.m. briefings and applicable to future communications in times of emergencies or large-scale events. They also pertain to decisions on what to do when using such texts for teaching and learning purposes. Table 7 (below) applies the steps in the Chung and Nation (2004) scale and gives suggestions on what to do for vocabulary that occurs at each step. For example, the first step relates to everyday words such as function words, which have no specific connection to the pandemic. These words do not have a technical meaning; therefore, the suggestion is to do nothing at this step. In contrast, Step 4 identifies highly technical terms that are very closely related to the pandemic but unlikely to be known outside medical research or the medical profession. These words highlight the two key audiences for COVID-19 briefings: medical professionals and people outside the field of medicine.

This work has implications for government communications and language teaching and learning. Table 7 suggests that one way to deal with technical vocabulary might be to replace specialised terms with everyday words to help comprehension, particularly in the early briefings if possible. For example, *cluster* could be replaced by *group*; *asymptomatic* could be replaced by *showing no symptoms*. If there are no general alternatives for the specialised terms, providing short definitions may help. For example, *genomic sequencing* could be defined as *analysing and comparing samples of the virus*. It is possible to expand abbreviations by saying *protective clothes* instead of *PPE* or referring to the equipment itself, for example, *face shields*. Some technical acronyms such as *Community-Based Assessment Centres* (*CBACs*) could be referred to just as *testing centres*. It can be difficult to keep track of vocabulary in a fast-moving situation such as a pandemic. However, it is important to remember that technical vocabulary can include everyday words that might be familiar to people but not in a particular usage or context. Simple explanations or definitions from time to time might also be useful.

4.1 | Limitations and future research

The decision to use both the written and spoken COCA corpus was made based on the potential issues of vocabulary in written-to-be-spoken texts. In hindsight, a comparison between the combined corpora compared to just spoken texts may have been both expeditious and useful for illuminating the nature of written-to-be-spoken texts and lexis in this context. This would be a useful area of future research but is beyond the scope of the current study. It would also be useful to develop a second sample of briefings to validate the results of the analysis in this paper. A larger corpus would also provide more opportunities for the technical single words and MWUs to occur. In addition, it would be ideal to test the understanding of listeners of the technical single and MWUs that were identified in this study and how this understanding develops over time. Research into other kinds of government communications would also be valuable.

TABLE 7 Applying Chung And Nation's (2004, p. 105) technical word scale to identify and decide what to do with technical vocabulary.

Step	Examples from COVID-19 briefings	What to do for learning and teaching?	What to do for govt. communications
Step 1—Words (e.g., function words) with no specific connection to a field.	*the, it, a, to*	Do nothing.	Do nothing.
Step 2—Words that are connected to a field such as anatomy, for example, related to aspects of the body including movement.	*Testing centre; lockdown; vaccination centres; chain/s*	Make sure there are plenty of encounters in reading and listening about the topic; ensure connections between existing understandings and new uses are made (e.g., *chain*, as in a series of linked objects or events); ensure there are written forms/images to help to understand.	Use the terms several times if possible to increase exposure; provide non-technical meanings of key terms where possible; ensure there are written forms/images to help to understand.
Step 3—Words that are used in the specific field as well as in everyday English; they could also be technical in another field or not.	*Alert level/s; bubble/s; contact tracing; case; distancing; social distancing; PPE; swab*	Focus on the technical meaning of the terms; have learners draw on any first language knowledge of these terms; provide simple definitions; ensure that acronyms are spelled out regularly and examples are used; provide opportunities for use in discussion and writing where possible.	Focus on the technical meaning of the terms and ensure simple definitions are given often; ensure that acronyms are spelled out regularly and examples are used.
Step 4—Words with a specific meaning in a field; unlikely to be in everyday English; use is restricted by the field.	*Genomic sequenc(e) -ing; chain/s of transmission; don and doff; asymptomatic*	Ensure that any multiword units are kept as a chunk or word string; check with the learners which items they consider worth learning for future language use (e.g., *doff and don* are not likely to be useful in future language use); practise strategies for dealing with unknown words in texts.	Replace with everyday words (e.g., *genomic sequencing = analysing and comparing samples of the virus*) or provide short definitions using less technical vocabulary regularly as a reminder (e.g., *cluster = group*).

Future research could also consider how to best identify technical MWUs using different techniques and comparing the results. Examining collocates found around specific discourse markers could be a useful area of future research. This research encompassed 20 speeches from 2020 and 2021, collected during 2021. Further temporal analysis of changes in technical language over time, at a more granular level, with the resources and information available today is a great idea. Finally, reviewing a larger pandemic corpus to evaluate measures of speaking in plainer language would be an excellent area to study.

4.2 | Conclusion

This relatively small-scale study has identified technical single words and MWUs used in government briefings in New Zealand between 2020 and 2021. Findings showed that 6.02% of this corpus is technical, that the two speakers used different amounts of this lexis and that the vocabulary changed as the pandemic progressed. We can see the use of technical single and multiword units may pose a challenge to audiences and those involved in teaching and learning,

particularly in the early stages of the pandemic when technical vocabulary was perhaps less recognisable than in later times. Technical MWUs certainly provide a challenge in research and are worthy of more investigation.

AUTHOR CONTRIBUTIONS

Timothy Rossiter: Conceptualisation; methodology; formal analysis; data curation; writing—original draft; writing—review and editing. Averil Coxhead: Conceptualisation; methodology; writing—original draft; writing—review and editing; supervision.

ACKNOWLEDGEMENTS

We would like to acknowledge the valuable feedback from our two anonymous reviewers.

FUNDING INFORMATION

There is no funding for this article.

CONFLICT OF INTEREST STATEMENT

There is no conflict of interest for this article.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author Timothy Rossiter upon reasonable request.

ORCID

Timothy Rossiter https://orcid.org/0009-0003-8283-7743

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ijal.12581

REFERENCES

Anthony, L. (2019). *AntConc (Version 3.5.8.0)* [Computer software]. Waseda University. https://www.laurenceanthony.net/software

Asif, M., Zhiyong, D., Iram, A., & Nisar, M. (2021). Linguistic analysis of neologism related to coronavirus (COVID-19). *Social Sciences & Humanities Open*, *4*(1), 00201. https://doi.org/10.1016/j.ssaho.2021.100201

Bateup, K., & Henderson, C. (2020). The language of coronavirus. *Modern English Teacher*, *29*(4), 43–45.

Beattie, A., & Priestley, R. (2021). Fighting COVID-19 with the team of 5 million: Aotearoa New Zealand government communication during the 2020 lockdown. *Social Sciences & Humanities Open*, *4*(1), 2590–2911.

Beehive.govt.nz—The official website of the NZ Government. (n.d.). *COVID-19 updates*. https://www.beehive.govt.nz/feature/COVID-19-updates

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

Byrd, P., & Coxhead, A. (2010). On the other hand: Lexical bundles in academic writing and in the teaching of EAP. *University of Sydney Papers in TESOL*, *5*, 31–64.

Chung, T. M., & Nation, P. (2003). Technical vocabulary in specialised texts. *Reading in a Foreign Language*, *15*(2), 103–116.

Chung, T. M., & Nation, P. (2004). Identifying technical vocabulary. *System*, *32*(2), 251–263.

Coxhead, A., & Walls, R. (2012). TED talks, vocabulary, and listening for EAP. *TESOLANZ Journal*, *20*(1), 55–67.

Coxhead, A., Demecheleer, M., & McLaughlin, E. (2016). The technical vocabulary of Carpentry: Loads, lists and bearings. *TESOLANZ Journal*, *24*, 38–71.

Coxhead, A. (2018). Working with multiword units in ESP/EAP. In *Teaching essential units of language* (pp. 36–54). Routledge.

Coxhead, A., & Demecheleer, M. (2018). Investigating the technical vocabulary of Plumbing. *English for Specific Purposes*, *51*, 84–97.

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

Coxhead, A., Parkinson, J., Mackay, J., & McLaughlin, E. (2020). *English for vocational purposes: Language use in trades education*. Routledge.

Coxhead, A. (2021). Technical vocabulary in Automotive Technology. *Journal of Applied Linguistics and Professional Practice*, *14*(3), 329–350.

Davies, M. (n.d.). *The Coronavirus corpus*. https://www.english-corpora.org/corona/

Davies, M. (2009). *The corpus of contemporary American English*. https://www.english-corpora.org/coca/

Davies, M. (2021). The Coronavirus Corpus: Design, construction, and use. *International Journal of Corpus Linguistics*, *26*(4), 583–598.

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.

Durrant, P. (2009). Investigating the viability of a collocation list for students of English for academic purposes. *English for Specific Purposes*, *28*(3), 157–169.

EnglishClub.com. (n.d.). *EnglishClub Coronavirus COVID-19 vocabulary*. https://www.englishclub.com/vocabulary/coronavirus-COVID19.php

Hassan Jamaal, S. M., & Friedman, D. J. (2021). Batting against mental health in elite cricket. *British Journal of Sports Medicine*, *55*(14), 1–2. https://doi.org/10.1136/bjsports-2020-103319

Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). Range [Software]. https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/vocabulary-analysis-programs

Hyland, K., & Jiang, J. (2021). The COVID infodemic: Competition and the hyping of virus research. *International Journal of Corpus Linguistics*, *26*(4), 444–468. https://doi.org/10.1075/ijcl.20160.hyl

Jiang, F. K., & Hyland, K. (2022). COVID-19 in the news: The first 12 months. *International Journal of Applied Linguistics*, *32*(2), 241–258. https://doi.org/10.1111/ijal.12412

Kearns, N., Shortt, N., Kearns, C., Eathorne, A., Holliday, M., Mackle, D., … Braithwaite, I. (2021). How big is your bubble? Characteristics of self-isolating household units ('bubbles') during the COVID-19 Alert Level 4 period in New Zealand: A cross-sectional survey. *BMJ Open*, *11*, e042464. https://doi.org/10.1136/bmjopen-2020-042464

Liu, D. (2012). The most frequently-used multi-word constructions in academic written English: A multi-corpus study. *English for Specific Purposes*, *31*, 25–35.

Merriam-Webster Medical Dictionary. (n.d.) https://www.merriamwebster.com/medical

Ministry of Health—Manatū Hauora. (n.d.) YouTube page. https://www.youtube.com/c/minhealthnz/videos

Ministry of Health—Manatū Hauora. (n.d.) *COVID-19: Protecting Aotearoa New Zealand*. https://www.health.govt.nz/our-work/diseases-and-conditions/covid-19-novel-coronavirus/covid-19-response-planning/covid-19-protecting-aotearoa-new-zealand

Morrison, S. (2021). *The Raupō phrase book of modern Māori*. Raupō.

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.

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

Nation, I. S. P. (2017). The BNC/COCA Level 6 word family lists (Version 1.0.0) [Data file]. Available from http://www.victoria.ac.nz/lals/staff/paul-nation.aspx

Nation, P., & Webb, S. A. (2011). *Researching and analyzing vocabulary*. Heinle Cengage Learning.

Quero, B. (2015). Estimating the vocabulary size of L1 Spanish ESP learners and the vocabulary load of medical textbook *[Unpublished doctoral dissertation]*. University of Wellington.

Quero, B., & Coxhead, A. (2018). Using a corpus-based approach to select medical vocabulary for an ESP course: The case for high-frequency vocabulary. In Y. Kirkgöz & K. Dikilitaş (Eds.), *Key issues in English for specific purposes in higher education* (pp. 51–75). Springer.

Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research*, *27*(2), 251–272.

Siyanova-Chanturia, A., & Pellicer-Sánchez, A. (Eds.) (2019). *Understanding formulaic language: A second language acquisition perspective*. Routledge.

How to cite this article: Rossiter, T., & Coxhead, A. (2024). Technical vocabulary in government spoken communications: The team of five million in bubbles, PPE and CBACs. *International Journal of Applied Linguistics*, 1–19. https://doi.org/10.1111/ijal.12581

APPENDIX

TABLE A Technical word families from a total frequency of 30 to 7 occurrences in the COVID-19 briefings corpus by year.

Items	Frequency 2020	Frequency 2021	Total
bubble (-s)	13	17	30
quarantine (-ed/-ing)	20	7	27
clinic (-s/-al/-ian)	12	14	26
swab (-s/-bing/-bed)	5	21	26
genome/genomic	10	12	22
GP (-s)	8	12	20
trace (-s/-ed)	14	4	18
source (-s)	13	5	18
DHB (-s)	11	7	18
recover (-ery/-ed/-ing)	12	5	17
Exposure	5	11	16
surge/resurgence (-s)	12	3	15
chain (-s)	9	5	14
Hygiene	12	2	14
detect (-ed/-ion/-ions)/undetected	2	11	13
lab (-s-/-oratory/-oratories)	8	4	12
Wastewater	0	12	12
ICU	7	5	12
hospitalisation (-s/ised)	1	10	11
Delta	0	11	11
influenza/-flu	5	5	10
tracer (-s/-COVID-)	3	6	9
Modelling	4	4	8
eliminate (-ion/-ing/-ed)	6	2	8
epidemic/epidemiological/-ly)	2	6	8
Pandemic	6	2	8
breach (-es)	2	5	7
ESR	4	3	7
MIQ	4	3	7

Abbreviations: ESR, *Institute of Environmental Science & Research*; MIQ, *Managed Isolation and Quarantine.*

TABLE B The top 20 technical words and their two, three and four-word technical multiword units (MWUs) by frequency/year and speaker/year.

Technical term	Two-word MWU	Three-word MWU	Four-word MWU	2020	2021	Ard.	Blfd.	Tot.
case (-s)								
case	case numbers			3	10	8	5	26
cases	new cases			8	22	4	26	60
test (-s/-ing/-ed/retested)								
test	negative test			2	4	1	5	12
tests		number of tests		4	0	0	4	8
testing	testing centre/s			7	18	3	22	50
tested	tested positive			16	3	9	10	38
level (-s)								
level	alert level			51	47	60	38	196
levels	alert levels			12	11	23	0	46
contact (-s/-less)								
contact	contact tracing			25	16	24	17	82
contacts	close contacts			12	20	6	26	64
COVID/COVID-19								
COVID-19		cases of COVID-19		12	3	2	13	30
COVID		Re-emergence of COVID		2	1	3	0	6
vaccine (-s/vaccination/-s/-vaccinate(-ing/-ed/-unvaccinated)								
vaccine	COVID vaccine			0	2	1	1	4
vaccination	vaccination centre/s			0	24	22	2	48
vaccinations		walk in vaccinations		0	2	1	1	4
vaccinated	fully vaccinated			0	6	1	5	12
alert (-s)								
alert	alert level/s			63	58	83	38	242
isolate (-ing/-ed/-ion)								
isolate	self isolate			4	5	5	4	18

(Continues)

TABLE B (Continued)

Technical term	Two-word MWU	Three-word MWU	Four-word MWU	2020	2021	Ard.	Blfd.	Tot.
isolating	self isolating			2	2	1	3	8
isolation	managed isolation			20	13	8	25	66
symptom (-s/-atic/-asymptomatic)								
symptoms		people with symptoms		3	1	1	3	8
symptomatic	symptomatic people			2	1	0	3	6
virus			spread of the virus	3	2	4	1	10
cluster (-s/subclusters)								
cluster	college cluster			2	0	0	2	4
clusters	significant clusters			4	0	0	4	8
trace (-s/-ing/-ed)								
trace	contact trace			5	3	7	1	16
tracing	contact tracing			25	16	24	17	82
positive (-ity)								
positive	tested positive			16	3	2	17	38
positivity	positivity rate			0	2	0	2	4
link (-s/-ed/-age)/unlinked								
linked		cases are linked		1	4	0	5	10
			unlinked at this point	0	2	0	2	4
distance (-ed/-ing)								
distance	physical distance			1	1	1	1	4
distancing	physical distancing			10	7	14	3	34
infect (-ed/-ive/-ious/-iousness/-ion/-ions)								
infectious period				0	8	1	7	16
infection	COVID-19 infection			3	1	0	4	8

(Continues)

TABLE B (Continued)

Technical term	Two-word MWU	Three-word MWU	Four-word MWU	2020	2021	Ard.	Blfd.	Tot.
transmit (-ting/-ted/-ssion/-table)								
transmission		chain/s of transmission		5	5	8	2	20
			transmissible strains of COVID-19	0	2	2	0	4
spread (-ing)								
spread			spread of the virus	3	2	4	1	10
lockdown (-s)								
lockdown			lockdown under alert level	2	0	0	2	4
sequence (-s/-ing/-ed/)								
sequence	genome sequence			3	0	0	3	6
sequencing	genome sequenc-ing			5	10	2	13	30

AUTHOR BIOGRAPHIES

Timothy Rossiter completed an MA in Applied Linguistics (with Distinction) at Te Herenga Waka—Victoria University of Wellington, while working in the public service. He is now working as an English lecturer at Van Hien University, Ho Chi Minh City, Vietnam. He has a particular interest in plain language communication, etymology, and idioms.

Averil Coxhead teaches undergraduate and postgraduate courses in Applied Linguistics at the School of Linguistics and Applied Language Studies, Victoria University of Wellington, Aotearoa/New Zealand. She is a co-author of *English for Vocational Purposes* (2020; Routledge) and *Measuring Native Speaker Vocabulary Size* with Paul Nation (2021, John Benjamins). Averil is currently researching specialised vocabulary in the trades and higher education, as well as aspects of teaching and learning vocabulary in EFL/ESL contexts.