High Frequency Medical Spanish Terminology: A Corpus-Based Study of Textbooks and Reference Materials

Alyssia Miller De Rutté
Colorado State University

Follow this and additional works at: https://gbl.digital.library.gwu.edu/

Recommended Citation

Alyssia Miller De Rutté
Colorado State University

High Frequency Medical Spanish Terminology:
A Corpus-based Study of Textbooks and Reference Materials

Abstract: The purpose of this study is to identify high frequency medical Spanish terminology by building a corpus of medical Spanish textbooks and resources. By discovering the most common terms, instructors and curriculum designers can include salient topics in courses, which will better serve learners in these courses. To do so, medical Spanish textbooks and resources from 2012-2022 were uploaded to the corpus software AntConc. The final corpus consisted in a total of 26 resources and 1,421,163 tokens (or words). Word frequencies were calculated, and the top 3,000 most frequent terms were determined. A Semantic Rating Scale (Chung & Nation, 2003; Quero & Coxhead, 2018) was used to categorize words as general or technical. The top 3,000 words covered 38.9% of the entire corpus, and the most frequent words found in the medical Spanish resources were general Spanish vocabulary at 68.2%. The remaining 31.8% were medical Spanish technical terminology. The generation of these lists gives instructors and curriculum designers a tool to use in the development of medical Spanish courses that when combined with a patient-centered, communicative medical Spanish curriculum has the potential to enhance learners' medical Spanish proficiency.

Keywords: corpus research, lexicon, medical Spanish, technical vocabulary, Spanish for Specific Purposes

Introduction

Spanish for Specific Purposes (SSP) is a growing field of study. Particularly, one of its subfields, medical Spanish, has seen an increase in research and teaching in the last few years (Sánchez-López et al., 2017). However, there is still no consensus or standardization on medical Spanish education at the medical school level (Morales et al., 2015) or the undergraduate university level (Miller De Rutté et al., 2023). This is further exemplified by medical Spanish instructors reporting difficulty in finding resources to teach these classes (Miller De Rutté et al., 2023; Ortega et al., 2020). There have been recommendations from an expert panel on recommended core competencies and performance objectives for medical Spanish at the medical school level (Ortega et al., 2020). These recommendations include medical Spanish knowledge regarding organ systems, medical interviewing, and disease entities; patient-centered explanation of medical diagnoses/assessment and treatment/evaluation plan; and self-assessment of confidence and limitations. However, many of these recommendations may not be appropriate for undergraduate, pre-health, pre-clinical students, who are enrolled in Spanish courses and not medical ones. There remains a gap in undergraduate medical Spanish, and in medical Spanish generally, when deciding what content and vocabulary to include in these courses (Miller De Rutté et al., 2023). As such, this study seeks to add to the field by systematically discovering the most frequent vocabulary terms used in medical Spanish textbooks and resources to give learners, instructors, and curriculum designers evidence to use in their own learning and course design.

Literature Review

Much previous research on vocabulary usage in specific contexts has occurred within the field of English for Specific Purposes (ESP) (see Nesi, 2012). In SSP there is less research on vocabulary usage in specific contexts. In medical Spanish, to the author's knowledge, only one study exists on vocabulary usage in medical Spanish. This study focused on developing a method to identify and classify Spanish medical jargon found in transcripts of recordings made by medical students while they were explaining diagnoses, treatment, and follow-up care during medical Spanish clinical scenarios (Ortega et al., 2023). The authors discussed how Spanish medical jargon has the potential to be added "to the assessment of Spanish-language patient-centered communication" (Ortega et al., 2023, p. 1).

Corpus-based research is common in ESP, and research has shown that "it is important for teachers and learners to find out what vocabulary is encountered in different contexts, what company it keeps, and what might be the most efficient way of teaching and learning it" (Coxhead, 2012, p. 1). Additionally, word lists are of importance as they can provide the vocabulary that is needed to enhance proficiency in different content areas (Coxhead & Nation, 2001). Research has found that specialized word lists provide a means to determine technical vocabulary found in a text or content area and help in the understanding of the size of specialized vocabulary (Nation et al., 2016). Woodward-Kron (2008) notes:

> the specialist language of a discipline is intrinsic to students' learning of disciplinary knowledge; students need to show their understanding of concepts, phenomena, relations between phenomena, etc. by incorporating the specialist language and terminology of their discipline into their writing accurately. They also need to adopt the specialist language in order to make meaning and engage with disciplinary knowledge. (p. 246)

Many studies have focused on the analysis of written texts using corpora (Coxhead, 2019). An example of such a study is by Chung and Nation (2003) in which they researched technical vocabulary and found that one in three words in an anatomy textbook was classified as being a technical or discipline-specific word. The researchers pointed out that the instructors teaching courses on ESP usually do not have "specialist knowledge of the learners' technical areas" (Chung & Nation, 2003, p. 114), and so, instructors may not have a knowledge base from which to draw when determining vocabulary that should be included in an ESP course. Most times, teachers of these courses become researchers as well to determine what content to include in their courses (Hall, 2013). By systematically recognizing the most frequent vocabulary words, instructors can provide technical vocabulary instruction without being experts in the discipline-specific field. However, it is important to note that authors of textbook and resource materials may not be area experts, and most do not state how they determined the vocabulary (i.e., from a language needs analysis) to include in their resource, which is an inherent limitation.

Academic word lists in English have been created in many disciplines such as chemistry (Valipouri & Nassaji, 2013), computer science (Lam, 2001), engineering (Mudraya, 2006, Nekrasova-Beker, 2020), fabrication and welding (Coxhead et al., 2019), finance and accountancy (Ha & Hyland, 2017), nursing (Yang, 2015), and plumbing (Coxhead & Demecheleer, 2018). Coxhead and Demecheleer (2018) found that their technical word list for plumbing covered 32.17% of their corpus while Ha and Hyland (2017) found a coverage of 23.94% for their word list in finance and accountancy. These findings indicate that many words in these corpora were non-technical. In fact, Chung and Nation (2003) applied a semantic rating

scale to categorize words as general or technical. This scale has been used in subsequent studies (Quero & Coxhead, 2018), is used in the present study, and will be explained further. Quero and Coxhead (2018) conducted a corpus analysis of English textbooks used in medical schools to determine the most frequent terminology. They found that the first 3,000 words "constitute a vital set of vocabulary for ESP medical students at the beginning of their studies" (p. 51). Other studies have investigated word frequency in medical texts. There are few specialized word lists that have been developed that are relative to medicine. The Nursing Academic Word List (Yang, 2015) and the Medical Academic Word List (Wang et al., 2008) focus on establishing word lists related to research articles in nursing or medicine to aid English learners in reading and publishing articles in their respective fields. Hsu (2013) created the Medical Word List, which was derived from published eBooks. This corpus crossed numerous medical subject areas and excluded highly specialized Greek and Latin sources to develop a word list that encompassed "various sub-technical and lay-technical vocabularies" (Hsu, 2013, p. 454). These three word lists only occur in English, however. Therefore, this study aims to identify the most frequently used words in a discipline-specific context, in this case medical Spanish. This identification can provide both instructors of medical Spanish courses and students taking these courses with a starting point to learning and curriculum development as it pertains to medical Spanish vocabulary and content. This study was guided by the following research questions:

1. What are the 3,000 Spanish words with the most frequency found in a medical Spanish corpus compiled from textbooks and other resources?
2. What is the lexical coverage of the top 3,000 Spanish words?
3. How are the most frequent Spanish words categorized—as general words or as content—specific/technical terminology?

The decision to use the first 3,000 most frequent terms was chosen based on Quero and Coxhead's (2018) study as these terms yield a high lexical coverage, which defined as "the percentage of running words in the text known by the reader" (Nation, 2006, p. 61)

Methodology

Corpus

A corpus was compiled for this study that consisted of medical Spanish textbooks and resources published between January 2012 and July 2022. To identify resources for inclusion in the corpus, EBSCO, Google Books, Google Scholar, Web of Science, and World Cat were systematically searched. Shin et al. (2021) conducted a scoping review in which they used several search terms to find medical Spanish textbooks. Their list of search terms was modified for this study and included the following: "medical Spanish," "health Spanish," "health care Spanish," "healthcare Spanish," "Spanish for physicians," "Spanish for doctors," "Spanish for nurses," "Spanish for nursing," "español médico," "español para la salud," and "español para profesiones médicas." Figure 1 shows the search process. First, databases were searched using the indicated search terms. A total of 164 sources were identified. An additional four sources were identified from the Shin et al. (2021) study, which yielded a total of 168 sources. Duplicates were removed (n=123), and a remaining 45 records were screened. Of these 45, an additional 19 were removed as they were self-published, were not published within the last 10 years, or were audio only. Twenty-six textbooks and resources were included in the study.

Figure 1

The Search Process for Medical Spanish Textbook Selection
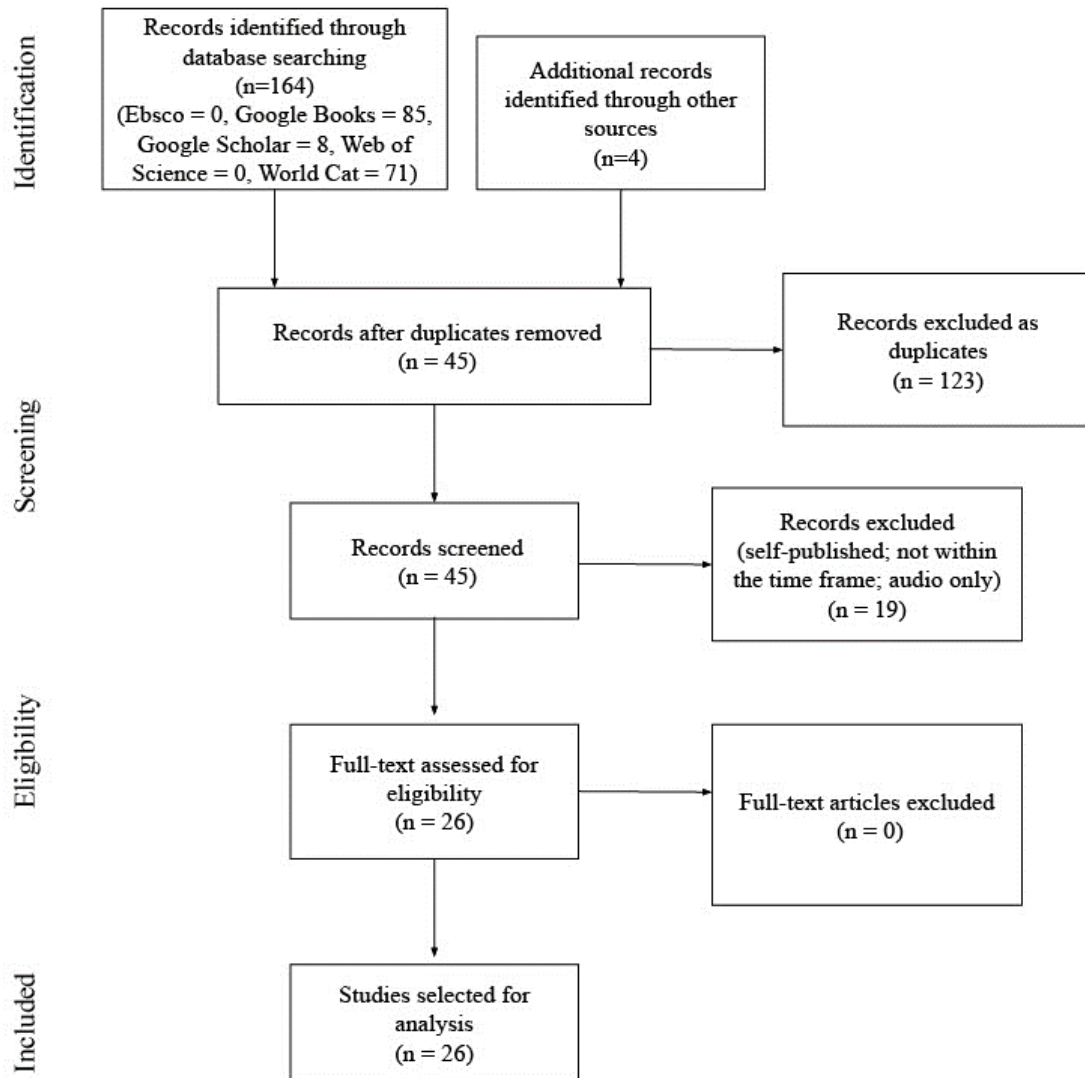
Digital versions, when available, were uploaded to AntConc, a free corpus analysis software (Anthony, 2022). When digital versions were not available, textbooks were scanned and run through Optical Character Recognition (OCR) software (Adobe, Inc., 2022) and then uploaded to AntConc. Once files were uploaded to AntConc, the corpus needed to be cleaned. Many of the resources included pronunciation of words, so individual letters (that did not equate to words) and sounds were removed from the corpus. Additionally, because the textbooks were written in both Spanish and English, there was a mix of both languages represented in the list. Moreover, the OCR software was only able to recognize one language at a time, and so, some words with written accent marks appeared in both forms, one with the accent and one without the accent. For example, the word *año* (year) and *ano* (anus) both appeared in the corpus as they are two relevant terms within a medical context. To ensure the meaning of these words, the Key Word in Context (KWIC) option was used to understand the definition of the word in question. Other words would also appear with and without written accent marks, such as *medico* and *médico* (doctor). These words and frequencies were combined under one term: *médico*. Cognates in both languages were also present, for example, *cultural* (in English) and *cultural* (in Spanish). Again, the KWIC option was used to determine if the word was being used in English or Spanish. Furthermore, word types were used as the counting mechanism. For example, *enfermedad* (sickness) and *enfermedades* (sicknesses) were counted as two separate tokens. This decision was made based on previous research (Chung & Nation, 2003; Quero & Coxhead, 2018).

Using the Semantic Rating Scale

As the purpose of this study was to understand vocabulary in Spanish, only the Spanish words were categorized as general terminology or content area/technical terminology using the semantic rating scale. As Quero and Coxhead (2018) state, "these potential specialized (medical) words need to be checked systematically to decide if they are truly specialized words" (p. 59). Following Quero and Coxhead (2018), the semantic rating scale was divided into two main levels, which included general purpose vocabulary and content area/technical vocabulary. General purpose vocabulary "refers to the words needed to write or talk about a wide range of topics, disciplines or content areas" and include words such as "*because, outside, ignore*" (Quero & Coxhead, 2018, p. 59). Content area/technical vocabulary "refers to words whose meaning is related to a particular topic, discipline, field or subject domain" (Quero & Coxhead, 2018, p. 60). Within the content area or technical vocabulary, there are different subfields that help in the categorization of terms as seen in Table 1. These subfields were used to categorize the content area/ technical vocabulary. As can be seen, there are different levels of technicality. These levels include words of high specialization as well as words that could be considered "general" because they are commonly used in everyday life (such as *doctor* or *health)*; however, because they are also used in the medical field, they are classified as content area/ technical vocabulary. If words fell outside of these subfields, they were considered to be general vocabulary. This way of coding terminology has been conducted and validated in previous studies (Chung & Nation, 2003; Quero & Coxhead, 2018). Therefore, all words generated by the corpus were compared to the four subfields to determine if they were general purpose vocabulary or content area/ technical vocabulary.

Table 1

Subfields of Content Area/Technical Vocabulary Related to the Medical Field (Quero & Coxhead, 2018)

Subfield 1. Words are also general words but are used in the medical field. Typically, these words maintain the same meaning they have in everyday use. Some examples are *doctor* or *health.*
Subfield 2. Words can also be general words used in the medical field. However, these words have a specialized meaning that is not as frequent in everyday use. Some examples are *pressure* or *history.*
Subfield 3. Words are associated with a specialized area and are often used to talk and write about health and medicine. Additionally, an expert in the specialized area would indicate that these words are part of their discipline-specific vocabulary. Some examples are *heart* or *skin.*
Subfield 4. Words are unique to the medical field and are associated with highly technical medical topics. Typically, these words are not found in other disciplines or in everyday use. Some examples are *embolism* or *conjunctivitis.*

Results

Overview

There was a total of 26 resources used to create the corpus, which had 1,421,163 tokens (i.e., words). The total tokens include both English and Spanish. Table 2 shows the 26 resources, the total number of tokens in each resource, and the percentage of coverage that each resource had in the corpus (see Appendix A for full bibliographic list of corpus sources.) To determine the percentage of coverage for each resource, the number of tokens in each resource was divided by the total number of tokens in the corpus (1,421,163).

Table 2

Number of Tokens and Percentage of the Whole by Textbook or Resource

Resource Author (Year)	Resource Name	Total number of tokens	Percentage of tokens compared to corpus
Block et al. (2013)	*El mundo hispano y la salud: Texto de español para nivel intermedio*	29,928	2.1%
Bobenhouse (2014)	*McGraw-Hill Education's Medical Spanish Visual Phrasebook*	18,583	1.3%
Bradley Williams (2015)	*Medical Spanish: A Pronto Reference and Study Guide*	8,604	0.6%
Chase & Medina de Chase (2019)	*An Introduction to Medical Spanish: Communication and Culture*	99,362	7.0%
Dejbord Sawan (2020)	*Beginning Medical Spanish: Oral Proficiency and Cultural Humility*	44,678	3.1%

Galarreta-Aima et al. (2021)	*Intermediate Medical Spanish: A Healthcare Workers' Guide for Communicating with the Latino Patient*	75,710	5.3%
Giralt (2012)	*Español médico y sociedad: Un libro para estudiantes de español en el tercer año de estudios*	81,847	5.8%
Guillet & Echavarría (2015)	*Spanish, Culture, and Health: An Introduction for Healthcare Professionals*	51,528	3.6%
Hardin & Irom (2018)	*Español onversacional para profesiones médicas*	49,600	3.5%
Harvey (2016)	*Spanish for Health Care Professionals*	82,145	5.8%
Jany & Mayberry (2017)	*Advanced healthcare Spanish language and culture*	46,135	3.2%
Jarvis & Lebredo (2014)	*Basic Spanish for Medical Personnel*	69,302	4.9%
Lingo Mastery (2019)	*Medical Spanish: Real Spanish Medical Conversations for Healthcare Professionals*	26,755	1.9%
Machtinger & Nigrović (2019)	*Spanish for Pediatric Medicine: A Practical Communication Guide*	73,951	5.2%
Morris García et al. (2019)	*Profesionales de la salud: Curso de español*	55,749	3.9%
Ortega (2016)	*Spanish and the Medical Interview: A Textbook for Clinically Relevant Medical Spanish*	133,485	9.4%
Ortega & Alemán (2022)	*Spanish and the Medical Interview: Clinical Cases and Exam Review*	31,787	2.2%
Rebar et al. (2019)	*Medical Spanish Made Incredibly Quick*	32,358	2.3%
Retter (2017)	*Medical Spanish: Fast Track Learning for English Speakers*	8,240	0.6%
Rios et al. (2021a)	*McGraw Hill's Complete Medical Spanish*	117,317	8.3%
Rios et al. (2021b)	*McGraw-Hill's Spanish for Healthcare Providers*	104,994	7.4%

Spark Notes (2014)	*Medical Spanish SparkCharts*	3,260	0.2%
Touri (2020)	*Learn Medical Spanish in 100 Days: Spanish Words and Phrases for Healthcare Professionals to Become Fluent Faster*	26,755	1.9%
Traverso et al. (2017)	*Medical Spanish for Nurses: A Self-teaching Guide*	68,516	4.8%
Valenzuela (2013)	*Spanish for Nurses*	30,353	2.1%
Vox (2012)	*Vox Super-mini Medical Spanish and English Dictionary*	50,221	3.5%

The resources used to construct the corpus varied in length (as seen by the number of tokens in Table 2) and in language. Most resources were written primarily in English and had Spanish translations, interviews, readings, etc. Six resources were primarily in Spanish (Block et al., 2013; Jany & Mayberry, 2017; Galarreta-Aima et al., 2021; Giralt, 2012; Hardin & Irom, 2018; Morris García et al., 2019). Eight resources were only vocabulary lists or dictionaries in both English and Spanish (Bobenhouse, 2014; Lingo Mastery, 2019; Machtinger & Nigrovic, 2019; Rebar et al., 2019; Retter, 2017; Spark Notes, 2014; Touri, 2020; Vox, 2012).

Top 3,000 Terms

The first research question was to determine the 3,000 words with the most frequency in the medical Spanish corpus, and to establish this list, frequency was calculated. The top 30 most frequent terms, their rank, frequency, and range across the 26 resources can be found in Table 3. The top 100 terms can be found in Appendix B, and all 3,000 Spanish terms can be found online (https://www.iris-database.org/details/Teeig-9dgCM).

The majority of the top 100 terms are ubiquitous in nature and are terms that students would learn in novice level courses, including prepositions and articles such as *de* (of), *la* (the), *el* (the), etc. These are followed by conjunctions (*que* – that), other prepositions (*a* – to), and commonly used verbs (*es* – is) with occasional technical words that one would expect to use in a medical context, such as *paciente* (patient), *doctor* (doctor), and *dolor* (pain).

Table 3
Top 30 Spanish Words

Rank	Spanish (English) word	Freq	Range	Gen v Tech
1	de (of)	29702	26	gen
2	la (the)	27470	26	gen
3	el (the)	22421	26	gen
4	a (to)	13972	26	gen
5	que (that)	13034	25	gen
6	en (in)	12933	26	gen
7	y (and)	12425	26	gen

8	los (the)	9358	26	gen
9	o (or)	8275	26	gen
10	las (the)	7318	26	gen
11	es (is)	7016	26	gen
12	un (a/an)	6755	26	gen
13	no (no)	6175	26	gen
14	para (for)	5898	26	gen
15	con (with)	5654	26	gen
16	una (a/an)	5625	26	gen
17	se (pronoun)	5218	25	gen
18	del (from the/of the)	5102	26	gen
19	su (his/hers/theirs/yours-formal)	4943	26	gen
20	por (for through)	4589	26	gen
21	al (to the)	3963	26	gen
22	me (me/to me)	2972	25	gen
23	tiene (he/she/you-formal has)	2910	26	gen
24	le (him/her/you-formal)	2903	25	gen
25	más (more)	2301	26	gen
26	paciente (patient)	2123	25	tech
27	doctor (doctor)	2011	26	tech
28	dolor (pain)	2001	26	tech
29	lo (him/it/you-formal)	1994	26	gen
30	son (they/you all-formal are)	1974	24	gen

Lexical Coverage

Mean frequency percent was calculated for the top 3,000 Spanish words to determine the amount of coverage. The frequencies of the 3,000 words were summed together, and the total was divided by the total number of tokens (1,421,163). The total frequency coverage of the top 3,000 words equaled 38.9% of the total corpus.

General Terminology vs. Content Area/Technical Terminology

Of the top 3,000 most frequent Spanish words found in the medical Spanish resources, 68.2% (n=2045) were general vocabulary. The remaining 31.8% (n=955) were content area or technical terminology. Tables 4 and 5 show the top 30 general terms and the top 30 content area/ technical terms along with the rank, frequency, and range of each token. The top 100 general and technical terms can be found in Appendix C and D.

The top 30 general Spanish words (listed in Table 4) consist of prepositions (*de* – of; *a* – to; *en* - in), articles (*la* – the; *el* – the; *las* – the), conjunctions (*que* – that; *y* – and; *o* – or), verbs (*es* – is; *tiene* – has; *ha* – has), pronouns (*se* – reflexive pronoun; *me* – me/to me; *le* – to him/her/you formal), and adverbs (*más* – more; *cuando* – when). The top 30 technical words (in

Table 5) are comprised of mostly nouns (*paciente* – patient; *doctor* – doctor; *dolor* – pain) and one verb (*duele* – hurts/hurt).

Table 4

Top 30 General Spanish Words

Rank	Spanish (English) word	Freq	Range	Gen v Tech
1	de (of)	29702	26	gen
2	la (the)	27470	26	gen
3	el (the)	22421	26	gen
4	a (to)	13972	26	gen
5	que (that)	13034	25	gen
6	en (in)	12933	26	gen
7	y (and)	12425	26	gen
8	los (the)	9358	26	gen
9	o (or)	8275	26	gen
10	las (the)	7318	26	gen
11	es (is)	7016	26	gen
12	un (a/an)	6755	26	gen
13	no (no)	6175	26	gen
14	para (for)	5898	26	gen
15	con (with)	5654	26	gen
16	una (a/an)	5625	26	gen
17	se (pronoun)	5218	25	gen
18	del (from the/of the)	5102	26	gen
19	su (his/hers/theirs/yours-formal)	4943	26	gen
20	por (for through)	4589	26	gen
21	al (to the)	3963	26	gen
22	me (me/to me)	2972	25	gen
23	tiene (he/she/you-formal has)	2910	26	gen
24	le (to him/her/you-formal)	2903	25	gen
25	más (more)	2301	26	gen
26	lo (him/it/you-formal)	1994	26	gen
27	son (they/you all-formal are)	1974	24	gen
28	ha (has)	1945	25	gen
29	cuando (when-adverb or conjunction)	1903	25	gen
30	si (if)	1846	26	gen

Table 5

Top 30 Technical Spanish Words

Rank	Spanish (English) word	Freq	Range	Gen v Tech
1	paciente (patient)	2123	25	tech
2	doctor (doctor)	2011	26	tech
3	dolor (pain)	2001	26	tech
4	médico (doctor)	1405	26	tech
5	salud (health)	1392	23	tech
6	medical (medical)	1108	24	tech
7	sangre (blood)	876	26	tech
8	diabetes (diabetes)	873	26	tech
11	medicina (medicine)	836	24	tech
12	enfermedad (sickness)	800	25	tech
13	ejercicio (exercise)	781	22	tech
14	hospital (hospital)	773	26	tech
15	síntomas (symptoms)	710	26	tech
16	enfermedades (sicknesses)	678	22	tech
17	corazón (heart)	645	21	tech
18	duele (he/she/you-formal hurts/hurt)	624	24	tech
19	doctora (doctor)	578	20	tech
20	presión (pressure)	568	26	tech
21	tratamiento (treatment)	562	21	tech
22	sistema (system)	553	21	tech
23	exámen (exam)	543	25	tech
24	enfermera (nurse)	543	22	tech
25	cabeza (head)	542	25	tech
26	cuerpo (body)	541	23	tech
27	piel (skin)	538	25	tech
28	pacientes (patients)	476	21	tech
29	boca (mouth)	463	25	tech
30	peso (weight)	444	25	tech

Discussion

The purpose of this study was to determine the most frequent Spanish words and lexical coverage of those words in a medical Spanish corpus. A total of 1,421,163 tokens came from 26 medical Spanish textbooks and resources. Many resources were written primarily in English with Spanish translations, and the top 3,000 Spanish words only covered 38.9% of the corpus, which is an indicator of English being a focus in these resources. The top 3,000 words may constitute a vital set of vocabulary for learners of medical Spanish, as was the case in previous research for

ESP medical students (Quero & Coxhead, 2018). Furthermore, 68.2% of the 3,000 most frequent Spanish terms were general vocabulary indicating the remaining 31.8% were technical. This percentage of 31.8% corresponds to previous studies that found that technical word lists for plumbing and finance and accountancy covered 32.17% and 23.94%, respectively, of the researchers' established corpora (Coxhead & Demecheleer, 2018; Ha & Hyland, 2017). This finding also aligns with previous medical Spanish related research that has shown that medical Spanish courses should start with intermediate language learners so that they have a basic knowledge and grasp of the language before adding technical vocabulary and concepts (Hardin, 2012; Miller De Rutté et al., 2023). It is important to note that the focus of medical Spanish courses should not be learning a certain number of vocabulary items but rather on the proficiency needed to have complex and effective communication that accounts for linguistic and cultural needs of the target language community (Hardin, 2015). Simply learning medical Spanish vocabulary does not mean that one can effectively communicate in Spanish, however, vocabulary lists can provide a starting point that can be built upon through immersive and patient-centered medical Spanish curriculum. As Quero and Coxhead (2018) assert, the creation of high frequency vocabulary lists can help students pay attention to "vocabulary that is worth learning" (p. 65), or frequently used terminology. While courses on medical Spanish entail much more than rote vocabulary memorization, the generation of these word lists can help guide the creation or reimagination of courses on medical Spanish. The planning stage, according to Quero and Coxhead (2018), is the most important stage in course design especially when designing the lexical component of the course. Using the top technical medical Spanish vocabulary terms can allow instructors to see the common medical Spanish terms that are used, which could then form not only vocabulary lists but also units of study in a course. For example, one of the top technical medical Spanish terms found in this study was *diabetes*. *Diabetes* could then form a unit of study, a discussion, a research topic, etc. for inclusion in the course. Clearly, corpus work is a component in language needs analysis research, a type of data collection and analysis that gathers multiple data points from different stakeholders and other data sources (i.e., already published materials) to determine relevant tasks and language skills needed for effective communication in a professional field (see Flowerdew, 2012; Long, 2005; Serafini & Torres, 2015; Zeller & Velazquez-Castillo, 2018 for a thorough discussion of language needs analysis). The present study may serve as one of the data sources in future language needs analyses.

These lists are not to be used as the "end all be all" to terminology in courses on undergraduate medical Spanish, but perhaps, they can be used as a reference or a starting point to implement high frequency vocabulary in courses on medical Spanish. These lists may aid both instructors and curriculum designers in the development of units, modules, courses, etc. Due to the extensive nature of these lists, they may even help in the design of sequences of courses by moving from the intermediate level to the more advanced levels through the introduction of both the general and technical vocabulary.

Furthermore, while these high frequency terms may be specialized enough for undergraduate pre-health students, it is likely that a different list would need to be created for medical Spanish education at the medical school level. Previous research has found that 95% of lexical coverage results in minimally acceptable comprehension while 98% lexical coverage has optimal comprehension (Laufer & Ravenhorst-Kalovski, 2010; Nation, 2006). These two percentages represent the possible thresholds for levels of understanding. A higher rate of lexical coverage would be needed at the graduate medical education level, and these thresholds should be considered. In fact, Hsu (2013) states, "the more restricted, specialized words with high

frequency of occurrences may be the next set of vocabulary for medical undergraduates[1] to learn after the top 3,000-word level" (p. 456). However, there is a case to be made for reducing medical jargon to improve patient-centered communication (Ortega et al., 2023), and while knowing technical vocabulary is needed for medical school students and clinicians, the ability to reduce medical jargon when speaking directly with a patient is equally important. Furthermore, this points to the fact that knowing a large set of vocabulary does not indicate effective patient communication.

Limitations and Future Research

This study is not without its limitations. First, some textbooks may have been missed due to indexing reasons. Second, due to software limitations, some accent marks in words were not read by the OCR software while at other times they were read correctly. This resulted in the author needing to manually search the corpus using the KWIC tool, and some words may have been missed. This study also looked at individual words and not phrases or expressions. Future research could examine both phrases and expressions to determine the most frequent phrases used in a medical Spanish context. Third, this current corpus analysis is based on a written corpus built from medical Spanish textbooks and resources. It does not account for any spoken language or consider the Spanish-speaking community member and their needs. Future research should develop a spoken corpus of medical Spanish encounters between health care professionals and Spanish-speaking patients and then compare the most frequent terms from the spoken corpus to this written corpus. Such analysis would help determine if the textbooks and resources are representative of spoken language used in the field. Finally, many of the authors of the textbooks and resources analyzed here are not medical specialists, and their determination of vocabulary and content inclusion was not discussed (i.e., the use of a language needs analysis or other evidence-based approaches to determine lexical inclusion). This indicates that much of the content in these textbooks and resources is most likely the result of authors' perceptions of vocabulary usage and not the result of evidence-based determinations. While these textbooks and resources are an excellent starting point, the medical Spanish field must advocate for high quality research that leads to evidence-based outcomes.

Future research should also determine the technical words beyond the first 3,000 words so that once students have mastered the first 3,000 words, they are able to go on to the next 3,000 words and have more lexical coverage of the content area, as suggested by Hsu (2013). This will also help determine which terms should be taught in medical Spanish courses at different levels. Additionally, future research should investigate how these frequent words can be used in proficiency development and testing especially as Coxhead and Nation (2001) indicate word lists provide the needed vocabulary to enhance proficiency in content areas. Future studies could also follow a similar approach in other disciplines within Spanish for Specific Purposes (i.e., Business Spanish or Legal Spanish) to determine the most frequent general and technical terminology in those areas. Furthermore, future research should compare English medical school textbooks and their most frequent vocabulary to the lists found in this study to determine if there are any discrepancies between the two sources. The results of future research may help instructors and

---

[1] "Medical undergraduates" refers to students at the beginning of their time in medical school and does not refer to the four years of undergraduate college for students who seek to earn a Bachelor's degree.

curriculum designers make more informed, evidenced-based decisions about vocabulary and content inclusion in courses on Spanish for Specific Purposes.

References

Adobe Inc. (2022). *Adobe Acrobat.* Available from https://www.adobe.com.

Anthony, L. (2022). AntConc (Version 4.1.1) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software.

Chung, T., & Nation, P. (2003). Technical vocabulary in specialized texts. *Reading in a Foreign Language, 15*(2), 103-116. https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/6c848810-a8bb-4153-8666-216791f4066d/content

Coxhead, A. (2019). Vocabulary and language for specific purposes. In C.A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (2nd ed.). https://doi.org/10.1002/9781405198431.wbeal1266.pub2

Coxhead, A. (2012). Vocabulary and language for specific purposes. In C.A. Chapelle (Ed.), *The encyclopedia of applied linguistics.* https://doi.org/10.1002/9781405198431.wbeal1266

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., McLaughlin, E., & Reid, A. (2019). The development and application of a specialized word list: The case of fabrication. *Journal of Vocational Education and Training, 71*(2), 175-200. https://doi.org/10.1080/13636820.2018.1471094

Coxhead, A., & Nation, I.S.P. (2001). The specialized vocabulary of English for academic purposes. In J. Flowerdew & M. Peacock (Eds.), *Research perspectives on English for academic purposes* (pp. 252-267). Cambridge, England: Cambridge University Press. https://doi.org/10.1017/CBO9781139524766.020

Flowerdew, L. (2012). Needs analysis and curriculum development in ESP. In B. Paltridge & S. Starfield (eds.), *The handbook of English for Specific Purposes* (pp. 325-346). Wiley. https://doi.org/10.1002/9781118339855.ch17

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

Hall, D. (2013). Teacher education for languages for specific purposes. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 5537-5542). Oxford: Wiley-Blackwell. https://doi.org/10.1002/9781405198431.wbeal1144

Hardin, K. (2012). Targeting oral and cultural proficiency for medical personnel: An examination of current medical Spanish textbooks. *Hispania, 95*(4), 698–713. https://www.jstor.org/stable/41756421

Hardin, K. (2015). An overview of medical Spanish curricula in the United States. *Hispania*, *98*(4), 640–661. https://www.jstor.org/stable/24572228

Hsu, W. (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

Lam, J. (2001). *A study of semi-technical vocabulary in computer science texts, with special reference to ESP teaching and lexicography.* Language Centre, Hong Kong University of Science and Technology. https://hdl.handle.net/1783.1/1056

Laufer, B., & Ravenhorst-Kalovski, G. (2010). Lexical threshold revisited: Lexical text

coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language, 22*(1), 15-30.

Long, M. (Ed.). (2005). *Second language needs analysis.* Cambridge University Press.

Morales, R., Rodriguez, L., Singh, A., Stratta, E, Mendoza, L., Valerio, M.A, & Vela, M. (2015). National survey of medical Spanish curriculum in US medical schools. *Journal of General Internal Medicine, 30*(10), 1434-1439. https://doi.org/10.1007/s11606-015-3309-3

Mudraya, O. (2006). Engineering English: A lexical frequency instructional model. *English for Specific Purposes, 25*(2), 235-256. https://doi.org/10.1016/j.esp.2005.05.002

Miller De Rutté, A., Kentengian, I., Pérez, M., & Ortega, P. (2023). The current state of undergraduate medical Spanish: Results of a national survey. *Foreign Language Annals*, *57(1*), 1-21. https://doi.org/10.1111/flan.12731

Nation, I.S.P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review, 63,* 59-82.

Nation, I.S.P., Coxhead, A., Chung, M., & Quero, B. (2016). Specialized word lists. In I.S.P. Nation (ed.), *Making and using word lists for language learning and testing* (pp. 145-151). Amsterdam, Netherlands: John Benjamins.

Nekrasova-Beker, T. (2020). Vocabulary demands for engineering students studying English in Russia: Comparing ESP course materials across three engineering disciplines. *Global Business Languages, 20*, 159-171. https://doi.org/10.4079/gbl.v20.9

Nesi, H. (2012). ESP and corpus studies. In B. Paltridge & S. Starfield (eds.), *The Handbook of English for Specific Purposes* (pp. 407-426). Wiley. https://doi.org/10.1002/9781118339855.ch21

Ortega, P., Cisneros, R., & Park, Y. (2023). Spanish medical jargon: A new metric for improving patient-centered communication with Spanish-speaking patients. *Patient Education and Counseling, 109,* 1-8. https://doi.org/10.1016/j.pec.2023.107644

Ortega, P., Diamond, L., Alemán, M.A., Fatás-Cabeza, J., Magaña, D., Pazo, V., Pérez, N., Girotti, J.A., Ríos, E. (2020). Medical Spanish standardization in U.S. medical schools: Consensus statement from a multidisciplinary expert panel. *Academic Medicine, 95*(1), 22-31. https://doi.org/10.1097/ACM.0000000000002917

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. Kirkgoz & K. Dikilitas (eds.), *Key issues in English for Specific Purposes in higher education* (pp. 51-75). Springer International Publishing. https://doi.org/10.1007/978-3-319-70214-8_4

Sánchez-López, L., Long, M., & Lafford, B. (2017). New directions in LSP research in US higher education. In M.K. Long (Ed.), *Languages for Specific Purposes: Trends in curriculum development* (pp. 13-34). Georgetown University Press.

Serafini, E. J., & Torres, J. (2015). The utility of needs analysis for nondomain expert instructors in designing task-based Spanish for the professions curricula. *Foreign Language Annals, 48*(3). 447-472. https://doi.org/10.1111/flan.12150

Shin, T.M., Hardin, K., Johnston, D., Rajadhyaksha, E., Diamond, L.C., Chang, R., Martinez, G.A., & Ortega, P. (2021). Scoping review of textbooks for medical Spanish education. *Medical Science Educator, 31*, 1519-1527.

Valipouri, L., & Nassaji, H. (2013). A corpus-based study of academic vocabulary in chemistry research articles. *Journal of English for Academic Purposes, 12*, 248-263. https://doi.org/10.1016/j.jeap.2013.07.001

Wang, J., Liang, S., & Ge, G. (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

Woodward-Kron, R. (2008). More than just jargon – the nature and role of specialist language in learning disciplinary knowledge. *Journal of English for Academic Purposes, 7*(4), 234-249. https://doi.org/10.1016/j.jeap.2008.10.004

Yang, M. (2015). A nursing academic word list. *English for Specific Purposes, 37*, 27-38. https://doi.org/10.1016/j.esp.2014.05.003

Zeller, S. K., & Velázquez-Castillo, M. (2018). Spanish for animal health and care: A case for expanding the LSP curriculum. In B. Lafford & C. King de Ramirez (Eds.), *Transferable skills for the 21st century: Preparing students for the workplace through world languages for specific purposes*. (pp. 287-316). Springville, UT: Sabio Books.

Appendix A

Textbooks and Reference Materials Used to Construct the Corpus

Block, C., Gray, A. E., & Willoughby, P. (2013). *El mundo hispano y la salud: Texto de español para nivel intermedio*. NCSU Campus Enterprises.

Bobenhouse, N. (2014). *McGraw-Hill Education's medical Spanish visual phrasebook*. McGraw Hill Education.

Bradley Williams, T. (2015). *Medical Spanish: A Pronto reference and study guide*. Pronto Spanish Services, LLC.

Chase, R. O., & Medina de Chase, C. B. (2019). *An introduction to medical Spanish: Communication and culture* (Fifth edition). Yale University Press.

Dejbord Sawan, P. T. (2020). *Beginning Medical Spanish: Oral proficiency and cultural humility*. Routledge, Taylor & Francis Group.

Galarreta-Aima, D. F., Segal, G., & Ruggiero, D. M. (2021). *Intermediate medical Spanish: A healthcare guide for communicating with the Latino patient*. Brown Walker Press.

Giralt, A. (2012). *Español médico y sociedad: Un libro para estudiantes de espanol en el tercer año de estudios*. Universal Publishers.

Guillet, N., & Echavarría, M. L. (2015). *Spanish, culture, and health: An introduction for healthcare professionals*. Forty Acres Press.

Hardin, K. J., & Irom, A. (2018). *Español conversacional para profesiones médicas: Manual de actividades* (Second edition). Stipes Publishing.

Harvey, W. C. (2016). *Spanish for health care professionals* (Fourth edition). Barron's.

Jany, C., & Mayberry, M. (2017). *Advanced healthcare Spanish language and culture*. World Languages and Literatures Faculty Publications. https://scholarworks.lib.csusb.edu/worldlang-publications/2

Jarvis, A. C., & Lebredo, R. (2014). *Spanish for medical personnel* (Enhanced second edition). Heinle / Cengage Learning.

Lingo Mastery. (2019). *Medical Spanish: Real Spanish medical conversations for healthcare professionals*. Lingo Mastery.

Machtinger, E., & Nigrović, P. A. (2019). *Spanish for pediatric medicine: A practical communication guide* (3rd edition). American Academy of Pediatrics.

Morris García, E., Ruiz Martínez, A., & García Arriola, M. (2019). *Profesionales de la salud: Curso de español, B1-B2*. En Clave-ELE.

Ortega, P. (2016). *Spanish and the medical interview: A textbook for clinically relevant medical Spanish* (Second edition). Elsevier.

Ortega, P., & Alemán, M. A. (2022). *Spanish and the medical interview: Clinical cases and exam review*. Elsevier.

Rebar, C. R., Heimgartner, N. M., & Gersch, C. J. (Eds.). (2019). *Medical Spanish made incredibly quick!* (Second edition). Wolters Kluwer.

Retter, S. (2017). *Medical Spanish: Fast track learning for English speakers*. UNITEXTO Publishing.

Ríos, J., Fernández Torres, J., & Rios, T. A. (2021a). *McGraw Hill's complete medical Spanish* (Premium fourth edition.). McGraw-Hill.

Ríos, J., Fernández Torres, J., & Rios, T. A. (2021b). *McGraw-Hill's Spanish for healthcare providers* (Premium fourth edition.). McGraw-Hill.

Spark Notes. (2014). *Medical Spanish SparkCharts*. SparkCharts.

Touri. (2020). *Learn medical Spanish in 100 days: Spanish words & phrases for healthcare professionals to become fluent faster* (2nd edition.). Touri Language Learning.

Traverso, S., Urraro, L., & McMahon, P. P. (2017). *Medical Spanish for nurses: A self-teaching guide*. Springer Publishing Company, LLC.

Valenzuela, R. (2013). *Spanish for nurses* (Third edition.). Kaplan Publishing.

Vox. (2012). *Vox super-mini medical Spanish and English dictionary* (3rd ed). McGraw-Hill.

Appendix B

Top 100 Spanish Words

Rank	Spanish (English) word	Freq	Range	Gen v Tech
1	de (of)	29702	26	gen
2	la (the)	27470	26	gen
3	el (the)	22421	26	gen
4	a (to)	13972	26	gen
5	que (that)	13034	25	gen
6	en (in)	12933	26	gen
7	y (and)	12425	26	gen
8	los (the)	9358	26	gen
9	o (or)	8275	26	gen
10	las (the)	7318	26	gen
11	es (is)	7016	26	gen
12	un (a/an)	6755	26	gen
13	no (no)	6175	26	gen
14	para (for)	5898	26	gen
15	con (with)	5654	26	gen
16	una (a/an)	5625	26	gen
17	se (pronoun)	5218	25	gen
18	del (from the/of the)	5102	26	gen
19	su (his/hers/theirs/yours-formal)	4943	26	gen
20	por (for through)	4589	26	gen
21	al (to the)	3963	26	gen
22	me (me/to me)	2972	25	gen
23	tiene (he/she/you-formal has)	2910	26	gen
24	le (him/her/you-formal)	2903	25	gen
25	más (more)	2301	26	gen
26	paciente (patient)	2123	25	tech
27	doctor (doctor)	2011	26	tech
28	dolor (pain)	2001	26	tech
29	lo (him/it/you-formal)	1994	26	gen
30	son (they/you all-formal are)	1974	24	gen
31	ha (has)	1945	25	gen
32	cuando (when-adverb or conjunction)	1903	25	gen
33	si (if)	1846	26	gen
34	puede (he/she/you-formal can)	1747	26	gen
35	usted (you-formal)	1679	26	gen

36	qué (what-interrogative)	1610	14	gen
37	como (as/like)	1533	24	gen
38	yo (I)	1469	26	gen
39	mi (my)	1416	25	gen
40	médico (doctor)	1405	26	tech
41	está (he/she/you-formal are)	1395	23	gen
42	salud (health)	1392	23	tech
43	ud. (abbreviation for usted/you-formal)	1254	13	gen
44	sus (his/hers/theirs/yours-formal)	1208	24	gen
45	hay (there is/there are)	1133	26	gen
46	medical (medical)	1108	24	tech
47	sobre (about/over/above)	1100	23	gen
48	dos (two)	1083	25	gen
49	hacer (to make/to do)	1081	24	gen
50	pero (but)	1070	23	gen
51	vez (time)	1020	25	gen
52	favor (favor)	1007	24	gen
53	muy (very)	988	24	gen
54	debe (he/she/you-formal should)	976	22	gen
55	te (you/to you)	966	26	gen
56	tu (your)	965	24	gen
57	tener (to have)	964	23	gen
58	va (he/she/it/you-formal go)	950	23	gen
59	sí (yes)	923	15	gen
60	bien (well)	923	25	gen
61	este (this-adjective)	896	25	gen
62	sangre (blood)	876	26	tech
63	diabetes (diabetes)	873	26	tech
64	problemas (problems)	868	23	gen
65	tomar (to take/to drink)	867	25	gen
66	cómo (how-interrogative)	856	24	gen
67	alguna (some)	853	26	gen
68	ser (to be)	851	24	gen
69	tenido (had)	844	24	gen
70	medicina (medicine)	836	24	tech
71	hace (he/she/you-formal does/do/makes/make)	826	26	gen

72	bebé (baby)	823	20	gen
73	tengo (I have)	814	24	gen
74	voy (I go)	811	26	gen
75	enfermedad (sickness)	800	25	tech
76	día (day)	786	26	gen
77	ejercicio (exercise)	781	22	tech
78	esta (this-adjective)	780	25	gen
79	hospital (hospital)	773	26	tech
80	tiempo (time)	765	25	gen
81	después (after)	755	25	gen
82	preguntas (questions)	753	20	gen
83	ir (to go)	736	25	gen
84	también (also)	724	22	gen
85	necesita (he/she/you-formal needs/need)	721	23	gen
86	síntomas (symptoms)	710	26	tech
87	comer (to eat)	705	24	gen
88	veces (times)	700	23	gen
89	personas (people)	699	22	gen
90	cada (each)	682	23	gen
91	enfermedades (sicknesses)	678	22	tech
92	mucho (many/much)	672	24	gen
93	antes (before)	667	24	gen
94	entre (among/between)	657	21	gen
95	corazón (heart)	645	21	tech
96	años (years)	642	26	gen
97	ella (she)	626	23	gen
98	duele (he/she/you-formal hurts/hurt)	624	24	tech
99	actividad (activity)	621	20	gen
100	estar (to be)	616	24	gen

Appendix C

Top 100 General Spanish Words

Rank	Spanish (English) word	Freq	Range	Gen v Tech
1	de (of)	29702	26	gen
2	la (the)	27470	26	gen
3	el (the)	22421	26	gen
4	a (to)	13972	26	gen
5	que (that)	13034	25	gen
6	en (in)	12933	26	gen
7	y (and)	12425	26	gen
8	los (the)	9358	26	gen
9	o (or)	8275	26	gen
10	las (the)	7318	26	gen
11	es (is)	7016	26	gen
12	un (a/an)	6755	26	gen
13	no (no)	6175	26	gen
14	para (for)	5898	26	gen
15	con (with)	5654	26	gen
16	una (a/an)	5625	26	gen
17	se (pronoun)	5218	25	gen
18	del (from the/of the)	5102	26	gen
19	su (his/hers/theirs/yours-formal)	4943	26	gen
20	por (for through)	4589	26	gen
21	al (to the)	3963	26	gen
22	me (me/to me)	2972	25	gen
23	tiene (he/she/you-formal has)	2910	26	gen
24	le (to him/her/you-formal)	2903	25	gen
25	más (more)	2301	26	gen
26	lo (him/it/you-formal)	1994	26	gen
27	son (they/you all-formal are)	1974	24	gen
28	ha (have)	1945	25	gen
29	cuando (when-adverb or conjunction)	1903	25	gen
30	si (if)	1846	26	gen
31	puede (he/she/you-formal can)	1747	26	gen
32	usted (you-formal)	1679	26	gen
33	qué (what-interrogative)	1610	14	gen
34	como (as/like)	1533	24	gen
35	yo (i)	1469	26	gen

36	mi (my)	1416	25	gen
37	está (he/she/you-formal are)	1395	23	gen
38	ud. (abbreviation for usted/you-formal)	1254	13	gen
39	sus (his/hers/theirs/yours-formal)	1208	24	gen
40	hay (there is/there are)	1133	26	gen
41	sobre (about/over/above)	1100	23	gen
42	dos (two)	1083	25	gen
43	hacer (to make/to do)	1081	24	gen
44	pero (but)	1070	23	gen
45	vez (time)	1020	25	gen
46	favor (favor)	1007	24	gen
47	muy (very)	988	24	gen
48	debe (he/she/you-formal should)	976	22	gen
49	te (you/to you)	966	26	gen
50	tu (your)	965	24	gen
51	tener (to have)	964	23	gen
52	va (he/she/it/you-formal go)	950	23	gen
53	sí (yes)	923	15	gen
54	bien (well)	923	25	gen
55	este (this-adjective)	896	25	gen
56	problemas (problems)	868	23	gen
57	tomar (to take/to drink)	867	25	gen
58	cómo (how-interrogative)	856	24	gen
59	alguna (some)	853	26	gen
60	ser (to be)	851	24	gen
61	tenido (had)	844	24	gen
62	hace (he/she/you-formal does/do/makes/make)	826	26	gen
63	bebé (baby)	823	20	gen
64	tengo (i have)	814	24	gen
65	voy (i go)	811	26	gen
66	día (day)	786	26	gen
67	esta (this-adjective)	780	25	gen
68	tiempo (time)	765	25	gen
69	después (after)	755	25	gen
70	preguntas (questions)	753	20	gen
71	ir (to go)	736	25	gen

72	también (also)	724	22	gen
73	necesita (he/she/you-formal needs/need)	721	23	gen
74	comer (to eat)	705	24	gen
75	veces (times)	700	23	gen
76	personas (people)	699	22	gen
77	cada (each)	682	23	gen
78	mucho (many/much)	672	24	gen
79	antes (before)	667	24	gen
80	entre (among/between)	657	21	gen
81	años (years)	642	26	gen
82	ella (she)	626	23	gen
83	actividad (activity)	621	20	gen
84	estar (to be)	616	24	gen
85	sin (without)	616	24	gen
86	días (days)	610	26	gen
87	familia (family)	572	25	gen
88	vocabulario (vocabulary)	567	14	gen
89	cuánto (how many)	565	26	gen
90	tipo (type)	551	26	gen
91	ahora (now)	548	24	gen
92	fue (he/she/it/you-formal was/went)	543	24	gen
93	persona (person)	538	23	gen
94	pueden (they/you all-formal can)	536	20	gen
95	da (he/she/it/you-formal gives/give)	532	22	gen
96	casa (house)	530	23	gen
97	cuál (interrogative-which)	528	26	gen
98	tres (three)	527	26	gen
99	tienen (they/you all-formal have)	551	26	gen
100	ya (already)	523	22	gen

Appendix D

Top 100 Technical Spanish Words

Rank	Spanish (English) word	Freq	Range	Gen v Tech
1	paciente (patient)	2123	25	tech
2	doctor (doctor)	2011	26	tech
3	dolor (pain)	2001	26	tech
4	médico (doctor)	1405	26	tech
5	salud (health)	1392	23	tech
6	medical (medical)	1108	24	tech
7	sangre (blood)	876	26	tech
8	diabetes (diabetes)	873	26	tech
11	medicina (medicine)	836	24	tech
12	enfermedad (sickness)	800	25	tech
13	ejercicio (exercise)	781	22	tech
14	hospital (hospital)	773	26	tech
15	síntomas (symptoms)	710	26	tech
16	enfermedades (sicknesses)	678	22	tech
17	corazón (heart)	645	21	tech
18	duele (he/she/you-formal hurts/hurt)	624	24	tech
19	doctora (doctor)	578	20	tech
20	presión (pressure)	568	26	tech
21	tratamiento (treatment)	562	21	tech
22	sistema (system)	553	21	tech
23	exámen (exam)	543	25	tech
24	enfermera (nurse)	543	22	tech
25	cabeza (head)	542	25	tech
26	cuerpo (body)	541	23	tech
27	piel (skin)	538	25	tech
28	pacientes (patients)	476	21	tech
29	boca (mouth)	463	25	tech
30	peso (weight)	444	25	tech
31	tos (cough)	433	24	tech
32	fiebre (fever)	431	26	tech
33	infección (infection)	428	26	tech
34	medicamento (medicine)	427	22	tech
35	médica (doctor)	419	26	tech
36	siente (feel)	414	26	tech
37	medicamentos (medicines)	409	24	tech

38	pecho (chest)	409	24	tech
39	mental (mental)	405	22	tech
40	dieta (diet)	398	25	tech
41	clínica (clinic)	387	26	tech
42	ojos (eyes)	378	24	tech
43	cáncer (cancer)	360	6	tech
44	edad (age)	357	25	tech
45	ataque (attack)	352	26	tech
46	cirugía (surgery)	344	26	tech
47	pie (foot)	334	26	tech
48	embarazo (pregnancy)	333	23	tech
49	sexual (sexual)	315	22	tech
50	orina (urine)	310	25	tech
51	parto (birth)	308	24	tech
52	seguro (sure/safe)	299	25	tech
53	virus (virus)	298	19	tech
54	drogas (drugs)	286	20	tech
55	estómago (stomach)	285	26	tech
56	respirar (to breathe)	282	25	tech
57	alcohol (alcohol)	278	21	tech
58	dientes (teeth)	277	22	tech
59	ojo (eye)	275	22	tech
60	pastillas (pills)	275	19	tech
61	brazo (arm)	270	24	tech
62	mano (hand)	264	25	tech
63	vacuna (vaccine)	259	23	tech
64	asma (asthma)	256	23	tech
65	cuello (neck)	256	25	tech
66	arterial (arterial)	253	21	tech
67	frecuencia (frequency)	253	24	tech
68	receta (prescription)	247	23	tech
69	nariz (nose)	246	23	tech
70	cita (appointment)	242	23	tech
71	anemia (anemia)	236	17	tech
72	emergencia (emergency)	231	21	tech
73	garganta (throat)	229	26	tech
74	enfermero (nurse)	224	21	tech
75	cerebral (cerebral)	219	21	tech
76	médicos (doctors)	218	19	tech

77	anestesia (anesthesia)	212	17	tech
78	espalda (back)	208	25	tech
79	dental (dental)	200	18	tech
80	dolores (pains)	200	23	tech
81	análisis (analysis)	198	19	tech
82	azúcar (sugar)	198	23	tech
83	orinar (to urinate)	196	22	tech
84	hepatitis (hepatitis)	195	19	tech
85	accidente (accident)	194	21	tech
86	temperatura (temperature)	193	23	tech
87	depresión (depression)	191	26	tech
88	gripe (flu)	190	24	tech
89	nasal (nasal	187	20	tech
90	seguro (insurance)	187	21	tech
91	consultorio (doctor's office)	185	19	tech
92	manos (hands)	183	22	tech
93	pulmones (lungs)	183	23	tech
94	operación (operation)	183	26	tech
95	muestra (sample)	181	23	tech
96	diarrea (diarrhea)	178	22	tech
97	nacimiento (birth)	178	24	tech
98	dedo (finger)	177	24	tech
99	resultados (results)	175	18	tech
100	pies (feet)	173	24	tech