Studies in Second Language Learning and Teaching
Department of English Studies, Faculty of Pedagogy and Fine Arts, Adam Mickiewicz University, Kalisz
SSLLT 13 (1). 2023. 179-217
https://doi.org/10.14746/sslt.31677
http://pressto.amu.edu.pl/index.php/sslt

Specialized vocabulary across languages: The case of traditional Chinese medicine

Cailing Lu ✉
Shanghai Jiao Tong University, Shanghai, China
https://orcid.org/0000-0002-9911-2824
luca@sjtu.edu.cn

Averil Coxhead
Victoria University of Wellington, Wellington, New Zealand
https://orcid.org/0000-0003-3392-6961
averil.coxhead@vuw.ac.nz

Abstract

This paper reports on the creation of specialized word lists in traditional Chinese medicine (TCM), which is a discipline using vocabulary across languages (i.e., Chinese and English) and involves learners with different L1 backgrounds. First, a TCM Word List of 2,778 specialized words was established from corpora of TCM textbooks and journal articles. Selection criteria included specialized meaning, keyness in a corpus of general written English compared to the TCM Corpora, and frequency. The resulting TCM list covered 36.65% of the TCM Corpora but had low coverage over corpora of general written English and medical English. The TCM Word List was then divided into three sub-lists based on frequency, and graded into three levels. Level 1 contains high-frequency lexical items in English (e.g., *organ*, *coating*); Level 2 contains items that are mid-, low-frequency, or beyond any frequency levels (e.g., *pericarpium*, *metabolism*); and Level 3 contains Chinese loan words (e.g., *qi*, *yang*). Last, there is an overlap of 309 word families between this list and an earlier TCM list by Hsu (2018), which excludes words from the 1st-3rd 1,000 word families in English. Suggestions for teachers and future research are provided.

Keywords: corpus analysis; keyword analysis; specialized vocabulary; word lists; Traditional Chinese Medicine

1. Introduction

Demand for skills in health professions accelerates at an unforeseen rate. More knowledge in medical areas such as traditional Chinese medicine (TCM) could probably provide us with more solutions for global wellness. TCM has an extensive history of over five millennia. In order to promote its international development, China introduced the concept of *internationalization* of TCM in 1996 (Lin et al., 2018). As a result, the TCM international trade and market share has been promoted and its legal status has been established internationally. With such development, a growing body of research in Traditional Chinese Medicine is being published in English, and a number of institutions in English-speaking countries (e.g., Aotearoa New Zealand, Australia) offer TCM as a degree course at the tertiary level. At the same time, the demand for TCM in Chinese universities has increased from both first language Chinese speakers and learners from countries such as the U.K., Vietnam, and Malaysia. The medium of instruction is usually English for non-Chinese speakers and Mandarin for native Chinese speakers. All learners need to develop their knowledge of TCM vocabulary in English as they progress through their studies. Some TCM university programs in China combine English and TCM majors, whereby students spend their first two years in English major before undertaking their TCM study. This means English-medium communication and instruction is of essential importance in this discipline. Similar to western medical education, students start with learning theory and then move on to clinical training (Wette & Hawken, 2016). Once working, TCM practitioners in countries such as New Zealand also have to communicate with patients in English (Patel & Toossi, 2016).

Vocabulary is often perceived as an area of need by English for Academic Purposes (EAP) learners and teachers (Evans & Morrison, 2011; Flowerdew & Miller, 1992). Insufficient specialized vocabulary knowledge is cited as being particularly important but also difficult (Ardasheva & Tretter, 2015; Evans & Morrison, 2011; Flowerdew & Miller, 1992). One reason why this vocabulary poses problems for Traditional Chinese Medicine students is that, like in the case of students in other scientific disciplines, it is important for them to understand complex relationships among various scientific phenomena. At the same time, they need to develop an ability to express these relationships in spoken and written English (Ardasheva & Tretter, 2015). Secondly, TCM texts have a high vocabulary load (Lu & Coxhead, 2020), which means learners need a large vocabulary to deal with reading in this subject area. This vocabulary can be high-, mid- or low-frequency (Lu & Coxhead, 2020; see also Nation, 2016a). Last but not least, TCM is a discipline which uses vocabulary across languages (i.e., Chinese and English). This means vocabulary in this field is quite complex. Hsu (2018)

found that over 10% of the vocabulary in a TCM textbook corpus were borrowed from Chinese (e.g., *qi*). She referred to these words as *transliterated* and noted that they illustrate the lexical and cultural origin of TCM from China (Hsu, 2018). TCM texts also contain words with Graeco-Latin origins (e.g., *Astragali membranacei*) and high-frequency words with specialized meanings, such as *heart* and *gall bladder*. For example, *heart* controls *mind* in Chinese culture (Hu & Fong, 2010; Maciocia, 2005), whereas *head* and *mind* overlap more in other cultures. Learners from different cultural and linguistic backgrounds may find such conceptual differences challenging over the course of their studies.

The present study is part of a larger investigation of technical vocabulary in TCM (Lu, 2018). Part of that study included investigating knowledge of technical words in TCM possessed by students from different cultural backgrounds. Lu et al. (2021) found that English L1 speakers encountered difficulty in understanding loan words, while Chinese L1 speakers found specialized vocabulary from low-frequency levels challenging. Their findings suggested that variation in learners' linguistic backgrounds should be taken into account in word list development to better target specialized vocabulary development, but until now, research into TCM vocabulary has not delineated vocabulary in this way. The current study intends to fill part of this gap by grouping the words in the list according to their linguistic backgrounds to make the list user-friendly for learners with different first languages. Specifically, this study has three goals. The primary aim is to identify specialized vocabulary in Traditional Chinese Medicine and develop a specialized word list from scratch, that is including high-, mid- and low-frequency vocabulary to support the learning of specialized vocabulary in this discipline. The second is to explore a way to divide the long list of specialized vocabulary into smaller sets of vocabulary to accommodate the needs of learners with various L1 backgrounds and vocabulary levels. Finally, this list is compared to Hsu's (2018) TCM English word list to bring to light overlaps and differences between the two TCM word lists to highlight the values of each list for learners.

2. Literature review

2.1. Vocabulary in English-medium TCM texts

Research into vocabulary in written medical English texts has found that they contain a wide range of vocabulary. Hsu (2013) developed a medical textbook corpus (15,016,553 running words) and found that it contained a large number of medical terms (3,474), proper nouns (5,952), and abbreviations (1,427). Another feature of medical textbooks is that they contain fewer high-frequency words than other kinds of texts (Quero & Coxhead, 2018), which is important

when considering the needs of language learners who might have a small vocabulary size and have not had much exposure to long, complex medical texts prior to taking up medical studies in English. This is also important when we consider Nation's (2016a) point that words can be both technical and high-, mid- or low-frequency. This means that high-frequency words might also be technical words in a particular subject area. In Traditional Chinese Medicine, this issue is exacerbated because the meanings of high-frequency words such as *warm* and *cold* are central to the field, but quite different to their meanings in English.

Another feature of medical texts is that they have a high vocabulary load, which means students need to possess a large vocabulary to read English-medium texts in this field. Hsu (2018) found that learners need to know the most frequent 7,000 word families in English to reach 95% coverage of TCM textbooks, and a vocabulary size of 10,000 to reach 98%.[1] Lu and Coxhead (2020) also found that 7,000-8,000 word families plus proper nouns, marginal words, transparent compounds, and abbreviations (four supplementary lists), and TCM-specific words were needed to reach 95% coverage in TCM textbooks and journal articles. To reach 98% of TCM journal articles, 12,000 word families plus supplementary and TCM lists were needed compared to a vocabulary size of 13,000 for textbooks. These figures are important because vocabulary researchers suggest that English as a Foreign Language (EFL) learners tend to have a limited vocabulary (e.g., Henriksen & Danelund, 2015; Lu & Dang, 2022). One way to help learners and teachers develop the vocabulary needed for a discipline is to develop lists of specialized words (Nation, 2016a). Such a list in TCM might be of great help to its learners as well as teachers and practitioners.

2.2. Two approaches to making specialized word lists

The first approach to developing specialized word lists uses a common core approach, meaning the specialized lists are built on top of general high-frequency word lists. This approach was adopted for the development of the Academic Word List (Coxhead, 2000), which excluded the first 2,000 word families of West's (1953) General Service List (GSL). This approach assumes that learners would have mastered the first 2,000 word families of English. Table 1 shows some examples of word lists that use the common core approach for building medical word lists of various kinds, including general medical English, Nursing,

[1] Laufer and Ravenhorst-Kalovski (2010) suggested 95% text coverage as the minimal reading comprehension threshold, while 98% as the optimal reading comprehension threshold. The former referred to the coverage point at which most readers can reasonably comprehend a text, while the latter was the possible coverage point at which a majority of learners can adequately comprehend the text.

and Traditional Chinese Medicine (e.g., Hsu, 2018; Yang, 2015). The final column shows the coverage of each list over its source corpora; in this case more than 10% for all the lists. Note that the word lists are all fairly small, at around 600 or more families, which makes their coverage fairly impressive overall. Smaller lists make for easier adaptation to classrooms. These specialized lists inherit the limitations of the high-frequency lists they are based on. By excluding high-frequency vocabulary, the lists potentially have not included lexical items that have a technical meaning.

Table 1 Medicine-related word lists using the common core approach

Study	Corpus size	Word list	List size	Coverage over source corpora
Yang (2015)	1,006,934	Nursing Academic List	676 word families	13.64%
Wang et al. (2008)	1,093,011	Medical Academic Word List	623 word families	12.24%
Hsu (2018)	13 million	TCM English Word List	605 word families	11.42%
Hsu (2013)	15 million	Medical Word List	595 word families	10.72%

The second approach builds word lists "from scratch" and includes high-frequency words (e.g., the Academic Vocabulary List, Gardner & Davies, 2014). This approach considers specialized vocabulary as a separate set of words which cut across frequency bands in general English (e.g., Dang, 2019b; Lei & Liu, 2016). Table 2 shows examples of such word lists. It is worth noting that the specialized word lists in this table are larger than those in Table 1, and in all cases, their coverage over their source corpora is much higher. This point is important, because it illustrates the powerful role of high-frequency vocabulary in the specialized corpora.

Table 2 Stand-alone medicine-related word lists

Study	Corpus size	Word list	List size	Coverage over source corpora
Fraser (2009)	360,000	Pharmacology Word List	2,000 word families	89.1%
Lei and Liu (2016)	6.2 million	New Medical Academic Word List	965 lemmas	over 19.44%
Dang (2019b)	556,074	Medical Spoken Word List	895 word types	13.44%

Word lists based on the discipline-specific core approach are at risk of repeating words which have been mastered by the learners (Dang et al., 2017). Further, the medically-related word lists in Tables 1 and 2 are mainly based on western medicine and may be of limited help to TCM learners because of its different lexical profile from that of Western medicine (Lu & Coxhead, 2020). One exception is the TCM English word list of Hsu (2018), which is the focus of Section 2.3.

2.3. Hsu's (2018) TCM English word list

Hsu (2018) developed a TCM English word list by identifying the most frequent and evenly distributed word families in a TCM Textbook Corpus that occur in the mid- and low-frequency bands based on the British National Corpus (BNC) and Corpus of Contemporary American English (COCA) and developed by Paul Nation (2012). Hsu selected 605 items for the list based on principles of frequency, dispersion, range, and expert and learner review. It is important to note the common core approach taken here: the most frequent 3,000 word families from the BNC/COCA lists were excluded from this list with the assumption that those words have been acquired by the EFL medical learners. There is, quite possibly, a considerable number of words in those high-frequency lists that could be technical. Quero and Coxhead (2018) found that more than 15% of the items in the first 1,000 of West's (1953) GSL and 17% in Coxhead's (2000) AWL had a medical meaning. Such figures make high frequency vocabulary difficult to ignore in studies of specialized vocabulary. The present study sets out to identify specialized vocabulary in a TCM corpus for a list which is developed from scratch and compare it with Hsu's (2018) headword list. It also aims to examine the coverage of this new list over journal articles, which the students in TCM need to read for their studies, and two types of textbooks. To achieve these goals, this study is guided by the research questions listed below:

1. Which specialized words in the TCM Corpora meet the criteria for inclusion for a TCM Word List?
2. What coverage is provided by Levels 1, 2, and 3 of the TCM Word List, respectively?
3. Does the TCM Word List coverage differ over journal articles and textbooks in the corpus?
4. What coverage does the TCM Word List provide over a second TCM corpus, general English texts, and other medical texts?
5. How does the TCM Word List compare with Hsu's (2018) TCM headword list?

3. Methodology

3.1. Developing the corpora

Three corpora[2] of Traditional Chinese Medicine were developed for this study as shown in Table 3: (1) a theory-based textbook corpus, (2) a practice-based

[2] The TCM Word List was developed, validated, and reported in a Ph.D thesis by the first author (Lu, 2018) under the supervision of the second author. For this article, the comparison with Hsu (2018) was carried out and the conceptualization of the research was extended throughout

textbook corpus, and (3) a journal article corpus. The total corpus size is 3,478,267 words. The textbook corpora represent textbooks used at two different stages of medical education. The theoretical stage (alternatively, pre-clinical stage) comes first, and is followed by the clinical practice stage. Textbooks used at those two stages are different in nature. Because textbooks are typically longer than journal articles, far more articles are needed to make up a similar-sized corpus.

Table 3 Overview of the TCM Corpora

Corpus	Size	Number of texts
Theory-based textbooks	1,171,625	6
Practice-based textbooks	1,109,701	5
TCM journal articles	1,196,941	342
Total	3,478,267	353

A corpus of fiction, magazines and newspapers was developed to serve as a comparison corpus to help identify specialized TCM lexis (see Table 4). It was made up of a roughly equal number of fiction, magazine and newspaper texts, which were not related to medicine. These texts came from the Corpus of Contemporary American English (COCA) (Davies, 2008).

Table 4 Overview of the comparison corpora

Genres	Running words
Fiction	5,980,535
Magazines	5,745,610
Newspapers	5,624,990
Total	17,351,135

To validate the TCM Word List, one validating TCM corpus and one Western Medical corpus were developed. As shown in Table 5, both corpora contain textbooks and journal articles for comparison with the TCM corpora. In principle, the structure of the validating TCM corpus should mirror that of the source corpora from which the word list was built (Dang et al., 2017). Because materials in this discipline are hard to come by, the validating TCM corpus (Table 5) was relatively small. In relation to the Western Medical corpus, the textbook section consists of 3,059,332 tokens in total. These textbooks were obtained from Dr. Betsy Quero (2015), who compiled the corpus for her own research on medical vocabulary. Journal articles were sourced from the COCA medical corpus (Davies, 2008-).

the writing process. The same target TCM corpora were used in Lu and Coxhead (2020) to examine the vocabulary load of these texts and this article appeared in the *ILT-International Journal of Applied Linguistics.* The TCM specialized list was used as a source for target items in an article in *Studies in Second Language Learning and Teaching* (Lu et al., 2021).

Table 5 Overview of the validation corpora

Corpus	Materials	Running words
TCM validating corpus	Textbooks (54,757) Journal articles (409,995)	464,752
Western Medical corpus	Textbooks (3,059,332) Journal articles (1,098,151)	4,157,483

3.2. Developing the TCM Word List

We used word types for the development of TCM Word List because members of word families may not all carry a technical or specialized meaning (Nation, 2016a). To be included in this list, a lexical item had to satisfy criteria of keyness, meaning, and frequency as described in the following paragraphs.

1. *Keyness*: more frequent in the TCM corpus than in the comparison corpus which is much larger (Scott, 1997). Wordsmith Tools 6.0 (Scott, 2012) was used to compute the keyness using chi-square statistics with a p-value of $p < .000001$.
2. *Meaning*: linked semantically to TCM. Based on a study of Chung and Nation (2004, p. 254), a TCM semantic scale was developed to gauge that connection (Table 6).
3. *Frequency*: occur 60 times or more in the TCM Corpora. A comparison of frequency cut-offs at 33, 40, 50, 55, 60 occurrences showed that fewer types at the cut-off point of 60 did not undermine coverage of this list. This process also led to a more manageable word list.

Table 6 The semantic rating scale with TCM examples (adapted from Chung & Nation, 2004, p. 254)

Scales	Description	Examples
Non-specialized	no TCM-related meaning	*due, summary, box, chapter*
Specialized	1. loan word with TCM-related meaning	*wan, yin, chai, fu*
	2. TCM-related meaning but the same or almost the same as general meaning	*pain, sweating, efficacy*
	3. TCM-related meaning but different from the general meaning	*tissue, liver, back*
	4. TCM-related meaning and only found in TCM or general medical science	*acupuncture, apoptosis, moxibustion*

The items which satisfied the above criteria were included in the TCM Word List. Then, RANGE (Heatley et al., 2002) was used to calculate the coverage figures of this list over the TCM Corpora, the validating TCM corpus, and the Western Medical corpus. The TCM Word List was then divided into three sub-

lists based on frequency in the TCM Corpora. Each sub-list was divided into three levels based on the work by Dang et al. (2017), who graded their Academic Spoken Word List into four levels to benefit EAP students with different proficiency levels. This approach was adopted because of the unique features of TCM vocabulary and variations in knowledge and language backgrounds among learners. To compare the TCM Word List with Hsu's (2018) TCM headword list, RANGE was used to create word families. This step resulted in a reduction of our list to a total of 1,958 word families.

4. Results

4.1. The TCM Word List

A total of 2,778 word types met the selection criteria for inclusion in the TCM Word List. It covers 36.64% of the TCM Corpora. Table 7 shows the top 20 specialized words arranged by alphabet and their frequency in the TCM Corpora. The table also illustrates features and origins of TCM vocabulary. The loan words (e.g., *qi*, *yang*) are in bold to make them more visible. They appear with extremely high frequency in the TCM texts and are thus key to all learners. Additionally, some obviously general words such as *group* and *used* also appear in the list because they show a degree of technicality when they appear as part of highly specialized multiword units in the TCM Corpora. For example, *group* is a node word in common collocations such as *treatment group, acupoint group,* and *acupuncture group*, whereas *used* is often found in combinations with acupoints and herbal medicine such as *SP-9 were used,* as well as other collocational patterns such as *widely used*, and *formula(herb) used*.

Table 7 The most frequent 20 items by alphabet in the TCM Word List

Specialized words	Occurrences	Specialized words	Occurrences
acupuncture	8,484	*pain*	10,252
blood	13,924	*patients*	7,627
cold	6,415	*points*	6,229
deficiency	8,956	*qi*	21,866
group	7,311	*radix*	7,294
heart	6,760	*spleen*	6,878
heat	12,039	*treatment*	13,433
kidney	6,945	*used*	7,749
liver	9,976	*yang*	9,800
medicine	6,276	*yin*	10,453

The full TCM Word List was then divided by their frequency of occurrences in the TCM Corpora into three sub-lists to make the list more manageable from a

pedagogical perspective. Table 8 shows that Sub-list 1 with the first 1,000 items has much greater coverage of the TCM corpus (29.97%) than Sub-list 2 with the next most frequent 1,000 types at nearly 5% coverage. Such a coverage pattern is typical of word lists in the target corpora which they represented (Coxhead, 2000; Nation, 2016a). Unsurprisingly, Sub-list 3 has the lowest coverage because it contains only 778 items which are less frequent overall. These coverage figures illustrate the power of high-frequency specialized vocabulary in TCM.

Table 8 The sub-lists of the TCM Word List

Sub-list	Word types	Coverage	Examples
1	1,000	29.97%	*qi, blood, treatment*
2	1,000	4.89%	*detected, laboratory, geng (根)*
3	778	1.78%	*ultrasound, vein, fuling (茯苓)*

Note. Sub-lists 1, 2, 3 contain the most frequently-occurring 1-1000, 1001-2000, and 2001-2778 word types in the TCM Corpora, respectively.

4.2. Levels of the TCM Word List

Table 9 presents the number and coverage of specialized vocabulary in each level of the sub-lists. The sub-lists of the TCM Word List were divided into three levels primarily according to their frequency in the BNC/COCA lists by Nation (2012). Level 1 contains items from the 1st-3rd 1,000 word families of the BNC/COCA; Level 2 contains those which fell outside the 1st-3rd 1,000; and Level 3 is for loan words from Chinese. It is important to mention that some loan words (e.g., *yin, chai*) appear at Level 2 because of semiotic coincidences[3] between English and Romanized Chinese (see Lu & Coxhead, 2020 for more details). Although a considerable number of items (1,325 word types out of 2,778) in the TCM Word List are at Level 1 from the first 3,000 BNC/COCA word families, the other 1,245 words fell outside the first 3,000 BNC/COCA word families at Level 2, and the remaining 218 items are loan words at Level 3. This result means that nearly half (47.7%) of the items from the list developed from the present study also occur in general high-frequency vocabulary.

Dividing the list into sub-lists and levels means that frequency is taken into account both in terms of TCM specialized vocabulary and general English vocabulary. Loan words are also taken into account in the levels. The top ten Chinese specialized lexical items at Level 3 by frequency are *geng, bei, xuehai, waiguan, mingmen, deqi, bian, shenmen, yue* and *shou*. The TCM specialized words in

[3] 3 Semiotic coincidences refer to items (e.g., *tang, ling*) which can either be regarded as an English word or a loan word borrowed from Chinese, depending on the context of use (Lu & Coxhead, 2020).

each sub-list and level are presented in Appendices A to I. It means the list can be adaptable to learners from different L1 backgrounds (e.g., English L1 speakers and Chinese L1 speakers) or at different vocabulary levels.

Table 9 Number and coverage of specialized words at each level of TCM sub-lists

Level	Level 1		Level 2		Level 3	
	number	coverage	number	coverage	number	coverage
Sub-list 1	643	19.39%	289	8.13%	68	2.45%
Sub-list 2	465	2.33%	468	2.23%	67	0.33%
Sub-list 3	217	0.49%	478	1.1%	83	0.19%
Total	1,325	22.21%	1,235	11.46%	218	2.97%

Note. Number refers to number of word types

4.3. Coverage of the TCM Word List in the theory- and practice-based textbooks, and journal articles

The TCM Word List provides the highest coverage over the two sets of textbooks in the corpora: 40.06% over practice-based and 38.54% over theory-based ones. Its coverage over the journal articles is somewhat lower at 31.43%. These patterns are consistent with the findings of Lu and Coxhead (2020), using the same corpora, who found that the journal articles were less demanding than the textbooks and had more high-frequency vocabulary. It is important to note that the TCM Word List in the three types of TCM texts is all over 30%, which is similar to findings of technical vocabulary coverage by Chung and Nation (2003) in an anatomy textbook and Coxhead et al. (2016) in pedagogical written carpentry texts.

4.4. Coverage of the TCM Word List in other corpora

The TCM Word List covers exactly the same percentage over the TCM validating corpus (31.43%) as it does over the TCM journal article corpus. Its coverage over the Western Medical texts is 25.32%, which is largely achieved by frequently-occurring words such as *antibiotics, diagnoses,* and *cell.* Its coverage over the corpus of general English (fiction, newspapers and magazines) is 4.11%, and items such as *brain*, *arm* and *chest* are largely responsible for that coverage.

4.5. Comparison with Hsu's (2018) headword list

In total, 309 words overlap between Hsu's (2018) TCM headword list (605 items) and our TCM Word List. *Acupoint, damp,* and *evil* are examples of words which are not in Hsu's list; along with high-frequency items such as *warm* and *wind*. Hsu's TCM headwords account for 4.87% and 3.71% of the tokens in the TCM

Corpora and the TCM validating corpus, respectively. A total of 579 and 480 out of 605 headwords occur in the TCM Corpora and validating corpus. Its coverage over the TCM Textbook Corpora from our study is slightly higher at 5.71% (570 out of 605). For interest, we also checked coverage of Hsu's TCM headword list over the Western Medical Corpora (3.81%), and the general written English Corpora (0.23%). Note that the coverage of Hsu's (2018) list is lower in the corpora of our study because we only have the headword list rather than the whole list. These coverage figures do not undermine the value of Hsu's (2018) list in any way. Appendix J presents a Common Core TCM List of shared items between Hsu's (2018) headword list and the TCM Word List. This list is of paramount importance to all learners in this discipline.

5. Discussion

In response to research question 1, the findings in this study showed that there is a large amount of specialized vocabulary (2,778 word types) in Traditional Chinese Medicine texts, and the lists discussed here did not include lexical items that are technical but did not meet the selection criteria for the final word list. Especially, the development of specialized word lists using either a common core or discipline-specific approach is important to discuss here. This study shows that starting from scratch for selecting items for a TCM word list results in, firstly, a much larger word list, and secondly, much higher coverage than a word list that does not include high-frequency technical items. Further, the TCM Word List presented in this article took a wide view of specialized words, including Chinese loan words as well as high-, mid- and low-frequency words that met the criteria for inclusion in the list. Discipline-specific word lists can provide a short-cut to the most useful words in various disciplines by identifying and ranking lexical items to help with academic reading texts that usually have heavy vocabulary loads (Nation, 2013, 2016b). The findings of this article provide strong support for this approach. TCM students can be motivated to learn words from such specialized lists because they can clearly see the relationship between what they can learn from this list and their subject courses, which is particularly important for English for Specific Purposes (ESP) learners (Coxhead & Hirsh, 2007; Hyland, 2016). These findings thus provide more solid evidence for how specialized word lists can better serve ESP learners' needs (Durrant, 2013, 2016; Hyland & Tse, 2007).

In answer to research question 2, the coverage figures of the three levels of TCM Word List can be found in Table 9. There are several points that are worth attention. First, high-frequency vocabulary which has a technical meaning in TCM (Sub-list 1, Level 1 of the TCM list) is vital for learners. The coverage of the TCM source corpus at just over 19% by 643 types (see Table 9) demonstrates

this point. Creating the TCM specialized list from scratch resulted in (1) a much larger list than Hsu's (2018) list, and (2) quite different coverage figures. Our TCM Word List covers more than one third of the TCM Corpora in this study, compared to nearly 5% of Hsu's (2018). The corpora used in the current study, however, is not large and the validation corpus could be better balanced to make it more similar to the source corpora from which the TCM Word List was developed. Second, the TCM Word List illustrates how different L1 backgrounds and vocabulary knowledge can be taken into account through dividing the list into levels and sub-lists. The Chinese loan words are at their own level and can be the target of the non-Chinese speakers, in particular the most frequent loan words. Last, the division also makes the longer word list more manageable for teachers and learners (Coxhead, 2000; Dang & Webb, 2016), and may help encourage learners to study so as to progress from items in their current vocabulary knowledge to items beyond it (Webb & Chang, 2012).

With regard to research question 3, the present study showed that the TCM Word List covers 36.64% of its source corpus, and over 30% of the individual corpora of journal articles, practice, and theory textbooks in TCM. This means the learners will encounter TCM Word List items frequently in these texts. As already noted, these findings resonate with other studies which use different approaches to identifying discipline-specific vocabulary for word lists (Chung & Nation, 2004; Coxhead & Demecheleer, 2018; Coxhead et al., 2016).

In answer to research question 4, this study revealed that the coverage of the TCM Word List in the TCM validating corpus and the TCM journal article corpus is the same. This is possibly because the validating corpus contains mostly journal articles. This finding demonstrates that this list captures the most useful specialized words in this field. In addition, the TCM Word List provides higher coverage over the TCM Corpora than the Western Medical corpus. It suggests that this list better represents specialized vocabulary in TCM rather than in Western Medicine. Further, the lower coverage of the list over the general written English corpus indicates that this list contains words which are relatively rare in general written English. Again, these results are consistent with results of other specialized word lists over validating corpora (e.g., Coxhead, 2000; Dang et al., 2017; Gardner & Davies, 2014).

Concerning research question 5, the TCM Word List developed from this study has a number of advantages over that of Hsu (2018). First, it included high-frequency vocabulary. This is of paramount importance as recent studies have consistently reported that EFL learners are unlikely to have mastery of high-frequency vocabulary in various EFL contexts (e.g., China, Iran, Denmark, Vietnam) at secondary school (e.g., Laufer, 1998; Nguyen, 2020; Olmos, 2009; Stæhr, 2008; Sun & Dang, 2020), and university levels (e.g., Lu & Dang, 2022; Nguyen & Webb, 2017; Nurweni & Read, 1999). In particular, those studies have shown that EFL

university students have not mastered the first 2,000 word families in English (e.g., Akbarian, 2010; Dang, 2019a; Matthews & Cheng, 2015). Moreover, Lu et al. (2021) revealed that some TCM learners have difficulty understanding technical meanings of the 1st-3rd 1,000 word families (e.g., *channel*). The TCM Word List can help address this issue by ensuring that high-frequency items are included in Level 1. Second, Level 2 of this list also includes items which fell outside the high-, mid-, and low-frequency BNC/COCA levels. Such items are important in TCM but are not so frequent, meaning learners might not meet them often in general language use (Ardasheva & Tretter, 2015; Nation, 2013). Last, the inclusion of the loan words at Level 3 makes this unique set of vocabulary salient to learners from different linguistic backgrounds. Their relative value in this discipline is now more visible in comparison to other specialized words.

Overall, this study is among the few word list studies which systematically investigated specialized vocabulary in Traditional Chinese Medicine, including vocabulary across the Chinese and English languages. It also provides insights into the possible ways in which the specialized word list can be adapted to TCM students from different L1 backgrounds and existing vocabulary levels.

6. Pedagogical implications

This study has several implications for a range of possible users of the frequency-based TCM Word List, including materials writers, teachers, and learners. For example, the list can guide the setting of goals for teachers, learners and course designers over short and long terms (Nation, 2013). Working with any list has to be more than recitation and rote learning, although such memorization activities can be useful. A principled approach is needed to take into account the features of the vocabulary in the TCM Word List and the learners' lexical knowledge and language backgrounds. First, the current TCM Word List contains high-frequency specialized words that were not included by Hsu's (2018) list, meaning the current list may be more useful for learners at the initial stages of study. At the beginning of a TCM course, the teachers can establish the learning goals depending on their students' L1 backgrounds or existing levels of general vocabulary. New learners to this field could begin, for example, with Sub-list 1, Level 1 and identify words and the TCM meanings they already know. Importantly, all learners would need to ensure that they understand and are able to express the TCM meaning of high-frequency words, rather than relying on the everyday meaning of these words. Examples of such kind of high-frequency items with specialized meanings are *fire*, *water,* and *flow*. Second, speakers of languages other than Chinese might want to focus on learning to recognize and understand the Chinese loan words in particular, since they would need to learn both the

word form and TCM-related concept. Examples of such words include *shenshu*, *zao* and *xuan*. As suggested by one of the reviewers, Chinese loan words in relation to acupoints and herbs may be better learnt when they are accompanied by pictures, video, or realia to non-Chinese native speakers. Another way to group the learning of the specialized items would be through the main subfields, such as herbal medicine (e.g., *wallichii, moxa*), acupuncture (e.g., *acupressure*, *needle*) and Chinese massage (e.g., *qi*, *manipulation*). Third, the TCM Word List includes specialized words that are shared with Western medicine. English as L1 speakers, L2 learners, and EFL learners who are learning TCM would need help with learning these words because they might not have had much exposure to these long, Graeco-Latin items previously. There are clearly several such categories or groups of items in the list, such as organs and parts of the body (e.g., *kidney*, *heart*, *knees*) and adjectives (e.g., *empty*, *fried*, *choppy*), that many people might already recognize in general English or from exposure to medical English. These lexical items would require checking to ensure that the discipline-specific meaning is also known. Medical terms that are shared with Western medicine, such as *anti-inflammatory*, *cardiovascular*, and *analgesic*, could be learned through developing knowledge of roots and affixes (Nation, 2013), because these words are often made up of word parts which learners need to learn to recognize and use. For example, the root *cardi(o)-* (meaning *heart*), the affix *-vascul* (meaning *vessel*), and the affix *-ar* (meaning *pertaining to*) make up *cardiovascular*. *Cardi(o)-* can also combine with *-tocography* to form *cardiotocography* (Hutton, 2006). Knowing the meaning of roots and affixes and being able to identify them in longer medical words is important for word recognition skills. Learners can use this knowledge to guess the meaning of an unfamiliar word from context, and then check whether the meaning of this word has been successfully inferred (Nation, 2013).

The TCM Word List provides a point of reference for teachers. This is important in ESP for teachers who might not have expertise in TCM but are required to teach it. These practitioners can use the list to identify specialized words in texts according to the levels, design materials so that they contain and repeat specialized words to help with learning, and organize learning activities in relation to specialized words. This can help the learners to develop and consolidate TCM specialized vocabulary knowledge in meaningful ways. They can highlight specialized words in the TCM texts, thereby causing the words to stand out. Such technique will likely draw the attention of the learners. It is important to ensure that learners work on developing their TCM vocabulary across Nation's (2007) Four Strands of meaning-focused input (reading and listening), language-focused learning, meaning-focused output (writing and speaking), and fluency development. The list and the strands can provide a framework for ESP

courses for activities, course design and instruction (see Hirsh & Coxhead, 2009; Nation & Yamamoto, 2012 for more).

Last, items in the TCM Word List can act as a guide in the assessment and testing of specialized vocabulary knowledge (Nation, 2016a). For instance, test developers could sample items from each sub-list and level to write the lexical components of a test. The testing and assessment, however, should be aligned with the overall learning goals as well as the actual activities in teaching and learning specialized vocabulary (Malmström et al., 2018). For example, if the teaching activities only involve receptive learning (e.g., form-meaning recognition) of certain specialized words, the testing should reflect such receptive aspects of knowing the target words.

7. Limitations and directions for future research

Limitations of this study include the relatively small corpora and the focus on single words over multiword units. Future studies could use larger corpora to further validate the list. There are several possible directions for future research, including analyzing technical abbreviations in various academic disciplines, and looking into the vocabulary used in other types of specialized language, including class handouts and multimodal texts. Similarly, studies in medical contexts could also investigate vocabulary in spoken discourse (see Dang, 2018 and Lessard-Clouston, 2010 for examples). These types of discourse are of great importance to the development of disciplinary literacy. In addition, it is equally important to explore specialized multiword units in ESP contexts, which might provide some new insights into the nature of specialized vocabulary. Finally, research that looks into the learning and testing of specialized vocabulary would be valuable.

8. Conclusion

This study reported on the development and validation of a specialized word list using a discipline-specific approach in Traditional Chinese Medicine, which draws on vocabulary across languages (i.e., Chinese and English). The TCM Word List is large, with 2,778 items, but has been classified into various sub-lists and levels to make it more manageable. It has similar coverage over a second TCM corpus and low coverage over a corpus of general English, which suggests that it captures items that are specific to this discipline. By comparison with Hsu's (2018) TCM headword list, 309 common core headwords were identified. This study illustrates the large amount of specialized vocabulary, and suggests ways in which that teachers and learners might work with a specialized word list to help with learning this essential vocabulary in that discipline. Further, the approach

used to develop and divide a discipline-specific word list provides a possible direction for future work on specialized vocabulary which targets at learners from various linguistic backgrounds and with different vocabulary levels.

ACKNOWLEDGEMENTS

We would like to thank Professor Mirosław Pawlak, Professor Joanna Zawodniak and the two anonymous reviewers for their useful feedback, which has improved the quality of this article in different ways. We would also like to express our sincere thanks to their support on the topic of specialized vocabulary in the discipline of Traditional Chinese Medicine.

References

Akbarian, I. H. (2010). The relationship between vocabulary size and depth for ESP/EAP learners. *System*, *38*(3), 391-401. https://doi.org/10.1016/j.system.2010.06.013

Ardasheva, Y., & Tretter, T. R. (2015). Developing science-specific, technical vocabulary of high school newcomer English learners. *International Journal of Bilingual Education and Bilingualism*, *20*(3), 1-20.

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

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

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

Coxhead, A., & Hirsh, D. (2007). A pilot science-specic word list. *Revue Française de Linguistique Appliqueé*, *7*(2), 65-78.

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

Dang, T. N. Y. (2019a). High-frequency words in academic spoken English: Corpora and learners. *ELT Journal*, *74*(2), 146-155. https://doi.org/10.1093/elt/ccz057

Dang, T. N. Y. (2019b). 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. https://doi.org/10.1002/tesq.552

Dang, T. N. Y., & Webb, S. (2016). Evaluating lists of high-frequency words. *ITL - International Journal of Applied Linguistics, 167*(2), 132-158.

Dang, T. N. Y., Coxhead, A., & Webb, S. (2017). The Academic Spoken Word List. *Language Learning*, *67*(4), 959-997. https://doi.org/10.1111/lang.12253

Davies, M. (2008-). *The Corpus of Contemporary American English (COCA): 560 million words, 1990-present*. http://corpus.byu.edu/coca/

Durrant, P. (2013). Discipline and level specificity in university students' written vocabulary. *Applied Linguistics*, *35*(3), 328-356. https://doi.org/10.1093/applin/amt016

Durrant, P. (2016). To what extent is the Academic Vocabulary List relevant to university student writing? *English for Specific Purposes*, *43*, 49-61.

Evans, S., & Morrison, B. (2011). Meeting the challenges of English-medium higher education: The first-year experience in Hong Kong. *English for Specific Purposes*, *30*(3), 198-208. https://doi.org/10.1016/j.esp.2011.01.001

Flowerdew, J., & Miller, L. (1992). Student perceptions, problems and strategies in second language lecture comprehension. *RELC Journal*, *23*, 60-80. https://doi.org/10.1177/003368829202300205

Fraser, S. (2009). Breaking down the divisions between general, academic and technical vocabulary: The establishment of a single, discipline-based word list for ESP learners. *Hiroshima Studies in Language and Language Education*, *12*, 151-167.

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

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

Henriksen, B., & Danelund, L. (2015). Studies of Danish L2 learners' vocabulary knowledge and the lexical richness of their written production in English. In P. Pietilä, K. Doró, & R. Pípalová (Eds.), *Lexical issues in L2 writing* (pp. 29-56). Cambridge Scholars Publishing.

Hirsh, D., & Coxhead, A. (2009). Ten ways of focussing on science-specific vocabulary in EAP. *English Australia Journal*, *25*(1), 5-16.

Hsu, W. H. (2013). Bridging the vocabulary gap for EFL medical undergraduates: The establishment of a medical word list. *Language Teaching Research*, *17*(4), 454-484. https://doi.org/10.1177/1362168813494121

Hsu, W. (2018). The most frequent BNC/COCA mid- and low-frequency word families in English-medium traditional Chinese medicine (TCM) textbooks. *English for Specific Purposes*, *51*, 98-110. https://doi.org/10.1016/j.esp.2018.04.001

Hu, Y.-H., & Fong, Y.-Y. (2010). Obstacles to CM-guided L2 idiom interpretation. In S. D. Knop, F. Boers, & A. D. Rycker (Eds.), *Fostering language teaching efficiency through cognitive linguistics* (pp. 293-316). DeGruyter.

Hutton, A. R. (2006). *An introduction to medical terminology for health care: A self-teaching package*. Churchill Livingstone.

Hyland, K. (2016). General and specific EAP. In K. Hyland & P. Shaw (Eds.), *The Routledge handbook of English for academic purposes* (pp. 17-29). Routledge.

Hyland, K., & Tse, P. (2007). Is there an "academic vocabulary?" *TESOL Quarterly*, *41*(2), 235-253.

Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics*, *19*(2), 255-271.

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign. Language*, *22*(1), 15-30.

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. https://doi.org/10.1016/j.jeap.2016.01.008

Lessard-Clouston, M. (2010). Theology lectures as lexical environments: A case study of technical vocabulary use. *Journal of English for Academic Purposes, 9*(4), 308-321. https://doi.org/10.1016/j.jeap.2010.09.001

Lin, A., Chan, G., Hu, Y., Ouyang, D., Ung, C., Shi, L., & Hu, H. (2018). Internationalization of traditional Chinese medicine: Current international market, internationalization challenges and prospective suggestions. *Chinese Medicine*, *13*(9), 1-6.

Lu, C. (2018). *Investigating knowledge and use of technical vocabulary in Traditional Chinese Medicine* [Unpublished doctoral dissertation, Te Herenga Waka – Victoria University of Wellington].

Lu, C., & Coxhead, A. (2020). Vocabulary in Traditional Chinese Medicine: Insights from corpora. *ILT-International Journal of Applied Linguistics*, *171*(1), 34-61.

Lu, C., & Dang, T. N. Y. (2022). Vocabulary in EAP learning materials: What can we learn from teachers, learners, and corpora? *System, 106,* 102791. https://doi.org/10.1016/j.system.2022.102791

Lu, C., Boers, F., & Coxhead, A. (2021). Exploring learners' understanding of technical vocabulary in Traditional Chinese Medicine. *Studies in Second Language Learning and Teaching*, *11*(1), 71-101. https://doi.org/10.14746/ssllt.2021.11.1.4

Maciocia, G. (2005). *The foundations of Chinese medicine: A comprehensive text for acupuncturists and herbalists* (2nd ed.). Elsevier Churchill Livingston.

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.

Matthews, J., & Cheng, J. (2015). Recognition of high frequency words from speech as a predictor of L2 listening comprehension. *System*, *52,* 1-13. https://doi.org/10.1016/j.system.2015.04.015

Nation, P. (2007). The four strands. *Innovation in Language Learning and Teaching*, *1*(1), 1-12.

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

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

Nation, P. (2016a). *Making and using word lists for language learning and testing*. John Benjamins Publishing Company.

Nation, P. (2016b). Hyphenated words and transparent compounds. In P. Nation (Ed.), *Making and using word lists for language learning and testing* (pp. 65-70). John Benjamins.

Nation, I. S. P., & Yamamoto, A. (2012). Applying the four strands. *International Journal of Innovation in English Language Teaching and Research*, *1*(2), 167-181.

Nguyen, C.-D. (2020). Lexical features of reading passages in English-language textbooks for Vietnamese high-school students: Do they foster both

content and vocabulary Gain? *RELC Journal*, 1-14. https://doi.org/10.1177/0033688219895045.

Nguyen, T. M. H., & Webb, S. (2017). Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research*, *21*(3), 298-320.

Nurweni, A., & Read, J. (1999). The English vocabulary knowledge of Indonesian university students. *English for Specific Purposes*, *18*(2), 161-175.

Olmos, C. (2009). An assessment of the vocabulary knowledge of students in the final year of secondary education. Is their vocabulary extensive enough? *International Journal of English Studies*, *9*, 73-90.

Patel, A., & Toossi, V. (2016). Traditional Chinese medicine practitioners in New Zealand: Differences associated with being a practitioner in New Zealand compared to China. *NZMJ*, *129*(1444), 48-55.

Quero, B. (2015). *Estimating the vocabulary size of L1 Spanish ESP learners and the vocabulary load of medical textbooks* [Unpublished doctoral dissertation, Te Herenga Waka – Victoria 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.

Scott, M. (1997). PC analysis of key words: And key key words. *System*, *25*(2), 233-245. https://doi.org/10.1016/S0346-251X(97)00011-0

Scott, M. (2012). *WordSmith tools version 6*. Lexical Analysis Software Ltd.

Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, *36*(2), 139-152. https://doi.org/10.1080/09571730802389975

Sun, Y., & Dang, T. N. Y. (2020). Vocabulary in high-school EFL textbooks: Texts and learner knowledge. *System*, *93*, 1-13. https://doi.org/10.1016/j.system.2020.102279

Wang, J., Liang, S. L., & Ge, G. C. (2008). Establishment of a Medical Academic Word List. *English for Specific Purposes, 27*(4). 442-458. https://doi.org/10.1016/j.esp.2008.05.003

Webb, S., & Chang, A. C.-S. (2012). Second language vocabulary growth. *RELC Journal*, *43*(1), 113-126.

West, M. (1953). *A general service list of English words*. Longman.

Wette, R., & Hawken, S. J. (2016). Measuring gains in an EMP course and the perspectives of language and medical educators as assessors. *English for Specific Purposes*, *42*, 38-49. https://doi.org/10.1016/j.esp.2015.11.002

Yang, M.-N. (2015). A nursing academic word list. *English for Specific Purposes*, *37*, 27-38.

APPENDIX A

Sub-list 1 of the TCM Word List at Level 1

Note. The word types were arranged according to their frequency in the TCM Corpora. The concordance lines of word types falling into a single lemma (e.g., *point, points*) were checked to see if they had similar meanings and patterns of use in the TCM Corpora. If so, they were combined into a single entry in the word list as *point(s)*.

1. blood	47. skin	93. practice	139. period(s)
2. treatment	48. condition(s)	94. addition(s)	140. nature
3. heat	49. cause(s/d/ing)	95. primary	141. pressure
4. pain	50. level(s)	96. throat	142. form(s)
5. use(d)/using	51. treat(ed/ing/s)	97. signs	143. poor
6. patient(s)	52. organ(s)	98. excessive	144. difficult
7. group(s)	53. cell(s)	99. rat(s)	145. stimulation
8. heart	54. feeling	100. area(s)	146. loss
9. cold	55. factor(s)	101. expression(s)	147. lead(ing/s)
10. medicine(s)	56. compared	102. swelling	148. interior
11. point(s)	57. participants	103. full	149. performed
12. symptom(s)	58. severe	104. test(s)	150.practitioner(s)
13. study	59. excess	105. sweat(ing/s)	151. action(s)
14. body	60. therapy/therapies	106. mental	152. mean
15. wind	61. medical	107. upper	153. lack
16. tongue	62. significantly	108. brain	154. relationship(s)
17. stomach	63. manifestation(s)	109. mouth	155. intervention(s)
18. clinical	64. flow	110. powder	156. transformation
19. channel(s)	65. coating	111. pale	157. thirst
20. low(er)	66. health	112. difference(s)	158. development(s)
21. damp	67. painful	113.increase(d/s/ing)	159. outcome(s)
22. formula(s/e)	68. system(s)	114. location	160. resolve(s/d)
23. effect(s)	69. internal	115. stage(s)	161. invasion
24. disease(s)	70. food(s)	116. changes	162. process(es)
25. dampness	71. include(d/ing)	117. indication(s)	163. neck
26. fire	72. women	118. emotional	164. score(s)
27. result	73. depression	119. warm(ing)	165.supplement(ing)
28. pattern(s)	74. effective	120. movement	166. soul
29. control(led/s)	75. normal	121. thin	167. hot
30. case(s)	76. quality(ies)	122. combined	168. evidence
31. chest	77. principle(s)	123. descend(ing/s)	169. connecting
32. clear(s/ing)	78. indicate(d/s/ing)	124. rate	170. strong
33. vessel(s)	79. traditional	125. combination(s)	171. decrease(d/s)
34. add(ed)	80. yellow	126. fullness	172. protein(s)
35. related	81. activity/activities	127. deep	173. drug(s)
36. water	82. disorder(s)	128. rapid	174. similar
37. trail(s)	83. weak	129. min	175. reduce(d/ing/s)
38. mind	84. empty	130. muscle(s)	176. sleep
39. red	85. associated	131. root	177. spirit
40. data	86. shan	132. risk	178. theory
41. method(s)	87. cough(ing)	133. problems	179. major
42. significant	88. model(s)	134. external	180. source
43. function(s)	89. middle	135. criterion(a)	181. statistical
44. western	90. age	136. region	182. stroke
45. main	91. bleeding	137. cancer	183. energy(ies)
46. dry	92. burner(s)	138. physical	184. injury

185. modern	240. vision	295. spiritual	350. index
186. pregnancy	241. knee(s)	296. breath(ing)	351. negative
187. eye	242. difficulty	297. develop(ed)	352. presence
188. affect(ed/s/ing)	243. eating	298. breast	353. element(s)
189. observed	244. original	299. loose	354. surgery
190. characteristics	245. weakness	300.demonstrate(d/s)	355. modifications
191. ye	246. bitter	301. strength(s)	356. uncooked
192. cycle	247. comparison(s)	302. approach	357. complex
193. stress	248. characteriz(s)ed	303. injection(s)	358. identification
194. tissue(s)	249. ability	304. doctors	359. restore(s/ing)
195. ban	250. discharge(s)	305. anger	360. tip
196. applied	251. damage(s)	306. assess(ed/ing)	361. failure
197. follow	252. leg(s)	307. located	362. products
198. improvement(s)	253. resulting	308. defined	363. formation
199. response(s)	254. rising	309. mild	364. sang
200. drain(ing)	255. aspect(s)	310. oral	365. female(s)
201. influence(s)	256. moves	311. ingredients	366. divided
202. weight	257.randomiz(s)ed	312. reduction	367. experienced
203. frequent	258. events	313. desire	368. collected
204. standard(s)	259. measure(d/s)	314. average	369. immune
205. anxiety	260.preparation(s)	315. floating	370. male(s)
206. diet	261. sexual	316. prevent(s)	371. determine(d)
207. dryness	262. bone(s)	317. balance	372. examination
208. involved	263. memory	318. directly	373. relieve
209. mice	264. ear(s)	319. evaluation	374. heaven
210. conducted	265. position	320. stops	375. relevant
211. positive	266. exercise	321. production	376. scientific
212. transform(ing)	267.identify(ied/ing)	322. reinforcing	377. enters(ing)
213. surface	268. scale(s)	323. application	378. colour
214. benefit(s)	269. variation	324. accompanied	379. dependent
215. inability	270. animal(s)	325. administration	380. release(s)
216. phase(s)	271. swollen	326. remedy	381. gene(s)
217. hospital(s)	272. individual	327. basis	382. mainly
218. sample(s)	273. occur(s)	328. growth	383. produce(d)
219. subjects	274. calm(s)	329. safety	384. significance
220. potential	275. rebellious	330. fur	385. underlying
221. temperature	276. nose	331. breathlessness	386. centre
222. dose(s)	277. experimental	332. strategies	387. limited
223.improve(d/ing/s)	278. illness	333. direct(ing)	388. suffering
224. arm	279. provide(d/s)	334. provide(d/s)	389. concentration
225. healthy	280. infertility	335. consumption	390. drink
226. size(s)	281. conventional	336. daily	391. absence
227. mechanism(s)	282. circulation	337. alternative	392. network
228. receive(d/ing/s)	283. joint(s)	338. purple	393. redness
229. shoulder	284. stiffness	339. nerve(s)	394. opens
230. intension	285. current	340. irregular	395. shortness
231. governing	286. sudden	341. effectiveness	396. management
232. defensive	287. obtained	342. intensity	397. presented
233. sea	288. strengthen(s)	343. relatively	398. branch
234. infection(s)	289. active	344. restlessness	399. marked
235. taste	290. thick	345. evaluate(ing)	400. specifically
236. value(s)	291. eliminate(s/ed)	346. objective	401. population
237. frequency(ies)	292. heaviness	347. central	402. origin
238.promote(s/ing)	293. functional	348. emotion(s)	403. pronounced
239.regulate(s/ing/d)	294. tiredness	349. useful	404. expressed

405. sadness	431. technique	457. depth	483. users
406. conception	432. variable(s)	458. examine(d)	484. agents
407. constitution	433. excluded	459. transporting	485. linking
408. random	434. proper	460. medicinal	486. correspond(s)
409. severity	435. smooth	461. insufficient	487.measurement(s)
410. ache(s)	436. modified	462. responsible	488. psychological
411. moderate	437. core	463. birth	489. reaction
412. containing	438. fried	464. correlation(s)	490. substances
413. burning	439. status	465. consciousness	491. sensitivity
414. extract(s/ed)	440. gate	466. corresponding	492. confirmed
415. alcohol	441. understanding	467. prevention	493. extreme
416. selection	442. chemical	468. revealed	494. blinding
417. mental	443. distribution	469. nervous	495. contrast
418. aged	444. downward(s)	470. stepping	496. investigate(d)
419. importance	445. clinic(s)	471. chosen	497. sore
420. compounds	446. dosage(s)	472. recovery	498. consistent
421. extraordinary	447. fright	473. exhaustion	499. superior
422. visual	448. signaling	474. structure(s)	500. typical
423. carried	449. unable	475. fixed	501. clinically
424. components	450. reflect(s/ing)	476. procedure(s)	502. biological
425. relative	451. helps	477. repeated	503. guidelines
426. chills	452. establish(ed)	478. protective	504. wasting
427. classical	453. regulation	479. sequence	505. interaction(s)
428. inner	454. masses	480. manifest	506. initial
429. degree	455. suitable	481. complications	507. tendency
430. randomly	456. observation(s)	482. relation	

APPENDIX B

Sub-list 1 of the TCM Word List at Level 2

1. yin	51. induced	101. sclerotium	151. protocol(s)
2. liver	52. needling	102. duration	152. upward(s)
3. deficiency(ies)	53. moxibustion	103. menstrual	153. marrow
4. acupuncture	54. cortex	104. acid(s)	154. activation
5. radix	55. limb(s)	105. glycyrrhizae	155. nasal
6. kidney(s)	56. herba	106. appetite	156. physiological
7. spleen	57. moxa	107. lumbar	157. ovulation
8. phlegm	58. sensation(s)	108. disharmony(ies)	158. chemotherapy
9. pulse(s)	59. constipation	109. angelicae	159. diagnosed
10. decoction(s)	60. efficacy	110. cocos	160. spine
11. lung(s)	61. stool(s)	111. officinalis	161. sperm
12. herb(s)	62. uterus	112. chinensis	162. subtractions
13. syndrome(s)	63. nausea	113. poriae	163. agitation
14. tang	64. penetrating	114. citri	164. astragali
15. rhizoma	65. lateral	115. oedema	165. bi
16. chronic	66. insomnia	116. distil	166. prescribe(d)
17. stagnation	67. acupoint(s)	117. palpitation(s)	167.macrocephalae
18. fluid(s)	68. glucose	118. complexion	168. oppression
19. fu	69. pathology(ies)	119. atractylodis	169. physician(s)
20. herbal	70. adverse	120. numbness	170. pericarpium
21. stasis	71. pathological	121. median(s)	171. ascend(s)
22. fructus	72. vacuity	122. ethereal	172. pathogen(s)
23. abdominal	73. differentiation	123. sinensis	173. pathways
24. bladder	74. dong	124. oedema	174. epilepsy
25. pathogenic	75. sham	125. ginseng	175. tuber
26. tonify(ies/ing)	76. hypertension	126. aversion	176. nutritive
27. exterior	77. therapeutic	127. sinew(s)	177. metabolism
28. urine	78. mucus	128. paeoniae	178. ramulus
29. ling	79. tian	129. insulin	179. blurred
30. obstruction(s)	80. retention	130. administered	180. cerebral
31. intestine(s)	81. serum	131. medicinals	181. wheezing
32. prescription(s)	82. asthma	132. uralensis	182. febrile
33. deficient	83. placebo	133. greasy	183. medication(s)
34. essence(s)	84. fang	134. orifice(s)	184. seu
35. acute	85. distension	135. diabetic	185. apoptosis
36. diagnosis	86. sticky	136. flos	186. cinnamomi
37. fever	87. wiry	137. inflammatory	187. elbow
38. urination	88. diagnostic	138. turbid	188. contraction
39. pill(s)	89. fatigue	139. backache	189. uterine
40. nourish(es/ing)	90. rehmanniae	140. menstruation	190. inclusion
41. semen	91. distention	141. onset	191.invigorate(s/ing)
42. urinary	92. pericardium	142. dysfunction	192. itching
43. vomiting	93. headaches	143. tong	193. needled
44. dizziness	94. tinnitus	144. chai	194. anterior
45. gallbladder	95. expel(s)	145. gen	195. parameters
46. wan	96. diarrh(o)ea	146. inflammation	196. dull
47. diabetes	97. yuan	147. massage	197. respiratory
48. needle(s)	98. acupressure	148. irritability	198. praeparata
49. dang	99. scanty	149. bowel(s)	199. toxicity
50. abdomen	100. slippery	150. ascending	200. heterogeneity

201. profuse	218. reticulatae	235. renal	252. moutan
202.complementary	219. toxic	236. cardiac	253. upright
203. toxin(s)	220. cassiae	237. corporeal	254. oblique
204. superficial	221. ophiopogonis	238. incidence	255. fibrosis
205. medication	222. miscarriage(s)	239. disperse(s)	256. japonici
206. obstructing	223. midline	240. subdues	257. pinelliae
207. costal	224. tonification	241. gong	258. intestinal
208. radicis	225. perpendicular	242.harmonize(s/ing)	259. lactiflorae
209. deviation	226. stagnant	243. intake	260. obstructed
210. aetiology	227. diaphragm	244. allergic	261.miltiorrhizae
211. epigastrium	228. discomfort	245. oppositae	262. impairment
212. spinal	229. poria	246. auricular	263. wrist
213. artery(ies)	230. atrophy	247.gastrointestinal	
214. tumor(s)	231.acupuncturist(s)	248. glutinosae	
215. caulis	232. dampness	249. proliferation	
216. beneficial	233. dioscoreae	250. spontaneous	
217. facial	234. medial	251. fen	

Appendix C

Sub-list 1 of the TCM Word List at Level 3

1. qi	18. sanyinjiao	35. sha	52. lieque
2. yang	19. zusanli	36. gou	53. qihai
3. zi	20. xiao	37. luo	54. zhishi
4. shen	21. xie	38. xiong	55. cang
5. zhi	22. xia	39. xuan	56. baihui
6. cao	23. zhong	40. neiguan	57. yanglingquan
7. pi	24. zao	41. taixi	58. gua
8. jiao	25. qin	42. po	59. hou
9. mu	26. huo	43. quchi	60. dou
10. shi	27. hegu	44. niu	61. qiao
11. sheng	28. shenshu	45. pishu	62. xian
12. jing	29. jie	46. zhen	63. zhaohai
13. ji	30. zheng	47. xue	64. fengchi
14. shu	31. xing	48. xin	65. zangfu
15. cun	32. taichong	49. rou	66. fenglong
16. xiang	33. guanyuan	50. zang	67. feishu
17. lian	34. zhongwan	51. yinlingquan	68. zhongji

APPENDIX D

Sub-list 2 of the TCM Word List at Level 1

1. detected	51. soreness	101. cycles	151. reflected
2. survival	52. detection	102. electrical	152. stored
3. difficulties	53. locate	103. nan	153. behaviour
4. independent	54. tension	104. contribute(s)	154. elevated
5. laboratory	55. occur(red/ring)	105. coupled	155. manner
6. techniques	56. formed	106. similarity	156. peak
7. exposure	57. dietary	107. labour	157. resulted
8. ancient	58. functioning	108. manipulation	158. sufficient
9. allocation	59. movements	109. aim(s/ed)	159. blocked
10. childhood	60. environment	110. controlling	160. depressive
11. flows	61. extremities	111. recruited	161. examinations
12. modification	62. sessions	112. boost(s)	162. substantial
13. passage(s)	63. referred	113.eliminating	163. ingredient
14. requires	64. supplements	114. extraction	164. strengthening
15. variety	65.derive(d/ing/s)	115. progression	165. cure(d)
16. attacks	66. fall(s/ing)	116. tube(s)	166. percentage
17. properties	67. gathering	117. virus(es)	167. resolving
18. reactions	68. illnesses	118. exercises	168. disturbance(s)
19. characteristic	69. integrated	119. raw	169. angle
20. connect(s/ed)	70. minor	120. stable	170. gender
21. pregnant	71. simultaneously	121. accuracy	171.maintain(ed/ing)
22. normally	72. density	122. emission	172. composed
23.generate(d/ing)	73. concept(s)	123. experiences	173. relieve(ing/s/d)
24. component	74. fertility	124. involve(s/ing)	174. energetic
25. content(s)	75. capacity	125. molecular	175. present(s)
26. stimulate(s/d)	76. imbalance(s)	126. employed	176. approaches
27. limitation(s)	77. continuous	127. produce(s/ing)	177. prevent(ing/s)
28. increases	78. methodological	128. quantitative	178. counter
29. internally	79. portion	129. enhance(d)	179. concentrations
30. presentation	80. complicated	130. manual	180. cool(s/ing)
31. articles	81. infant(s)	131. inserted	181. distinguish
32. contain(s/ed)	82. intense	132. removes	182. markedly
33. tested	83. signal	133. assigned	183. strain
34. context	84. randomization	134. compare(ing)	184. aching
35. correct	85. substance	135. effectively	185. drains
36. unclear	86. achieve(d)	136. situations	186. principal
37. experiment(s)	87. enter	137. vital	187. invasions
38.manifest(s/ing/ed)	88. transform(ed/s)	138. thirsting	188. stimulate(ing)
39. resistance	89. positions	139. blind	189. constant
40. transportation	90. perspective(s)	140. explore(d)	190. healing
41. individuals	91. practices	141. linked	191. outer
42. solution	92. regions	142. persistent	192. remedies
43. variations	93. describe	143. blockage	193. obtain
44. depend(s/ing)	94. shock	144. curative	194. warmth
45. suffer(ed/s)	95. appearance	145. motor	195. damaged
46. adopted	96. develop(ing/s)	146. transport	196. exhibited
47. stiff	97. stronger	147. calming	197. fundamental
48. sinking	98. compound	148. collapse	198. strain(ing/ed)
49. watery	99. elderly	149. injure(s)	199. bacterial
50. participant	100. binding	150. produces	200. evil(s)

201. participate(ing)	246. influenced	291. paired	336. processed
202. restless	247. separation	292. radiating	337. steaming
203. variance	248. bacteria	293. screening	338. frontal
204.accumulate(d/s/ing)	249. presents	294. subgroup	339. isolated
205. mortality	250. involvement	295. respective	340. partially
206. reverse	251. profiles	296. sighing	341. symptomatic
207. regulatory	252. separated	297. damaging	342. taxation
208. emphasis	253. statistics	298. differ(ed)	343. emptiness
209. injected	254. chemicals	299. distributed	344. participation
210. adjusted	255. reaches	300. oils	345. bluish
211. disturbed	256. stones	301. rebellion	346. lowering
212. rapidly	257.accompanying	302. reinforce(d)	347. determination
213. constitutional	258. allocated	303. varied	348. discrimination
214. theories	259. equivalent	304. benefiting	349. guiding
215. validity	260. warms	305. lifting	350. relaxation
216. decline	261. balanced	306. disability	351. cools
217. blind	262. representative	307. drying	352. selective
218. correlated	263. rises	308. flavors	353. vary(ies)
219. orally	264. scatter(s/ing)	309. mouse	354. combining
220. tightness	265. perceived	310. retained	355. instrument(s)
221. courses	266. successfully	311. curve	356. integrative
222. stopping	267. ya	312. hypothesis	357. partial
223. quicken(s)	268. attributed	313. radical	358. reactive
224. treasures	269. frustration	314. recruitment	359. acquisition
225. respondents	270. confusion	315. coldness	360. cultured
226.correspodence(s)	271. thickness	316. documented	361. habits
227. inadequate	272. expert	317. emerges	362. combines
228. occurrence	273. sided	318. processing	363. decreasing
229. standardized	274. distinct	319. confirm(s)	364. disruption
230. arise(s)	275. fertile	320. govern(s)	365. palms
231. simultaneous	276. mode	321. imaging	366. behavioral
232. tenderness	277. preference	322. incomplete	367. comparable
233. proportion	278. quantity	323. specificity	368. disc
234. flowing	279. cloudy	324. sympathetic	369. inappropriate
235. assist(ed)	280. depressed	325. circulate(s/ing)	370. parallel
236. removing	281. presenting	326. interpreted	371. situated
237. fatty	282. psychiatric	327. pretreatment	372. transferred
238. focuses	283. toes	328. classics	373. withdrawal
239. helpful	284. generalized	329. mechanical	374. consultation
240. unblock	285. weaken(s)	330. observe	375. evident
241. hardness	286. definition	331. formulation(s)	376. indices
242. integration	287.prescription(s)	332. heavenly	377. infectious
243. collecting	288. sores	333. indicator(s)	378. measuring
244. excellent	289. consequence	334. stained	379.phenomenon(a)
245. marker(s)	290. implies	335. delayed	380. spreads

APPENDIX E

Sub-list 2 of the TCM Word List at Level 2

1. salviae	48. ulcer(s)	98. antibody(ies)	151. nocturnal
2. wen	49. corni	99. ischemia	152. eczema
3. tan	50. scutellariae	100.itching	153. seminal
4. zingiberis	51.electroacupuncture	101. wallichii	154. carthami
5. importance	52. lycii	102. choppy	155. abscess
6. induce(s)	53. palpation	103.inhibit(ed/ing/s)	156. moisten(s/ing)
7. membranacei	54. pungent	104. analgesic	157. p53
8. abnormal	55. incontinence	105. nourishment	158. premature
9. cardiovascular	56. metabolic	106. phellodendri	159. penetrate(s)
10. patent	57. exiting	107. saliva	160. scrophulariae
11. tract	58. manic	108. umbilicus	161. ziziphi
12. receptor(s)	59. rhinitis	109. cum	162. polygoni
13. impaired	60. cytokine(s)	110. insufficiency	163. coicis
14. alismatis	61. bile	111. kit	164. predominant
15. vertigo	62. prolapse	112. overwork	165. biochemical
16. inhibition	63. jaundice	113. infarction	166. matrix
17. membrane(s)	64. obesity	114. incubated	167. ding
18. vaginal	65. ischemic	115. prostate	168. extinguish(es)
19. digestive	66. ligustici	116. antioxidant	169. gastric
20. enuresis	67. median	117. dispersing	170. recurrence
21. cholesterol	68. schisandrae	118. fibromyalgia	171. codonopsis
22. concomitant	69. exogenous	119. atractylodes	172. qualitative
23. prevalence	70. recurrent	120. cornu	173. oxidative
24.activate(d/s/ing)	71. caution	121. psoriasis	174. recens
25. tendon(s)	72. neurons	122. embryo(s)	175. transverse
26. trichosanthis	73. angina	123. gardeniae	176. empirical
27. dementia	74. myocardial	124. plasma	177. prone
28. lipid(s)	75. pruni	125. propensity	178. wristbands
29. sour	76. mania	126. spasm(s)	179. tremor(s)
30. cognitive	77. pharmacological	127. residual	180. aggravate(d)
31. sputum	78. amenorrhoea	128. fetus	181. anus
32. arthritis	79. fasting	129. diameter	182. diagnose(s/d)
33. hui	80. genital(s)	130. accordance	183. magnoliae
34. paralysis	81. hepatitis	131. hormone(s)	184. analgesia
35. coronary	82. pivot	132. malaria	185. neurological
35. rhei	83. subjective	133. secretion(s)	186. platelet
36. saline	84. ternatae	134. forhead	187. sensory
36. curcumae	85. motility	138. intercostal	188. grief
37. vascular	86. antibiotics	139. proximal	189. systemic
37. defecation	87. granule(s)	140. threshold(s)	190. abnormalities
38. classified	88. obstruct(s)	141.anti-inflammatory	191. endometriosis
39. peripheral	89. prolonged	142. gall	192. trauma
40. rehmannia	90. hao	143. kirlowii	193. jasminoidis
41. expel(ling/led/s)	91. applicable	144. excretion	194. caspase
42. pulmonary	92. coptidis	145.inferior	195. implantation
43. cavity(ies)	93. invading	146. gypsum	196. buffer
44. collaterals	94. lesion(s)	147. physiology	197. cellular
45. posterior	95. divergent	148. tonic(s)	198. colon
46. subdue(s)	96. viral	149. pilosulae	199. enrich(es)
47. turbidity	97. alleviate(s/d)	150. ankle	200. inhibitory

201. rheumatoid	250. latent	299. systolic	348. emaciation
202. concealment	251. Bcl2	300. colic	349. engender(s)
203. endothelial	252. perimenstrual	301. dribbling	350. thermal
204. ningpoensis	253. demographic	302. inhibitor(s)	351. adjacent
205. pediatric	254. fibroids	303. coptis	352. atherosclerosis
206. preproductive	255. interval(s)	304. distending	353. bulbus
207. hippocampus	256. knotted	305. invigorating	354. latency
208. hypochondrium	257. luteal	306. scapula	355. platycoid
209. osteoarthritis	258. pituitary	307. jobi	356. vaccine
210. persicae	259. tinctorii	308. precursor(s)	357. capsule(s)
211. pus	260. astragalus	309. coma	358. endometrial
212. regurgitation	261. descent	310. irritable	359. gynaecological
213. stabbing	262. ode	311. laser	360. mast
214. haemorrhoids	263. ovary(ies)	312. migraine	361. oriental
215. occipital	264. rigidity	313. pancreas	362. puerariae
216. overflowing	265. tenuifoliae	314. pensiveness	363.antihypertensive
217. synthesis	266. tingling	315. pores	364. multiflori
218. focal	267. diminished	316. aggregation	365. obese
219. gums	268. immaturus	317. girdle	366. delirium
220. influenza	269. thready	318. intermittent	367. diffuse(ing)
221. jujubae	270. transcription	319. rehabilitation	368. leucorrhoea
222. progestrone	271. unilateral	320. sciatica	369. mediators
223. vivo	272. convulsions	321. consolidate	370. necrosis
224. axilla	273. thigh	322. lachryma	371. nephropathy
225. biomedical	274. Ccl4	323. pelvic	372. polygalae
226. carmichaeli	275. copious	324. vertebra€	373. scrofula
227. sprain	276. decoct(ed)	325. canthus	374. surgical
228. grind	277. eligible	326. cyperi	375. ting
229. pancreatic	278. eucommiae	327. hypogastric	376. ulmoidis
230. ping	279. minimal	328. hypoglycemic	377. amplitude
231. mellitus	280. contraindicated	329. mmHg	378. crude
232. RNA	281. harmony	330. opioid	379. exertion
233. sneezing	282. hyperactivity	331. tablet(s)	380. fragrant
234. achyranthis	283. invade(s/d)	332. collagen	381. horizontal
235. coefficient	284. prostatitis	333. plantaginis	382. japonicae
236. originate(s)	285. cervi	334. rotation	383. puncture
237. undigested	286. postoperative	335. soggy	384. cleft
238. amomi	287. perparata	336. terminal	385. epithelial
239. carcinoma	288. sedate	337. ventricular	386. fibrosum
240. hypertensive	289. enzyme(s)	338. anatomical	387. fluoxetine
241. moist	290. induction	339. aspheloidis	388. hippocampal
242. neuropathy	291. lassitude	340. gel	389. notoginseng
243. vulgaris	292. mitochondrial	341. radial	390. placenta
244. arterial	293. spinosae	342. modalities	391. simulation
245. armeniacae	294. cirrhosis	343. muscular	392. tolerance
246. bidentatae	295. depletion	344. ostreae	393. foul
247. blot	296.differentiate(d/ing)	345. pneumonia	394. haemorrhage
248. carotid	297. endogenous	346. upregulated	395. pernicious
249. cyrus	298. rash(es)	347. autonomic	

APPENDIX F

Sub-list 2 of the TCM Word List at Level 3

1. geng	18. ququan	35. zhigou	52. qiang
2. bei	19. geshu	36. zhe	53. zulinqi
3. xuehai	20. xinshu	37. qigong	54. lue
4. waiguan	21. suo	38. zha	55. jianyu
5. mingmen	22. weizhong	39. tianshu	56. shenting
6. deqi	23. gongsun	40. jue	57. tongli
7. bian	24. xingjian	41. duan	58. huantiao
8. shenmen	25. ganshu	42. jiu	59. pangguangshu
9. yue	26. sanjiaoshu	43. taiyang	60. shuidao
10. shou	27. taichi	44. zong	61. neiting
11. sanjiao	28. shenmai	45. qimen	62. huan
12. shuifen	29. weishu	46. taiyuan	63. jiuwei
13. suan	30. fengfu	47. pian	64. tiao
14. dazhui	31. kunlun	48. renzhong	65. danggui
15. lun	32. shanzhong	49. taibai	66. shangjuxu
16. fuliu	33. rong	50. yangming	67. rangu
17. houxi	34. daling	51. chize	

APPENDIX G

Sub-list 3 of the TCM Word List at Level 1

1. constitute	26. sciences	51. generates	76. immunity
2. redder	27. varying	52. tidal	77. positively
3. urgency	28. climatic	53. virtue	78. preventative
4. coping	29. dreaming	54. externally	79. altered
5. molecules	30. ineffective	55. fats	80. construct
6. concentrated	31. interiorly	56. radiates	81. energizer
7. connective	32.recommendations	57. heating	82. predictive
8. diseased	33. separates	58. progressive	83. urgent
9. extending	34. bodily	59. theoretical	84. advantages
10. lumps	35. constituents	60. untreated	85. arising
11. professionals	36. derive	61. antidepressant	86. completion
12. balancing	37. determining	62. evidenced	87. dominant
13. cupping	38. relies	63. gentle	88. enhancing
14. posttreatment	39. travels	64. indicative	89. episodes
15. translate(d)	40. applying	65. lacking	90. implications
16. nonspecific	41. binds	66. removal	91. radicals
17. probability	42. gradual	67. supportive	92. shivering
18. runny	43. reddish	68. swallowing	93. intensive
19. assessments	44. structural	69. complaint	94. participating
20. facilitate	45. conducting	70. formulated	95. poststroke
21. monitored	46. inspection	71. intention	96. prospective
22. mutually	47. transmitted	72. consumed	97. selecting
23. phenomenon	48. variability	73. developmental	98. correlated
24. settles	49. composition	74. continuously	99. undergoing
25. rabbit	50. extending	75. clinicians	100. dries
101. migration	120. screened	139. contracted	158. accurately
102. relaxes	121. applies	140. extremity	159. expenditure
103. rooted	122. output	141. implemented	160. liquids
104. underwent	123. rebelling	142. improper	161. referral
105. comprised	124. separating	143. interfere	162. segment
106. fainting	125. bleed	144. standardization	163. clustering
107. grouping	126. excessively	145. encountered	164. connectivity
108. gentlemen	127. resistant	146. gathers	165. essentials
109. institutional	128. absorption	147. localized	166. rootless
110. interact	129. colds	148. reproduction	167. yielded
111. anticancer	130. demonstrates	149. resides	168. emphasised
112. barriers	131. exception	150. sampling	169. institutes
113. brightens	132. restricted	151. uncontrolled	170. interference
114. fibers	133. grouped	152. worsening	171. normalized
115. immature	134. mindfulness	153. clarify	172. plots
116. injuring	135. outward	154. correction	173. powered
117. maturation	136. smelling	155. dislike	
118. references	137. associations	156. drainage	
119. salty	138. ceaseless	157. piglet	

APPENDIX H

Sub-list 3 of the TCM Word List at Level 2

1. ultrasound	51. rotundi	101. lethargy	151. aconite
2. vein	52. slimy	102. magnolia	152. aphasia
3. vitality	53. validated	103. polydipsia	153. codonopsitis
4. zedoary	54. blazing	104. amelioration	154. colitis
5. borborygmus(i)	55. cortical	105. endocrine	155. fracture
6. grandiflori	56. cystitis	106. fallopian	156. incubation
7. indirectly	57. dispel(s)	107. forsythiae	157. pectoris
8. lateralis	58. ephedra	108. metabolites	158. rectify
9. linear	59. fossa	109. penetration	159. spinous
10. cervix	60. frutescentis	110. seizures	160. stimuli
11. formalin	61. macrophages	111. static	161. acoustic
12. ginsenoside	62. outpatient	112. zizyphi	162. autophagy
13. polygonati	63. panacis	113. adjunctive	163. chrysanthemi
14. rhodiola	64. vitamin	114. algorithm	164. dysenteric
15. cones	65. diastolic	115. apoptotic	165. expulsion
16. dipsaci	66. guan	116. bowels	166.interrogation
17. kinase	67. hypothalamus	117. cannabis	167. lonicerae
18. orientalis	68. rectum	118. conquitae	168. regimen
19. perillae	69. appendix	119. fetal	169. turmoil
20. refined	70. calcium	120. flush	170. basel
21. sigma	71. distress	121. junction	171. cyathulae
22. cuscutae	72. exteriorly	122. menopause	172. differential
23. ejaculation	73. inducing	123.pneumothorax	173. ephedrae
24. hiccup(s)	74. lacrimation	124. sclerosis	174. gastritis
25. intellect	75. lucidi	125. thoracic	175. identical
26. reperfusion	76. potent	126. bland	176. murky
27. expectoration	77. salvia	127. bronchitis	177. proteinuria
28. itchy	78. somnolence	128. flaccidity	178. susceptible
29. nodules	79. albus	129. flaring	179. visceral
30. pertain(s/ing)	80. ethanol	130. goitre	180. dependence
31. respiration	81. follicle(s)	131. malignant	181. dissolved
32. sallow	82. phellodendron	132. resentment	182. draconis
33. temporal	83. pinellia	133. umbilical	183. enlarged
34. urticaria	84. prostatic	134. vagina	184. ingestion
35. alpha	85. rib	135. airway	185. nodes
36. Bax	86. scraping	136. cramps	186. relapse
37. conjunction	87. bupleurum	137. dysentery	187. sepsis
38. dyspnea	88. dioscorea	138. ion	188. benign
39. erection	89. hypogastrium	139. licorice	189. diuresis
40. etiology	90. rubrus	140. mucosa	190. dopamine
41. graminei	91. stagnate(s)	141. neural	191. rectal
42. malleolus	92. aggravate	142. rubra	192. traumatic
43. pathophysiology	93. anemarrhena	143.transplantation	193. urethra
44. spatholobi	94. clots	144. vertex	194. artemisiae
45. congenital	95. downregulated	145. alternation	195. cohort
46. corydalis	96. gentiana	146.cerebrovascular	196. follicular
47. dilution	97. optimal	147. fortify(ies)	197. funnel
48. gland(s)	98. alternating	148.proinflammatory	198. hyperplasia
49. holistic	99. disinhibit	149. stifling	199. myopia
50. retinopathy	100. hypochondriac	150. validation	200. phosphate

201. sui	256. flabby	311. extracellular	366. cumulative
202. ulceration	257. indirect	312. genitalia	367. infract
203. viability	258. morifolii	313. histone	368. inherent
204. absent	259. stubborn	314. littoralis	369. ligustri
205. actin	260. treatise	315. measles	370. lipoprotein
206.neuroprotective	261. autoimmune	316. nitrogen	371. metastasis
207. peritoneal	262. elucidated	317. nocturia	372. microarray
208. saussureae	263. glehniae	318. psoriatic	373. morindae
209. skeletal	264. hemoglobin	319. simulated	374. multivariate
210. subjected	265. histamine	320. suppressed(ed)	375.pharmacotherapy
211. acrid	266. noninvasive	321. afferent	376. podocyte
212. cholinergic	267. occiput	322. biomedicine	377. polarity
213. eruptions	268. palpable	323. fibroblasts	378. pubic
214. menses	269. prenatal	324. intraperitoneal	379. utilization
215. perineum	270. quantified	325. modulation	380. X-ray
216. polyuria	271. ascent	326. obstructive	381. asari
217. rigorous	272. astringe	327. oxalate	382. clonic
218. sequelae	273. corpus	328. regent(s)	383. comparative
219. thyriod	274. distilled	329. vexation	384. erectile
220. compliance	275. impair(s)	330. follicle(s)	385. glycyrrhiza
221. contrary	276. ovarian	331. harmonious	386. microscope
222. crataegi	277. Pon2	332. morbidity	387. muzziness
223. dermatomes	278. sacrum	333. morphine	388. nucleus
224. frankincense	279. subcutaneous	334. resonance	389. pacify
225. malar	280. throbbing	335. retrieved	390. usage
226. pancreatitis	281. tuberculosis	336. sacral	391. antiarrhythmic
227. postpartum	282. clonic	337. uncariae	392. antibiotic
228. sandalwood	283. dysmenorrhoea	338. verified	393. Ca1
229. artemisia	284. eligibility	339. adherence	394. cytosolic
230. ascites	285. ginkgo	340. attenuated	395. electro
231. bambusae	286. inhalation	341. cartilage	396. ligand
232. centrifuged	287. lignum	342. crease	397. massa
233. cocaine	288. nelumbinis	343. degeneration	398. plantago
234. congealed	289. opisthotonos	344. ectopic	399. pollen
235. fritillariae	290. overweight	345. harmonized	400. purging
236. guinea	291. sterile	346. impair(s)	401. soles
237. insulting	292. viscera	347. infantile	402. soothe
238. mitochondria	293. adjuvant	348. multiflorum	403. transient
239. p38	294. affinity	349. pediatrics	404. creatinine
240. testicles	295. ancestral	350. prominence	405. dorsal
241. augument(ed)	296.chromatography	351.pseudostellariae	406. macrocephala
242. denotes	297. dysmenorrhea	352. sinusitis	407. oxide
243. duct	298. elicited	353. anesthesia	408. sedation
244. gentianae	299. interleukin	354. cuscatae	409. spatial
245. ledebouriellae	300. location	355. exert	410. twitching
246. logistic	301. lobe	356. interstitial	411. amino
247. menthae	302. metformin	357. popliteal	412. appendicitis
248. phosphorylation	303. mucous	358. purulent	413. bloating
249. prevalent	304. otitis	359. serotonin	414. bouts
250. suspensae	305. polymerase	360. suppression	415. consolidation
251. aetiological	306. postnatal	361. susceptibility	416. gradient
252. anemia	307. anatomy	362. testudinis	417. immunoistochemistry
253. anesthetized	308. carbohydrates	363. apparatus	418. intercourse
254. bilaterally	309. cimicifugae	364. asthmatic	419. kaempferol
255. colorectal	310. electrophoresis	365. constriction	420. notopterygii

421. nutrients	423. sesloidis	425. stuffiness	
422. parching	424. sphere	426. supraclavicular	

APPENDIX I

Sub-list 3 of the TCM Word List at Level 3

1. fuling	22. chengshan	43. touwei	64. lidui
2. shanghuo	23. ciliao	44. weiyang	65. qiwei
3. zhangmen	24. laogong	45. yanhusuo	66. renying
4. zhongfu	25. qiuxu	46. yinbai	67. shenque
5. liang	26. shousanli	47. tianzhu	68. shenzhu
6. mang	27. yingu	48. xiaxi	69. jinmen
7. fengmen	28. dadun	49. yintang	70. shangwan
8. chuanxiong	29. jianjing	50. changqiang	71. tui
9. jiexi	30. yangxi	51. tou	72. yingxiang
10. dachangshu	31. guo	52. siman	73. daheng
11. yongquan	32. huai	53. tianjing	74. dashu
12. shangxing	33. shaoyang	54. miao	75. lao
13. wangu	34. tiantu	55. shimen	76. yangfu
14. dao	35. guizhi	56. zhengs	77. shaoze
15. guasha	36. jingming	57. guanchong	78. ao
16. ligou	37. bufei	58. shangyang	79. tuina
17. qichong	38. xiajuxu	59. shaoshang	80. ahshi
18. xuanzhong	39. yanggu	60. feiyang	81. jueyinshu
19. zhongzhu	40. jianshi	61. sanjian	82. shaohai
20. shangqiu	41. zuo	62. fengshi	83. yamen
21. juque	42. dazhong	63. huiyin	

APPENDIX J

The Common Core TCM List (arranged by alphabet)

Note. This list includes shared headwords between Hsu's (2018) and the present study.

1	abdomen	2	abnormal	3	abscess
4	absent	5	acid	6	acrid
7	activate	8	acupuncture	9	acute
10	adverse	11	affinity	12	aggravate
13	alleviate	14	amenorrhoea	15	analgesia
16	anatomy	17	angina	18	ankle
19	anterior	20	antibiotic	21	anus
22	appendix	23	appetite	24	artery
25	arthritis	26	ascend	27	ascend
28	asthma	29	atrophy	30	augment
31	beneficial	32	bilateral	33	bile
34	bladder	35	bland	36	bowel
37	bronchitis	38	calcium	39	cardiac
40	cartilage	41	caution	42	cavity
43	cerebral	44	cervix	45	chronic
46	cleft	47	colic	48	coma
49	comparative	50	complexion	51	concomitant
52	congenital	53	conjunction	54	consolidate
55	contraction	56	cord	57	coronary
58	cortex	59	decoct	60	deficiency
61	delirium	62	dementia	63	diabetes
64	diagnosis	65	diameter	66	diaphragm
67	diarrhoea	68	differential	69	differentiate
70	diffuse	71	discomfort	72	disharmony
73	dispel	74	disperse	75	distal
76	distress	77	diuresis	78	dorsal
79	dorsum	80	duct	81	dull
82	duration	83	dynamic	84	dysentery
85	dysfunction	86	dyspnea	87	eczema
88	efficacy	89	elbow	90	emperor
91	endocrine	92	endogenous	93	enuresis
94	enzyme	95	epigastric	96	epilepsy
97	epistaxis	98	essence	99	ethanol
100	etiology	101	exert	102	exogenous
103	expel	104	exterior	105	extinguish
106	facial	107	fang	108	fatigue
109	febrile	110	fetal	111	fever
112	fluid	113	flush	114	fossa
115	foul	116	fracture	117	gallbladder
118	gastrointestinal	119	gel	120	genital
121	ginkgo	122	ginseng	123	gland
124	glucose	125	grind	126	haemorrhage
127	hemiplegia	128	hepatic	129	hepatitis
130	herb	131	hiccup	132	holistic
133	horizontal	134	hormone	135	hypertension
136	impair	137	impotence	138	induce
139	infantile	140	inferior	141	influenza
142	inhibit	143	insomnia	144	intake

145	intermittent	146	interval	147	intestine
148	invade	149	invigorate	150	jaundice
151	junction	152	kidney	153	lassitude
154	lateral	155	lesion	156	limb
157	liver	158	lobe	159	lumbar
160	lung	161	malaria	162	malleolus
163	marrow	164	massage	165	medial
166	medication	167	membrane	168	menopause
169	meridian	170	metabolism	171	metastasis
172	midline	173	migraine	174	minimal
175	moist	176	mucus	177	muscular
178	nasal	179	nausea	180	necrosis
181	needle	182	nourish	183	obese
184	oblique	185	obstruct	186	occipital
187	oedema	188	onset	189	optimal
190	orifice	191	originate	192	ovary
193	pancreas	194	paralysis	195	pathogenic
196	pathology	197	penetrate	198	pericardium
199	perpendicular	200	pertain	201	pharmaceutical
202	phlegm	203	physician	204	physiological
205	pill	206	placebo	207	pneumonia
208	posterior	209	potent	210	predominant
211	premature	212	prenatal	213	prescribe
214	prescription	215	prevalent	216	prognosis
217	prolapse	218	prone	219	prostate
220	proximal	221	pulmonary	222	pulse
223	puncture	224	pungent	225	purulent
226	pus	227	qualitative	228	radial
229	rash	230	relapse	231	renal
232	reproductive	233	respiration	234	retention
235	rheumatoid	236	rhinitis	237	rib
238	sacral	239	saliva	240	salvia
241	sclerosis	242	scrofula	243	sedate
244	sensation	245	sensory	246	serum
247	sinew	248	slippery	249	soothe
250	sour	251	spasm	252	spine
253	spleen	254	spontaneous	255	sprain
256	sputum	257	stagnate	258	stasis
259	static	260	sterile	261	sticky
262	stool	263	subcutaneous	264	subjective
265	superficial	266	suppress	267	surgical
268	susceptible	269	syndrome	270	synthesis
271	systemic	272	tablet	273	temporal
274	tendon	275	therapeutic	276	thigh
277	thoracic	278	threshold	279	tinnitus
280	tolerance	281	tong	282	tonic
283	toxic	284	toxin	285	tract
286	transverse	287	trauma	288	tuberculosis
289	turbid	290	ulcer	291	urethra
292	urine	293	urticaria	294	uterine
295	uterus	296	vagina	297	vascular
298	vein	299	vertebra	300	vertex
301	vertigo	302	visceral	303	vitality
304	vitamin	305	vitro	306	wan
307	wen	308	wrist	309	yin