Lexical Familiarization in Medical Text: A Case Study of an Excerpt from *Harrison's Principles of Internal Medicine*

Somporn Handee
Chulalongkorn University Language Institute
E-mail: hsomporn@yahoo.com

Received: December 9, 2019
Revised: December 12, 2019
Accepted: December 16, 2019

Abstract

Teaching technical or specialist vocabulary, especially in medical sciences, has been of great concern among ESP practitioners and teachers due to their lack of subject-specific background knowledge. A key strategy often used to enable teachers and learners to cope with such vocabulary while reading medical texts is lexical familiarization. This study investigated the lexical familiarization techniques in the presentation of specialist vocabulary in a well-known standard medical textbook. The samples were collected from a major reference book currently on the reading list of the medical school in Chulalongkorn University in Bangkok, Thailand. The book comprises two volumes divided into 477 chapters. From this, only Part 5: Infectious Diseases in Volume 1 was chosen for further investigation, resulting in a corpus of 3,409,894 words. The results indicate a total of 142 lexical items in 150 instances, classified into seven main categories of lexical familiarization: definition, explanation, double lexical familiarization, complex lexical familiarization, synonym, stipulation and derivation. It is suggested that ESP teachers should take different categories of lexical familiarization into account to prepare learners to be more efficient readers.

Keywords: lexical familiarization, technical or specialist vocabulary, vocabulary recognition strategies, medical text

1. Introduction

Vocabulary plays a prominent role in the field of ELT as well as ESP. Several researchers (Nation & Coady, 1988; Woodward-Kron, 2008; Bravo & Cervetti, 2009; Al-Jamal, 2018) emphasize that vocabulary or lexical knowledge is an integral component for reading comprehension. Furthermore, vocabulary size and depth of lexical knowledge have been found to contribute to success in reading, writing and academic achievement in general (Laufer, 1997; Nation & Meara, 2002; Laufer & Goldstein, 2004).

According to Gablasova (2014), knowledge of technical or specialist vocabulary appears to have a strong effect on the academic development of learners. This is because learning the meaning of new technical or specialist words is part of learning new concepts or new content (Woodward-Kron, 2008). In fact, technical vocabulary occurs frequently in specialist texts or subject areas (Nation, 2001). In this respect, ESP learners need to exert greater effort than those taking general English courses since they have to read specialist texts with an abundance of technical vocabulary.

However, research has shown that the significant role of technical vocabulary in specialized texts has been downplayed and that very little information regarding how technical vocabulary relates to other types of vocabulary is available (Sutarsyah et al., 1994; Chung, 2003; Chung & Nation, 2003). Another concern regarding whether technical or specialist vocabulary should be taught in class or not has become a controversial issue among ESP practitioners and researchers. It is believed by some that English teachers should not teach technical vocabulary to their students as they can learn and absorb it from their subject instructors and textbooks (Robinson, 1980; Kennedy & Bolitho, 1984; Nation, 1990).

On the other hand, some researchers (e.g. Williams, 1985; Nation, 1990; Swales & Feak, 2004) argue that despite such considerations, language teachers still need to prepare learners to cope with specialist or technical vocabulary they will undoubtedly encounter in their specialized texts. To do so, learners should be taught some useful vocabulary recognition strategies. According to Bramki and Williams (1984), vocabulary recognition strategies refer to the natural process in which an efficient reader decodes the meaning of an unfamiliar lexical item. Williams (1985) strongly suggests five strategies for ESP learners: inferring or guessing from context, identifying lexical familiarization, unchaining nominal compounds (compound nouns), synonym search and word analysis. However, the present study will focus mainly on lexical familiarization since this type of vocabulary recognition strategy basically reveals the way in which textbook writers carefully and intentionally provide readers with assistance to enhance their comprehension of the meanings of technical terms (Gablasova, 2014).

In the field of medicine, most English technical terms have Greek or Latin roots or base forms (Chung & Nation, 2003). Moreover, it is commonly believed that English medical terminology is internationally recognized, which, in fact, is not true (Shiland, 2014). Although many medical terms are decodable, meaning that they can be broken into their Greek or Latin word parts and given a definition based on the meanings of such elements (Shiland, 2014), e.g. *chondritis* (chondro = cartilage; -itis = inflammation), such a rule is not applicable to certain medical terms. For example, *hypochondriasis* or *hypochondria* is not related to cartilage; it refers to a fear of illness or illness

anxiety disorder. Research has also shown that the English language plays an important role in medical lectures, textbooks and journal article publications (Hwang & Lin, 2010). This is because much scientific and technological information is universally and preferentially presented in English and often originates in countries with native English speaking majorities (Creswell, 2013).

Viewed as one vocabulary recognition strategy, lexical familiarization has been investigated in several disciplines and from various perspectives, such as in the fields of business (Manan, 2007), economics (Bramki & Williams, 1984), computer science (Handee, 1996), ESL/EFL and ESP (Fraser, 1999) and comparative analysis (Chung & Nation, 2003), to name just a few. However, lexical familiarization in medical textbooks has received very little attention and few researchers have investigated which techniques are effective in teaching medical vocabulary (Najafi & Talebinezhad, 2018). Most research to date focuses on how ESL/EFL learners tackle technical and specialist vocabulary in their academic careers or how vocabulary teaching helps them to comprehend and retain technical and specialist vocabulary. For instance, Al-Jamal (2018) examined 20 EFL medical students' use of contextual clues in reading medical resources, discovering that the subjects relied solely on morphological and syntactic devices in struggling to comprehend the technical texts they encountered. In other research, Najafi and Talebinezhad (2018) investigated the effect of EFL medical vocabulary training through the use of collocations on the vocabulary retention of 80 EFL medical students, finding that the subjects receiving such training fared better than their counterparts in tests of retained word knowledge.

Therefore, this study aims to examine how technical vocabulary in a widely recognized medical textbook is familiarized and how many categories of lexical familiarization are employed by medical textbook writers. A number of useful pedagogical implications are then suggested to help prepare medical students to deal with technical or specialist vocabulary in their field.

2. Literature review

2.1 The purpose and nature of lexical familiarization

Williams (1981) defines "lexical familiarization" as the author's intentional and explicit contextual device employed when writing for a specific group of readers. It is important to bear in mind that lexical familiarization is different from a dictionary or lexicographic definition. That is, a dictionary provides a full, precise and decontextualized definition while lexical familiarization is less complete and is more reader-oriented because the author intends to help the reader gain some understanding of unfamiliar words in a context (Bramki & Williams, 1984). As far as lexical familiarization is concerned, Huckin (1981) emphasizes that those who write to non-specialist readers should use familiar concepts to explain unfamiliar ones using definitions, examples or illustrations.

2.2 Previous studies on lexical familiarization in textbooks

Williams (1981) was probably the first to conduct research on lexical familiarization in content textbooks in which the term "lexical familiarization" was coined. Four categories were proposed in his study: definition and illustration, synonym, contrast and restatement. This was followed by Bramki and Williams (1984) focusing on lexical familiarization in an economics corpus consisting of the first four chapters of an economics textbook. They reported six categories of lexical familiarization: exemplification, explanation, definition, stipulation, synonymy and illustration. Their study paved the way for other studies (Nasr, 1994; Handee, 1996; Manan, 2007).

Nasr (1994) mainly focused on the role of lexical familiarization in understanding specialist texts. Handee (1996) carried out research based on a computer science corpus comprising a computer science textbook entitled *Oh! Pascal! An Introduction to Programming*. The findings revealed 11 categories of lexical familiarization: explanation, definition, synonymy, analogy, exemplification, paraphrase, reference, clarification, derivation, double lexical familiarization and complex lexical familiarization.

Research was later conducted by Manan (2007) who investigated lexical familiarization of terms from five business textbooks and proposed 10 categories of lexical familiarization: definition, description, exemplification, paraphrase, synonymy, contrast, explanation, comparison, stipulation and analogy. Bramki and Williams (1984) subcategorize "description" and "contrast" in the category of explanation familiarization.

3. Methodology

3.1 Criteria for textbook selection for the present study

The 20th edition of *Harrison's Principles of Internal Medicine* (Volume 1) (*Harrison's*), the latest edition, was chosen for this study for two reasons. The first is that according to Doody's Review Service (n.d.), it is an authoritative medical textbook regarded as a trusted, must-have medical compendium for all medical schools, bookstores and libraries. The other is that it is also used as a medical reference book for second-year students and above in the Faculty of Medicine, Chulalongkorn University, where this study was conducted.

3.2 The corpus and the selection of lexical items

Harrison's comes in two volumes which comprise 20 parts (477 chapters), covering 3,522 pages in total. From this, only Part 5: Infectious Diseases in Volume 1 is the main focus of this study since it covers 782 pages and is considered the most comprehensive part of the book (see Appendix 1). This part comprises 115 chapters with 3,409,894 words. Furthermore, more than 100 authors have contributed to this part. This should give a deeper insight into

how textbook writers cope with technical or specialist vocabulary. Moreover, infectious diseases are deemed a great threat to humanity especially given a changing climate and novel ecosystem adaptations, for example, the case of pandemic influenzas such as SARS and MERS. In 2018, WHO and partners launched a "Global Influenza Strategy" to develop seasonal influenza prevention and control capacities.

The selection of items for further examination of lexical familiarization began with initial screening. A great deal of research shows that technical vocabulary defined is usually signaled by typographical markers, e.g. italics or bold script (Williams, 1981; Bramki & Williams, 1984; Flowerdew, 1992; Chung & Nation, 2003). However, this notion may not be applicable to the system used in *Harrison's*. This is due to the fact that such technical or specialist words have already been listed as headings or subheadings of the chapters. Nevertheless, many pathogenic species are presented in italics. The lexical items derived from the initial screening were made up of both single and multi-word items, including the names of infectious or pathogenic diseases and infectious agents (pathogens), such as bacteria, viruses, fungi and parasites. From this part, a total of 142 different lexical items were manually obtained (see Appendix 2).

3.3 Categories of lexical familiarization

The categories of lexical familiarization found in Bramki and Williams (1984) and Handee (1996) are used as the framework for this study. It should be noted here that all the examples given will be based on *Oh! Pascal! An Introduction to Programming* (Cooper, 1993) except for the example of illustration taken from *Introductory Economics* (Stanlake, 1976) and the stipulation example based on the meaning given in www.thefreedictionary.com as these two categories were not found in the Handee study (Stanlake, 1976). Moreover, technical or specialist vocabulary in these examples is presented in italics as in the original text with the highlighted part for lexical familiarization underlined.

3.3.1 Exemplification

According to Bramki and Williams (1984), an example or instance is given to identify what the newly-introduced terminology means or refers to. Common signals used are *such as*, *for example*, *for instance*, *e.g.*, *including* and *like*. For example:

> "A file identifier *like Results* will typically show up in four different places. First, it appears as a *program parameter: ...*"
>
> (Cooper, 1993, p. 437)

3.3.2 Explanation

Bramki and Williams (1984) describe explanation as providing the meaning of a newly-introduced term using a sequence of words which is equivalent or opposite in meaning to the newly-introduced term. Explanation is most commonly signaled by terms such as: *is/are, mean that, i.e.* and *concern(s).* For example:

> "*Abstraction* is the separation of application from implementation."
>
> (Cooper, 1993, p. 114)

3.3.3 Definition

A definition is seen as "an attempt to show the meaning of one word (or other linguistic expression) by means of some other words which say the same thing" (Goddard, 2011, p. 33). Therefore, it is used to explain a newly-introduced term, especially when the term is abstract or specialized (Flowerdew, 1992; Anthony, 1999; Porte, 2002; Tittle, 2011).

It is generally agreed that the format of a definition comprises three components: the "term" (definiendum) + "class" (genus) and "characteristics or distinguishing features" (differentia) (Bramki & Williams, 1984; Flowerdew, 1992). For example:

> "*Palindromes* are words or phrases that read the same forward or backward."
>
> (Cooper, 1993, p. 309)

However, the format of definition mentioned earlier can present the information the other way round, i.e. "class" + "characteristics" + "term." Common signals are as follows: *X is/are Y which …* or *X is/are called … or X is/are known/defined as …* In this respect, Swales (1971) concludes that these two formulas of definition are deemed to be equivalent in function. For example:

> "Statements that make actions repeat are called *loops*."
>
> (Cooper, 1993, p. 121)

3.3.4 Stipulation

Konecni (1978, p. 378) points out that stipulation indicates that "the term being defined has its particular meaning only in a given situation and that it does not necessarily have the same meaning in other situations." This means that stipulation limits the use of a term and its specific meaning to a particular field or discipline. For example:

> "In computer science, a *mouse* is a hand-held device used to control the cursor movement and select computing functions without keying."
>
> (HarperCollins, 2014)

3.3.5 Synonymy or synonym

Flowerdew (1992) defines synonymy as the pairing of a lexical term with an item having the same or similar meaning. It is used to avoid word repetition in writing. According to Williams (1985), it is considered stylistically inappropriate to employ the same content word adjacent to its first occurrence. For example:

> "The *result*, or value, of a function call could, for instance, be assigned to a variable or passed to a value parameter."
>
> (Cooper, 1993, p. 85)

3.3.6 Illustration

Konecni (1978, p. 369) identifies illustration as: "a rhetorical technique which refers to the combination of text and visual aid used in informational writing to clarify concepts." In other words, it is a non-verbal method of lexical familiarization using tables, charts, diagrams and so on (Bramki & Williams, 1984). Illustration is usually used in scientific texts in which many complicated concepts can be better explained using diagrams or drawings than words to describe them (Nasr, 2015). The following example shows how an illustration is used to describe *circulating capital* in economics.

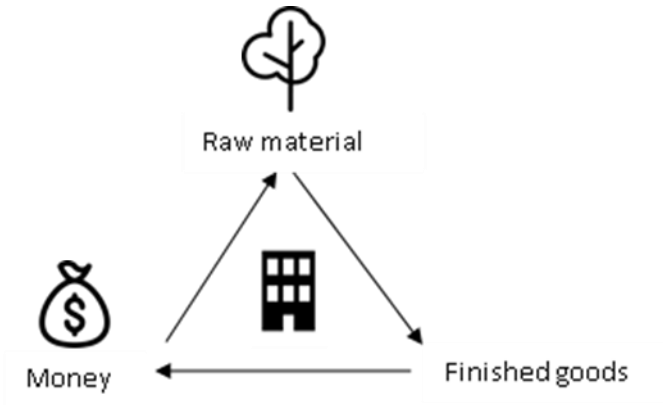

Fig. 2. Circulating capital

> "…This kind of capital is sometimes called *circulating capital* because it keeps moving and changing. Materials are changed into finished goods which are then exchanged for money and this in turn is used to buy more material" (Stanlake, 1976, p. 30-31).

3.3.7 Analogy

An analogy is usually used by textbook writers to introduce new concepts or difficult procedures by referring to something already known (Bloor, 1994). In the following example, the author explains the concept of algorithms by referring to the directions from a Tuna Helper Creamy Noodles package.

> Programmers use *algorithms* to lay out the steps a computer will take on the way to solving a problem. Now, the notion of an algorithm is hardly unique to computer science, since anybody who's written a recipe has been thinking algorithmically. But if you look at a recipe, like this one from the back of the Tuna Helper Creamy Noodles package, you'll notice something odd: the directions seem to have been extended each time a new problem came up.
>
> (Cooper, 1993, p. xxiii)

3.3.8 Paraphrase

According to Flowerdew (1992, p. 211), paraphrase is defined as "a phrase or phrases, greater than a lexical item, with the same or similar

meaning to the item it is paired with." In the following example, the writer restates the meaning of "an identical type" using a signal "i.e.", or "that is."

> "Remember that record-to-record assignments can be made only between records of an identical type, i.e. declared with the same type name."
>
> (Cooper, 1993, p. 381)

3.3.9 Reference

Bloor and Bloor (1995) suggest that an essential characteristic of cohesive reference is that the person or thing is not named when referred to for the second time. For example:

> ALTHOUGH IT'S JUST A FEW LINES LONG, program FirstRun, below, contains the essence of a thousand other programs. It's about as short and sweet as a piece of code can be, and I'm sure that you can already figure out what it does.
>
> ```
> program FirstRun (output);
>         {This is our first example program.}
> begin
>         writeIn ('Hello.  I love you.')
> end.
> ```
>
> As with every program printed in this book, the *source code* (what you see above) of FirstRun is available on diskette.
>
> (Cooper, 1993, p. xxxix)

3.3.10 Clarification

Halliday (1994, p. 226) defines clarification as: "The secondary clause clarifies the thesis of the primary clause, backing it up with some form of explanation or explanatory comment." In the following example, the textbook writer explains the meaning of a *procedure call* by specifying what it does.

> "The heart of every program generally holds at least one *procedure call*. Procedure calls do things: this particular call points a message for the program user."
>
> (Cooper, 1993, p. xi)

3.3.11 Derivation

According to Flowerdew (1992), derivation refers to a word or part of the term, usually from Greek or Latin, which is familiarized. The following example shows the prefix "mega", the word part, is familiarized.

> "Its size is given in *megabytes*, or millions of bytes."
> (Cooper, 1993, p. xxviii)

3.3.12 Double lexical familiarization

Blum-Kulka and Levenston (1983) emphasize the importance of double definition, i.e. "in explaining X it is necessary to use Y, Y being explained." Nasr (1994) highlights "synonym with proviso" or "synonym with explanation." However, Handee (1996) proposes the term "double lexical familiarization" in her computer science corpus since it is not restricted to double definition. Instead, it refers to cases where a new term is being familiarized along with another instance of lexical familiarization through different devices e.g. explanation and synonymy. The following example shows how the writer familiarizes "white space" while explaining the term "words."

> "We can define *words* as sequences of characters that are separated by *white space*–blanks, tabs, and control characters."
> (Cooper, 1993, p. 245)

3.3.13 Complex lexical familiarization

Complex lexical familiarization is used to describe cases where there is a combination of three or more lexical familiarization devices such as explanation and exemplification together with one or two embedded familiarizations within the term being familiarized (Handee, 1996). For example:

> "Using text harness is part of a more general approach to programming called *incremental program development.*
> An increment is a small amount; incremental development means programming by improvement."
> (Cooper, 1993, p. 329)

In this example, the term "incremental program development" is explained. Then, the writer specifies the meaning of "increment" in computer science and defines the meaning of "incremental development."

3.4 Classification of lexical familiarization and coding reliability

The 142 lexical items together with the corresponding concordances were uploaded onto AntConc (Version 3.5.8), concordancing software developed by Professor Lawrence Anthony (2019) of Waseda University, Japan. Following this step, the items were scrutinized to identify their lexical familiarization categories (cf. sections 3.3 and 4 below) following the frameworks postulated by Bramki and Williams (1984), Flowerdew (1992), and Handee (1996).

To ensure coding reliability, an assistant professor of English with specialization in English linguistics at Chulalongkorn University was enlisted to serve as the second coder. The coder received one-hour training during which AntConc (Anthony, 2019) and the categories of lexical familiarization were explained and clarified. Subsequently, the coder practiced classifying the first 35 lexical items, or one-fourth of all the lexical items in the study, into categories of lexical familiarization before possible areas of disagreement were negotiated and resolved. From there, the coder was given a file containing all the lexical items together with the corresponding concordances, including those previously rehearsed, while being advised to analyze the data carefully to avoid errors. Then, the coder classified the lexical items independently over a one-week period. Upon completion of these procedures, coding reliability was calculated using Cohen's Kappa (Cohen, 1960) to determine correspondence in the classification of lexical familiarization categories carried out by the present author and the second coder. The Cohen's Kappa coefficient was 77.47%, which indicates a substantial degree of coding agreement.

4. Results

In total, 150 instances of lexical familiarization which were related to 142 lexical items were found in the corpus. They were divided into seven main categories: definition, explanation, double lexical familiarization, complex lexical familiarization, synonym, stipulation and derivation (Table 1). Note that the discrepancy in the number of lexical items and lexical familiarization instances was due to the fact that some of the lexical items were familiarized more than once.

Table 1. Summary of categories and number of occurrences of lexical familiarization in the corpus

Category	Number of occurrences of lexical familiarization (instances) and percentage (%)
1. definition	80 (53.33%)
2. explanation	29 (19.33%)
3. double lexical familiarization	24 (16%)
4. complex lexical familiarization	8 (5.33%)
5. synonym	7 (4.67%)
6. stipulation	1 (0.67%)
7. derivation	1 (0.67%)
Total instances	**150**

4.1 Definition

Table 1 shows that familiarizing medical terms by means of definition had the highest distribution of all forms of lexical familiarization used in Part 5 (Infectious Diseases) of *Harrison's*. In all, 80 instances were found in the corpus, accounting for 53.33% of occurrences of lexical familiarization in the sample.

Common syntactic signals of this category are as follows:

1. X is/are Y + which/that …
   + whose …
2. X is/are Y + reduced relative clauses
   (present participle (V.ing) / past participle)
3. Y + which/that … is called …
4. X is defined as Y …
5. The term X refers to Y …
6. The term X is used to describe Y …

For example:

- *T. gondii* is an intracellular coccidian that infects both birds and mammals. (Jameson et al., 2018, p. 1609)
- Encephalitis is defined as an inflammation of the brain caused either by infection, usually with a virus, or from a primary autoimmune process. (Jameson et al., 2018, p. 991)

Note that any instance which does not correspond to the formulas mentioned above will be grouped in the category of explanation familiarization. In the corpus, one particular aspect regarding definition can be observed: authors

in *Harrison's* usually use a premodifier in front of "the class", or a postmodifier after "the class." In the following example, the premodifier and postmodifier are underlined.

- Rabies is a rapidly progressive, acute infectious disease of the central nervous system (CNS) in humans and animals that is caused by infection with rabies virus. (Jameson et al., 2018, p. 1484)

4.2 Explanation

Twenty-nine instances of lexical familiarization through explanation were found, making up 19.33% of the total. To explain the meaning of a medical term, textbook writers use a sequence of words, a phrase, a sentence or several sentences to achieve an exact meaning of that particular term. Common signals of this category are *is* and *are*. For example:

- Pneumonia is an infection of the pulmonary parenchyma. (Jameson et al., 2018, p. 908)
- Nematodes are elongated, symmetric roundworms. (Jameson et al., 2018, p. 1621)

4.3 Double lexical familiarization

Twenty-four instances of double lexical familiarization were found in the corpus, contributing another 16% to the total. Double lexical familiarizations are presented by a combination of different means such as synonym and definition and explanation and definition. For example:

- Trematodes, or flatworms, are a group of helminths that belong to the phylum Platyhelminthes. (Jameson et al., 2018, p. 1635)

In this example, the synonym of "trematodes" is given while the term is being defined.

- *Haemophilus ducreyi* is the etiologic agent of chancroid, a sexually transmitted disease characterized by genital ulceration and inguinal adenitis. (Jameson et al., 2018, p. 1131)

In this example, while the term "*Haemophilus ducreyi*" is being explained, another term "chancroid" is defined.

4.4 Complex lexical familiarization

Complex lexical familiarization refers to cases in which a combination of three or more lexical familiarization devices are used, for example, explanation and definition in conjunction with one or two embedded familiarizations within the term being familiarized (Handee, 1996). Eight instances of complex lexical familiarization were found in the corpus. For example:

> • Onchocerciasis ("river blindness") is caused by the filarial nematode *O. volvulus*, which infects an estimated 37 million individuals in 31 countries worldwide. (Jameson et al., 2018, p. 1632)

In this example, the synonym of "onchocerciasis" is given while the term is explained. Also, additional information about "*O. volvulus*" is provided. Note that a non-defining relative clause is grouped in the category of explanation as its function is to provide additional information about "*O. volvulus*."

> • Visceral Leishmaniasis (also known as *kala-azar*, a Hindi term meaning "black fever") is caused by the *Leishmania donovani* complex, which includes *L. donovani* and *Leishmania infantum* (the latter designated *Leishmania chagasi* in the New World). (Jameson et al., 2018, p. 1594)

In this example, the etymology of the term with its meaning is given while the term "Visceral Leishmaniasis" is explained while more explanation about another term "the *Leishmania donovani* complex" is provided.

4.5 Synonym or synonymy

Synonym or synonymy in this study means the pairing of a lexical item including multi-word units such as "German measles", phrasal verbs and set phrases with an item having the same or similar meaning (Flowerdew, 1992). Thus, any instances that do not meet this criterion will be grouped in the category of explanation familiarization.

In the analysis, seven synonyms which stood alone were found. Those which occurred with other means of lexical familiarization were grouped in the category of double lexical familiarization. Synonymy in the corpus is signaled by *or* and *punctuation marks* such as parentheses or brackets ( ) and quotation marks (" "). For example:

- Varicella-zoster virus (VZV) causes two distinct clinical entities: varicella (chickenpox) and herpes zoster (shingles). (Jameson et al., 2018, p. 1354)
- The word *helminth* is derived from the Greek *helmins* ("parasitic worm"). (Jameson et al., 2018, p. 1620)

4.6 Stipulation

According to Konecni (1978), stipulation restricts the use of the term and its meaning to a particular field or situation. Sometimes, it is also employed in the same field (Nasr, 2015). Only one instance of stipulation was found in the corpus.

- Although the same is true of many bacteria and viruses, the designation parasite is reserved, by convention, for helminths and protozoa. (Jameson et al., 2018, p. 1551)

The signal of stipulation familiarization used in this example is "*is reserved for*."

4.7 Derivation

Flowerdew (1992) describes derivation as a word or word element from the term, usually from Greek or Latin, which is familiarized. Only one instance of derivation was found in the corpus.

- The word *parasite* comes originally from the Greek *parasitos* (*para*, alongside of; and *sitos*, food), meaning someone who eats at another's table or lives at another's expense. (Jameson et al., 2018, p. 1551)

In this example, the word parts familiarized are the prefix "para" and the Greek root "sitos."

5. Discussion

The present study reveals seven main categories of lexical familiarization found in an excerpt of a medical textbook: definition, explanation, double lexical familiarization, complex lexical familiarization, synonym, stipulation and derivation. These seven categories are in line with those found in the computer science corpus studied in Handee (1996). The findings may reflect that textbook authors in scientific disciplines usually employ similar techniques when dealing with technical or specialist vocabulary. Moreover, the two most frequently encountered categories in the medicine and the computer science corpus are definition and explanation. The findings of definition familiarization in this

study are also consistent with the notion in the literature that definitions are a common means of lexical familiarization in textbooks and lectures (Bramki & Williams, 1984; Flowerdew, 1992; Nation, 2001). This is because the patterns of definitions can provide an adequate amount of semantic and syntactic information about a technical or specialist word, which can help the reader to construct the meaning of such a particular word (Gablasova, 2014). This notion also supports Krashen's claim (2003) that difficult terms can only be figured out by understanding them appropriately. Therefore, the role of definitions should be emphasized in ESP classes.

Apart from these two categories, other techniques such as using a combination of two types of familiarization or more also play an important role in *Harrison's*. Additionally, the findings suggest that the role of typographical devices i.e. bolding or italics, used to signal technical vocabulary to be defined, is not as significant as that in other disciplines (e.g. Bramki & Williams, 1984; Flowerdew, 1992; Chung & Nation, 2003). This is perhaps because most of the terms have been listed as headings of chapters; however, many newly-introduced terms tend to be italicized. Moreover, *Harrison's* is aimed at those who already have some background knowledge in medicine; accordingly, highlighting technical terms may be considered unnecessary. Still, the criteria for such a system is not known.

Several researchers (e.g. Williams, 1985; Nation, 1990; Chung & Nation, 2003) emphasize the teaching of technical or specialist vocabulary. Williams (1985) also highly recommends five strategies in ESP reading: inferring or guessing from context, identifying lexical familiarization, unchaining nominal compounds, synonym search and word analysis.

As far as teaching technical or specialist vocabulary is concerned, it seems rather dangerous to teach students to infer or guess from context because technical or specialist vocabulary, especially in the field of medicine, usually has a precise meaning which must be understood precisely (Blum-Kulka & Levenston, 1983); thus, students need to know the exact or the nearly exact meaning or concept of such particular technical vocabulary. Therefore, it can be concluded here that it is lexical familiarization that should be regarded as the most effective vocabulary recognition strategy supporting students when reading medical textbooks on their own because textbook writers intentionally employ various lexical familiarization techniques to help students work out the meaning of unfamiliar technical or specialist vocabulary (Bramki & Williams, 1984; Gablasova, 2014).

Since vocabulary recognition strategies, especially lexical familiarization, cannot be taught by means of vocabulary lists (Williams, 1985), common categories of lexical familiarization e.g. definition and explanation, should be introduced in ESP reading. Furthermore, an ESP course must include the explicit teaching of vocabulary which occurs in authentic written texts (Donesch-Jezo &

Pachonska-Wolowska, 2014). In this respect, data from a corpus, a collection of authentic texts stored electronically for the purpose of language study (Parltridge, 2006), can be used. In fact, a corpus can be a very useful authentic data source for language teachers (Li, 2015) and, simultaneously, can help students see the authentic and contextualized uses of vocabulary.

A corpus-based concordance, a list of occurrences of the words with their textual environment (Sinclair, 1991), can also be employed to encourage students to discover linguistic aspects by themselves (Donesch-Jezo & Pachonska-Wolowska, 2014). This notion corresponds to the principles of data driven learning, an approach to foreign language learning in which language is treated as data and students as researchers (Johns, 1991). For instance, a printed version of a concordance with the keyword "influenza" may be distributed to the class. Students would then be asked to identify the definition of "influenza." This activity would enable them to see how "influenza" in a medical textbook is familiarized by means of a definition (see Fig. 1).

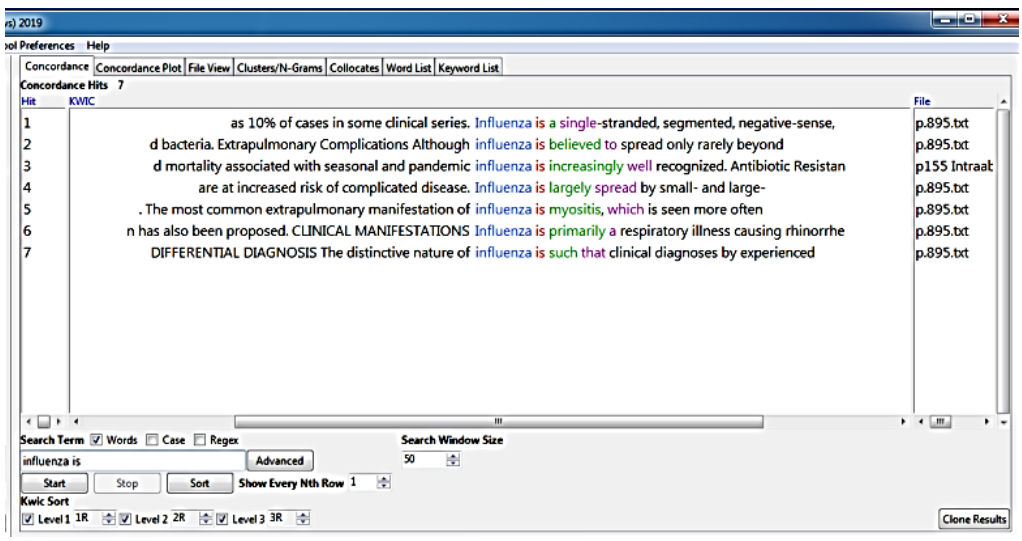

Fig. 1. Screenshot showing a concordance for the keyword "influenza"

The present study has two limitations. The first is that only a certain part of *Harrison's* was investigated, so it may not be inclusive of all techniques by which lexical familiarization has been developed throughout the entire textbook. The other is the number of medical terms used in the study may not cover other types of lexical familiarization employed by writers in other chapters of the textbook. Further research, if possible, should investigate the whole textbook of *Harrison's Principles of Internal Medicine* to determine whether other categories of lexical familiarization remain to be explored. Also, future research should look into other disciplines to see how textbook writers in other fields cope with technical or specialist vocabulary.

6. Conclusion

The findings of this study confirm the importance of lexical familiarization in a medical textbook and, at the same time, reveal the methods medical textbook writers employ to communicate technical or specialist vocabulary to the reader. Therefore, lexical familiarization can be viewed as a significant vocabulary recognition strategy used by textbook authors to facilitate the reader when encountering unfamiliar medical terms.

Strictly speaking, the complexity of medical terminology can be intimidating to ESP instructors who have to teach ESP courses to medical students. However, when scrutinizing technical or specialist vocabulary in detail, we, as English teachers, can adopt a variety of techniques to understand and communicate it to learners. One solution is to use lexical familiarization as a tool to provide the student with a more efficient reading strategy in a highly technical field such as medicine. From this, it can be concluded that lexical familiarization ought to be integrated into reading courses in ESP classes.

References

Al-Jamal, D. (2018). The role of linguistic clues in medical students' reading comprehension. *Psychology Research and Behavior Management, 11*, 395-401.

Anthony, L. (1999). Writing research article introductions in software engineering: How accurate is a standard model? *IEEE Transactions on Professional Communication*, *42*(1), 38-46.

Anthony, L. (2019). *AntConc* (Version 3.5.8) [Computer Software]. Waseda University. https://www.laurenceanthony.net/software

Bloor, M. (1996). Academic writing in computer science: A comparison of genres. In E. Ventola & A. Mauranen (Eds.), *Academic Writing: Intercultural and textual issues* (pp. 59-87). John Benjamins Publishing Company.

Bloor, T., & Bloor, M. (1995). *The functional analysis of English: A Hallidayan approach.* Edward Arnold.

Blum-Kulka, S., & Levenston, E. A. (1983). Universals of lexical simplification. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 119-139). Longman.

Bramki, D., & Williams, R. C. (1984). Lexical familiarization in economics text and its pedagogic implications in reading comprehension. *Reading in a Foreign Language, 2*(1), 169-181.

Bravo, M. A., & Cervetti, G. N. (2009). Teaching vocabulary through text and experience in content areas. In M. F. Graves (Ed.), *Essential reading on vocabulary instruction* (pp. 141-152). International Reading Association.

Chung, T. M. (2003). *Identifying technical vocabulary* [Unpublished doctoral dissertation]. Victoria University of Wellington.

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

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37-46.

Cooper, D. (1993). *Oh! Pascal!* (3rd ed.). W. W. Norton and Company.

Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). SAGE.

Donesch-Jezo, E., & Pachonska-Wolowska, A. (2014). Teaching of vocabulary to medical students in ESP courses. *LangLit: An International Peer-Reviewed Open Access Journal, 1*(1), 72-87.

Doody's Review Service. (n.d.). *Editorial Reviews*. https://www.barnesandnoble.com/w/harrisons-principles-of-internal-medicine-19-e-dennis-kasper/1124310287

Flowerdew, J. (1992). Definition in science lectures. *Applied Linguistics, 13*(2), 202-221.

Fraser, C. A. (1999). Lexical processing strategy use and vocabulary learning through reading. *SSLA, 21*(2), 225-241.

Gablasova, D. (2014). Learning and retaining specialized vocabulary from textbook reading: Comparison of learning outcomes through L1 and L2. *The Modern Language Journal, 98*(4), 976-991.

Goddard, C. (2011). *Semantic analysis: A practical introduction* (2nd ed.). Oxford University Press.

Halliday, M. A. K. (1994). *An introduction to functional grammar (*2nd ed.). Edward Arnold.

Handee, S. (1996). *Lexical familiarisation in computer science text: A preliminary investigation* [Unpublished master's dissertation]. University of Warwick.

HarperCollins. (2014). *Mouse*. https://www.thefreedictionary.com/mouse

Huckin, T. (1981). Readability from a cognitive perspective. In P. Anderson, C. Miller, & J. Brookman (Eds.), *New essays in technical and scientific communication: Theory, research and criticism* (pp. 25-44). Baywood Press.

Hwang, Y., & Lin, S. (2010). A study of medical students' linguistic needs in Taiwan. *The Asian ESP Journal, 6*(1), 35-58.

Jameson, J. L., Kasper, D. L., Longo, D. L., Fauci, A. S., Hauser, S. L., & Loscalzo, J. (Eds.). (2018). *Harrison's Principles of Internal Medicine* (20th ed., Vol. 1). McGraw-Hill Education.

Johns, T. (1991). From printout to handout: Grammar and vocabulary teaching in the context of data driven learning. *ELR Journal: Classroom concordancing, 4*, 27-45.

Kennedy, C., & Bolitho, R. (1984). *English for specific purposes.* Macmillan.

Konecni, E. (1978). Scientific and technical rhetoric. In L. Trimble & M. T. Trimble (Eds.), *English for specific purposes: Science and technology* (pp. 359-386). Oregon State University Press.

Krashen, S. D. (2003). *Explorations in language acquisition and use: The Taipei lectures*. Heinemann.

Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 20-34). Cambridge University Press.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning, 53*(3), 399-436.

Li, K. (2015). The use of concordance programs in English lexical teaching in high school. *Higher Education of Social Science, 8*(1), 60-65.

Manan, A. (2007). Lexical familiarization of terms from five business textbooks. *ESP Malaysia*, *13*, 55-70.

Najafi, M., & Talebinezhad, M. R. (2018). The impact of teaching EFL medical vocabulary through collocations on vocabulary retention of EFL medical students. *Advances in Language and Literary Studies, 9*(5), 24-27.

Nasr, A. (1994). *The role of lexical familiarization in understanding specialist texts* [Unpublished master's dissertation]. University of Warwick.

Nasr, A. A. M. (2015). Vocabulary recognition strategies: Theory and practice within ESP domains. *English for Specific Purposes World*, *1*6(1), 1-19. http://www.philologician.com/1_Special_Articles/Atef_Vocabulary%20.pdf

Nation, I. S. P. (1990). *Teaching and learning vocabulary.* Newbury House.

Nation, I. S. P. (2001). *Learning vocabulary in another language.* Cambridge University Press.

Nation, P., & Coady, J. (1988). Vocabulary and reading. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 97-110). Longman.

Nation, P., & Meara, P. (2002). Vocabulary. In N. Schmitt (Ed.), *An introduction to applied linguistics* (pp. 35-54). Arnold.

Paltridge, B. (2006). *Discourse analysis: An introduction.* Continuum.

Porte, G. K. (2002). *Appraising research in second language learning: A practical approach to critical analysis of quantitative research.* John Benjamins.

Robinson, P. (1980). *ESP (English for specific purposes): The present position.* Pergamon.

Shiland, B. J. (2014). *Medical terminology & anatomy for ICD-10 coding* (2nd ed.). Mosby. https://www.elsevier.com/books/medical-terminology-and-anatomy-for-icd-10-coding/unknown/978-0-323-26017-6

Sinclair, J. (1991). *Corpus, concordance, collocation.* Oxford University Press.

Stanlake, G. F. (1976). *Introductory Economics* (3rd ed.). Longman.

Sutarsyah, C., Nation, P., & Kennedy, G. (1994). How useful is EAP vocabulary for ESP? A corpus based study. *RELC Journal*, *25*(2), 34-50.

Swales, J. M. (1971). *Writing scientific English*. Nelson.

Swales, J., & Feak, C. (2004). *Academic writing for graduate students: Essential tasks and skills* (2nd ed.). University of Michigan Press.

Tittle, P. (2011). *Critical thinking: An appeal to reason.* Routledge.

Williams, R. C. (1981). Lexical familiarization in content area textbooks. In L. J. Chapman (Ed.), *The reader and the text* (pp. 49-59). Heinemann.

Williams, R. C. (1985). Teaching vocabulary recognition strategies in ESP reading. *The ESP Journal, 4*(2), 121-131.

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.

Appendix 1

Part 5: Infectious Diseases

Section 1: Basic Considerations in Infectious Diseases

115: Approach to the Patient with an Infectious Disease
116: Molecular Mechanisms of Microbial Pathogenesis
117: Approach to the Acutely Ill Infected Febrile Patient
118: Immunization Principles and Vaccine Use
119: Health Recommendations for International Travel
120: Climate Change and Infectious Disease

Section 2: Clinical Syndromes: Community-Acquired Infections

121: Pneumonia
122: Lung Abscess
123: Infective Endocarditis
124: Infections of the Skin, Muscles, and Soft Tissues
125: Infectious Arthritis
126: Osteomyelitis
127: Intraabdominal Infections and Abscesses
128: Acute Infectious Diarrheal Diseases and Bacterial Food Poisoning
129: *Clostridium difficile* Infection, Including Pseudomembranous Colitis
130: Urinary Tract Infections, Pyelonephritis, and Prostatitis
131: Sexually Transmitted Infections: Overview and Clinical Approach
132: Encephalitis
133: Acute Meningitis
134: Chronic and Recurrent Meningitis
135: Brain Abscess and Empyema
136: Infectious Complications of Bites

Section 3: Clinical Syndromes: Health Care-Associated Infections

137: Infections Acquired in Health Care Facilities
138: Infections in Transplant Recipients

Section 4: Therapy for Bacterial Diseases

139: Treatment and Prophylaxis of Bacterial Infections
140: Bacterial Resistance to Antimicrobial Agents

Section 5: Diseases Caused by Gram-Positive Bacteria

141: Pneumococcal Infections
142: Staphylococcal Infections
143: Streptococcal Infections
144: Enterococcal Infections
145: Diphtheria and Other Corynebacterial Infections

146: *Listeria monocytogenes* Infections
147: Tetanus
148: Botulism
149: Gas Gangrene and Other Clostridial Infections
Section 6: Diseases Caused by Gram-Negative Bacteria
150: Meningococcal Infections
151: Gonococcal Infections
152: *Haemophilus* and *Moraxella* Infections
153: Infections Due to the HACEK Group and Miscellaneous Gram-Negative Bacteria
154: *Legionella* Infections
155: Pertussis and Other *Bordetella* Infections
156: Diseases Caused by Gram-Negative Enteric Bacilli
157: *Acinetobacter* Infections
158: *Helicobacter pylori* Infections
159: Infections Due to *Pseudomonas*, *Burkholderia*, and *Stenotrophomonas* Species
160: Salmonellosis
161: Shigellosis
162: Infections Due to *Campylobacter* and Related Organisms
163: Cholera and Other Vibrioses
164: Brucellosis
165: Tularemia
166: Plague and Other *Yersinia* Infections
167: *Bartonella* Infections, Including Cat-Scratch Disease
168: Donovanosis
Section 7: Miscellaneous Bacterial Infections
169: Nocardiosis
170: Actinomycosis
171: Whipple's Disease
172: Infections Due to Mixed Anaerobic Organisms
Section 8: Mycobacterial Diseases
173: Tuberculosis
174: Leprosy
175: Nontuberculous Mycobacterial Infections
176: Antimycobacterial Agents
Section 9: Spirochetal Diseases
177: Syphilis
178: Endemic Treponematoses
179: Leptospirosis
180: Relapsing Fever
181: Lyme Borreliosis

Section 10: Diseases Caused by Rickettsiae, Mycoplasmas, and Chlamydiae
182: Rickettsial Diseases
183: Infections Due to Mycoplasmas
184: Chlamydial Infections
Section 11: Viral Diseases: General Considerations
185: Medical Virology
186: Antiviral Chemotherapy, Excluding Antiretroviral Drugs
Section 12: Infections Due to DNA Viruses
187: Herpes Simplex Virus Infections
188: Varicella-Zoster Virus Infections
189: Epstein-Barr Virus Infections, Including Infectious Mononucleosis
190: Cytomegalovirus and Human Herpesvirus Types 6, 7, and 8
191: Molluscum Contagiosum, Monkeypox, and Other Poxvirus Infections
192: Parvovirus Infections
193: Human Papillomavirus Infections
Section 13: Infections Due to DNA and RNA Respiratory Viruses
194: Common Viral Respiratory Infections
195: Influenza
Section 14: Infections Due to Human Immunodeficiency Virus and Other Human Retroviruses
196: The Human Retroviruses
197: Human Immunodeficiency Virus Disease: AIDS and Related Disorders
Section 15: Infections Due to RNA Viruses
198: Viral Gastroenteritis
199: Enterovirus, Parechovirus, and Reovirus Infections
200: Measles (Rubeola)
201: Rubella (German Measles)
202: Mumps
203: Rabies and Other Rhabdovirus Infections
204: Arthropod-Borne and Rodent-Borne Virus Infections
205: Ebolavirus and Marburgvirus Infections
Section 16: Fungal Infections
206: Diagnosis and Treatment of Fungal Infections
207: Histoplasmosis
208: Coccidioidomycosis
209: Blastomycosis
210: Cryptococcosis
211: Candidiasis

212: Aspergillosis
213: Mucormycosis
214: Superficial Mycoses and Less Common Systemic Mycoses
215: *Pneumocystis* Infections
Section 17: Protozoal and Helminthic Infections: General Considerations
216: Introduction to Parasitic Infections
217: Agents Used to Treat Parasitic Infections
Section 18: Protozoal Infections
218: Amebiasis and Infection with Free-Living Amebae
219: Malaria
220: Babesiosis
221: Leishmaniasis
222: Chagas Disease and African Trypanosomiasis
223: *Toxoplasma* Infections
224: Protozoal Intestinal Infections and Trichomoniasis
Section 19: Helminthic Infections
225: Introduction to Helminthic Infections
226: Trichinellosis and Other Tissue Nematode Infections
227: Intestinal Nematode Infections
228: Filarial and Related Infections
229: Schistosomiasis and Other Trematode Infections
230: Cestode Infections

Appendix 2

The list of 142 lexical items based on order of occurrence

1. Disease
2. Measles
3. Pneumonia
4. *Lung abscess*
5. *Infective endarteritis*
6. *Acute endocarditis*
7. *Pseudomonas infection*
8. Cellulitis
9. Osteomyelitis
10. Vertebral osteomyelitis
11. Peritonitis
12. *Clostridium difficile* infection (CDI)
13. *C. difficile*
14. Urinary tract infection (UTI)
15. *Uncomplicated urinary tract infection*
16. Mucopurulent cervicitis (MPC)
17. *Pelvic inflammatory disease*
18. Encephalitis
19. SSPE (Subacute sclerosing panencephalitis)
20. *Bacterial meningitis*
21. Brain abscess
22. Subdural empyema (SDE)
23. Bacterial resistance
24. Pneumococci
25. *S. aureus*
26. *Pyomyositis*
27. Impetigo
28. Diphtheria
29. *C. diphtheriae*
30. *C. urealyticum*
31. Erythrasma
32. *C. xerosis*
33. Tetanus
34. *C. tetani*
35. Botulism
36. *Enteritis necroticans*
37. *Necrotizing enterocolitis*
38. *Crepitant cellulitis*
39. *N. meningitidis*

40. Gonorrhea
41. *Neisseria gonorrhoeae*
42. *H. influenzae*
43. *Epiglottitis*
44. *Haemophilus ducreyi*
45. *M. catarrhalis*
46. HACEK organisms
47. *P. multocida*
48. *Achromobacter xylosoxidans*
49. *Elizabethkingia meningoseptica*
50. *Shewanella species*
51. *Chromobacterium violaceum*
52. *Legionellosis*
53. Pontiac fever
54. Pertussis
55. *Bordetella* species
56. *Acinetobacter* species
57. *H. pylori*
58. Pseudomonads
59. *P. aeruginosa*
60. *B. cepacia*
61. *B. pseudomallei*
62. *S. maltophilia*
63. Salmonellae
64. *Shigella*
65. *Campylobacter*
66. Cholera
67. Brucellosis
68. Tularemia
69. *F. tularensis*
70. Plague
71. *Yersiniosis*
72. *Bartonella* species
73. Donovanosis
74. Nocardiosis
75. Actinomycosis
76. Whipple’s disease
77. Empyema
78. Tuberculosis (TB)
79. Leprosy
80. Syphilis
81. Endemic treponematoses

82. Leptospirosis
83. Rickettsiae
84. *Brill-Zinsser disease*
85. Ehrlichioses
86. Chlamydiae
87. Viruses
88. Varicella-zoster virus (VZV)
89. Herpes zoster
90. Molluscum contagiosum virus
91. Parvoviruses
92. Influenza
93. HIV
94. HIV infection
95. *Rhodococcus equi*
96. *Coccidioides immitis*
97. Microsporidia
98. Acute infectious gastroenteritis
99. Norwalk virus
100. *Picobirnaviruses*
101. Rubella
102. Mumps
103. Rabies
104. Zika virus
105. *Histoplasma capsulatum*
106. Coccidioidomycosis
107. Blastomycosis
108. *Cryptococcus*
109. *Candida*
110. *Aspergillosis*
111. Mucorales
112. *Sporothrix schenckii*
113. *Talaromyces marneffei*
114. *Pneumocystis*
115. *Parasite*
116. *Protozoa*
117. Roundworms
118. *Entamoeba histolytica*
119. *Toxoplasma gondii*
120. Amebiasis
121. *Balamuthia mandrillaris*
122. Malaria
123. Babesiosis

124. VL (Visceral Leishmaniasis)
125. Chagas disease
126. Toxoplasmosis
127. *Congenital* toxoplasmosis
128. *T. gondii*
129. *Giardia intestinalis*
130. Giardiasis
131. Microsporidia
132. *Balantidium coli*
133. *Helminth*
134. Nematodes
135. Visceral larva migrans
136. Cutaneous larva migrans
137. *Trichostrongylus species*
138. Anisakiasis
139. Filarial worms
140. Tropical pulmonary eosinophilia
141. Onchocerciasis
142. Trematodes