

4-2018

Linguistic Features of Medical English for Curricular and Instructional Purposes

Yi Shi

Grand Valley State University

Follow this and additional works at: <https://scholarworks.gvsu.edu/theses>

 Part of the [Linguistics Commons](#)

Recommended Citation

Shi, Yi, "Linguistic Features of Medical English for Curricular and Instructional Purposes" (2018). *Masters Theses*. 886.
<https://scholarworks.gvsu.edu/theses/886>

This Thesis is brought to you for free and open access by the Graduate Research and Creative Practice at ScholarWorks@GVSU. It has been accepted for inclusion in Masters Theses by an authorized administrator of ScholarWorks@GVSU. For more information, please contact scholarworks@gvsu.edu.

Linguistic Features of Medical English for Curricular and Instructional Purposes

Yi Shi

A Thesis Submitted to the Graduate Faculty of

GRAND VALLEY STATE UNIVERSITY

In

Partial Fulfillment of the Requirements

For the Degree of

Master of Arts in Applied Linguistics

English Department

April 2018

Abstract

In China, medical English as a specific target language use domain has been of high need to Chinese doctors and college medical students. It is mainly used for publishing research articles in international medical journals and presenting at international medical conferences. For this reason, the study of linguistic characteristics of medical English will benefit Chinese doctors and medical students in their efforts to learn English in the field of medicine.

Nevertheless, “medical English” is recognized as a concentration of study as part of the English major at some universities in China while in other colleges, medical students just learn generic English due to a lack of a clear operational definition of “medical English”. The problem arises when colleges and universities purport to teach “medical English” to their students in the field of medicine because their curricula can have high variance and instructors have no guidance for “medical English”. Under the circumstance, this study examines the specific features of medical English in an effort to clearly delineate medical English for curricular and instructional purposes. Specifically, the linguistic features that characterize medical English are explored. This exploration relies on corpus linguistic data collected from 15 English-medium medical research articles (RAs) selected from five medical journals published in 2017 in order to examine four specific lexical and syntactic features of medical English. The procedure was executed manually and by computer. The findings suggest that medical English can be characterized as a legitimate register of English with the linguistic features of proportional specialized medical vocabulary, nominalization,

passivization, and sentence complexity. Instructors of English are pedagogically recommended to pay attention to specialized medical terminology and nominalized forms, introduce passive voice, and focus on clauses and complex sentence structure when teaching English to medical students.

Keywords: English for Specific Purposes, corpus linguistics, medical English

Table of Contents

Abstract	3
Table of Contents	5
List of Tables.....	6
Chapter 1 Introduction	7
Chapter 2 Review of Literature.....	11
2.1 Development of English for Specific Purposes (ESP).....	11
Background introduction	11
Development	13
2.2 Previous research on medical English in the EMP field.....	18
Lexical features	18
Syntactic features and rhetorical analysis	23
Use of medical English in real life.....	27
Chapter 3 Methodology	31
3.1 Data collection	31
3.2 Data analysis	33
Frequency of occurrence of specialized medical vocabulary	33
Nominalization.....	34
Passivization	34
Sentence complexity	35
3.3 Rating procedure	35
Chapter 4 Results	37
4.1 Frequency of occurrence of specialized medical words	37
4.2 Frequency of occurrence of nominalized words	38
4.3 Proportions of voices	40
4.4 Ratio of clause to sentence.....	43
Chapter 5 Discussion and Conclusion	46
5.1 Discussion	46
Lexical features	46
Syntactic features	48
5.2 Pedagogical implications	49
5.3 Limitations	50
5.4 Conclusion	51
Bibliography	52

List of Tables

Table 1 Frequency of occurrence of specialized medical words in medical RAs (words per 1,000 words).....	38
Table 2 Frequency of occurrence of nominalized words ending with <i>-tion/sion</i> , <i>-ment</i> , <i>-ness</i> , and <i>-ity</i> (and their plural forms) in medical RAs (words per 1,000 words).....	39
Table 3 Frequency of occurrence of active/passive/“null” voice in medical RAs (normed per 1,000 words).....	41
Table 4 Percentages of voices (active, passive, and null) in medical RAs.....	42
Table 5 Ratio of clause to sentence.....	43

Chapter 1 Introduction

The discourse of English for Specific Purposes (ESP) has emerged since the 1960s. It has been developing as an academic discipline with historical and social changes across the globe. After the end of World War II (WWII) in 1945, the United States arose as the only super power in the world. English, as a national language of the United States, has thus been valued and further transformed to an international language. As the traumatized countries (e.g., Japan, Great Britain, the United States, China, and countries in Europe) began to rebuild in nation construction and obtain material resources to revitalize their economies, international business and commerce have prospered. Globalization appeared as well after WWII and brought about the rapid growth and expansion of science and technology. From the 1990s, many international students committed themselves to studying abroad in English-medium countries, leading to an increased need for English for specific purposes in their professions. With these historical and socio-cultural factors, English for Specific Purposes (ESP) has been highlighted in English language teaching (ELT) contexts.

English as a language of science and technology emerged largely due to the influence of the United States (which has dominated the field for much of last century and bulk of this century). In this case, English for Medical Purposes (EMP) has developed as a sub-category of ESP and received attention because most medical and health-related research stemmed from English-speaking regions. It has gradually formed its own specific linguistic and rhetorical system to guide target discourse communities in medical English teaching and

learning. As the study of ESP discourse has transformed itself from a pure linguistic to a linguistic-rhetorical approach, research in English for Medical Purposes (EMP) has been drawn more of rhetorical and register analysis since the 1980s. Medical research articles (RAs), as a remarkable sub-register, have been perceived as representative of medical English and given more attention.

In China's EFL (English as a foreign language) context, medical English is mainly used for publishing research articles in international medical journals and presenting at international medical conferences (Chia et al., 1999). The ability to read and write medical RAs is therefore a fundamental language skill for most Chinese learners of English for Medical Purposes (e.g., college medical students and Chinese doctors). Research in determining specific linguistic characteristics of medical English will benefit Chinese undergraduate medical students and doctors in their efforts to learn English in the field of medicine. Medical English can be conceptualized as a genre of ESP (English for Specific Purposes) and even as a sub-variety of EAP (English for Academic Purposes) since medical science is also a part of the academic research community (to be discussed below); however, linguistically constituting the concept remains an area of empirical research. "Medical English" is recognized as a concentration of study as part of the English major at some universities in China while in other colleges, medical students just learn generic English (called "Basic English") due to a lack of a clear operational definition of "medical English", although medical English as a concept of ESP is not disputed. When colleges and universities purport to teach "medical English" to their students in the field of medicine, no curricula is recognized and instructors have no guidance to ensure fidelity of implementation. Under this

backdrop, this study examines the linguistic features of medical English in an effort to clearly delineate medical English for curricular and instructional purposes. Additionally, the linguistic features characterizing medical English in medical research articles (RAs) are explored. The research questions are described as below:

1. What lexical features of medical English are characteristically deployed in medical RAs that instructors of English should attend to when teaching English to medical students?
2. What syntactic features of medical English are characteristically deployed in medical RAs that instructors of English should attend to when teaching English to medical students?

To answer these questions, this research aims to examine four specific linguistic features of medical RAs, which are: 1) the frequency of occurrence of specialized medical words, 2) the frequency of occurrence of nominalized words, 3) proportions of voices, and 4) sentence complexity. The ultimate goal of the research is to clarify the features that distinguish medical RA as a register or genre (there is no distinct difference between register and genre in the linguistic field) that can be studied in its own right by undergraduate medical students as a legitimate part of the curriculum in their professional training.

To examine the linguistic features described above, this research requires collection of written data from medical journals. Corpus linguistic is collected from fifteen English-medium medical research articles (RAs) selected from five criteria-meeting medical journals published in 2017. All textual samples selected for the corpus followed the Introduction-Method-Results-Discussion (IMRD) structure, with a length ranging from 3000

to 6000 ± 500 words. It is assumed that medical English is characterized as a legitimate register of English with the linguistic features of proportional specialized medical vocabulary, nominalization, passivization, and high sentence complexity. As a representative register in the medical field, medical RAs carry the rhetorical functions of formality, technicality, objectivity, descriptiveness, and logicity. In addition to Salager's (1983) finding of three broad linguistic features of fundamental medical English (i.e., word-compounding process, affixation process, doublet phenomenon), the general linguistic features of medical English for curricular and instructional purposes are operationally conceptualized.

One major significant contribution of the research is pedagogical instruction. Findings from the research will give English instructors concrete ideas to focus on when teaching medical English classes, and guidance to design curricula and tests, not only to teach medical English and assess students' English language ability for academic achievement, but also to assist medical students in preparing for medical conferences and research article publications. This research can provide guidance for teachers when they teach grammar with college medical students in mind (e.g., they may pay more attention to the teaching of different types of sentences, clause types in them, and discourse functions of voice). In addition, some medical-associated teaching strategies are discussed such as the methods used to teach specialized medical vocabulary or even to use in testing. Some practical projects, such as attending a medical conference and writing a reflective journal, are also recommended in curriculum design to promote English use in the medical setting. The review of literature is represented in Chapter 2, followed by the discussion of research methodology in Chapter 3 and results in Chapter 4. Based on the findings, Chapter 5 discusses pedagogical implications.

Chapter 2 Review of Literature

To construct a clear concept of medical English in the ESP/EAP field, and to further establish the concept of English for Medical Purposes (EMP), the history and development of English for Specific Purposes (ESP) are examined in this chapter, in order to present a comprehensive knowledge of what ESP entails and how ESP has developed over time as a broad register. Previous research on medical English as a sub-category of ESP from lexical, syntactic and rhetorical perspectives is further investigated.

2.1 Development of English for Specific Purposes (ESP)

Background introduction

English for Specific Purposes (ESP) has emerged and been growing as an independent discipline in the field of English Language Teaching (ELT) since the 1960s. The emergence of ESP has been influenced by multiple different incidents such as the end of WWII, the rapid growth of global economies, and internationalization (Rahman, 2015). The United States came out as a big winner in the post-war period, dominating the world stage politically, militarily, economically, and certainly also in science and technology. In addition, it is learners' need for specific purposes of English that strengthened the ESP teaching movement. Hutchinson and Waters (1987) state that the outcomes in the ESP context appear when a large number of people in the world intend to learn English language for the fields of science, technology, and commerce. Students in other countries aspire to go to the United States to

study, and former British colonies continue to use English as the official language because of its accessibility to trade and commerce.

Hutchinson and Waters (1987) describe ESP as a learner-centered approach, which means it “does not involve a particular kind of language, teaching material, or methodology” (Rahman, 2015, p. 25), instead, the basic needs, learners, the required language and the learning context. Robinson (1991) also suggests the importance of needs analysis in ESP. She defines ESP criteria as “normally ‘goal-directed’ and that ESP courses develop from a needs analysis, which aims to specify as closely as possible what exactly it is that students have to do through the medium of English” (p. 3). Different from Hutchinson and Waters’ point of view, Robinson (1991) depicts ESP as an enterprise with education, training and practice. She characterizes ESP courses as those in which people receive specialized knowledge and achieve objectives in homogeneous classes for the purpose of work or specialist areas of interest. Moreover, Dudley-Evans and St John (1998) use absolute and variable characteristics to define ESP in the following way as Rahman (2015) described them below:

Absolute characteristics:

1. ESP is designed to meet specific needs of the learner;
2. ESP makes use of the underlying methodology and activities of the disciplines it serves;
3. ESP is centered on the language (grammar, lexis, register), skills, discourse and genres appropriate to those activities.

Variable characteristics:

1. ESP may be related to or designed for specific disciplines;
2. ESP may use, in specific teaching situations, a different methodology from that of ‘General English’;
3. ESP is likely to be designed for adult learners; either at a tertiary level institution or in a professional work situation. It could, however, be used for learners at secondary school level;
4. ESP is generally designed for intermediate or advanced students. Most ESP courses assume basic knowledge of the language system, but it can be used with beginners.

(p. 25)

Based on the description that Dudley-Evans and St John provide, individuals must infer in order to distinguish what belongs to the ESP category and what does not. However, the boundary is not easily delineated, considering language is a social construct utilized in different target language use domains.

ESP has been divided into two main branches: English for Academic Purposes (EAP) and English for Occupational Purposes (EOP). EAP refers to English language teaching (ELT) with the purpose of academic achievement such as English for Science and Technology (EST), English for Medical Purposes (EMP) and English for Legal Purposes (ELP), whereas EOP mainly involves work-related needs and training (Robinson, 1991). According to Hutchinson and Waters (1987), EOP is also known as English for Vocational Purposes (EVP) and Vocational English as a Second Language (VESL). There is an overlap between EAP and EOP when individuals apply what they are taught in academic contexts to the workplace after they graduate from school. For instance, although medical English is taught in undergraduate English classes for academic achievement in research publication and conference presentation, it could focus on doctor training and practice in the institutional environment (i.e., hospital) as well.

Development

As Hutchinson and Waters (1987) traced the early origins of English for Specific Purposes (ESP) from the end of WWII, the distinct ESP discipline has emerged with the development of world economy. During the 1960s, the register analysis of scientific and technical writing first spurred the movement. It focused on the technical and sub-technical vocabulary (i.e., academic vocabulary), language form, and other linguistic properties, yet

had “little explanation about why and how the sentences were formed and combined as they were” (Ramirez, 2015, p. 380). It dwelled on the study of language in specific registers; however, the rhetorical functions, skill-based approaches, and communicative purposes were less considered.

Later, rhetorical and discourse analysis were intended to make up the shortage. In the 1960s and 1970s, researchers’ emphasis on ESP shifted from conventional grammar and lexis to rhetorical studies. Mauranen (1993) states that “the study of rhetoric has been rediscovered not only as a means of improving efficiency in verbal presentation, but as an analytical tool that can be used by different disciplines uncovering certain aspects of discourse” (p. 20). Dudley-Evans (2001) commented as well, that rhetorical analysis “introduced the idea of relating language form to language use, making use the main criterion for the selection of ESP teaching materials” (p. 22). Rhetorical analysis works as a helper not only to assess learners’ needs, but also to produce practical teaching materials in the ESP field. Hutchinson and Waters (1987) also described rhetorical analysis as “a logical development of the functional/notional view of language which had shown that there is more to meaning than just the word in the sentence” (p. 33). This new movement was more concerned with rhetorical functions and the context of language use than with language form. Contrastive rhetorical studies and other significant studies have been carried out in the late 1970s and early 1980s. Nonetheless, the skills focus was not given much attention until the end of the 1970s.

With the development of skill-based courses, ESP has led to a revival of needs analysis. The term was first mentioned in the 1920s when Michael West explained what learners should do with the foreign language in the target context and how they could acquire the

target language in learning (as cited in Rahman, 2015). After ESP has been put forward at the Makerere Conference (Commonwealth Education Committee in 1961), as West (1994) states, needs analysis has again aroused general concern. Dudley-Evans and St John (1998) define needs analysis as “the process of establishing the *what* and *how* of a course,” and argue that “needs analysis is neither unique to language teaching—needs assessment, for example, is the basis of training programs and aid-development programs—nor, within language training, is it unique to LSP (Language for Special Purposes) and thus to ESP” (p. 126). They also claim three aims of needs analysis, that “to know learners as people, as language users and as language learners...to know how language learning and skills learning can be maximized for a given learner group...to know the target situations and learning environment so that data can appropriately be interpreted” (Dudley-Evans & St John, 1998, p. 126). Needs analysis is essential before language teaching in an ESP context is well acknowledged, as it is an aid to find out learners’ needs in the specific area and to develop appropriate language courses and teaching materials as a basis. Three fundamental components of needs analysis are suggested by many ESP scholars to investigate particular issues. That is, Target Situation Analysis (TSA), Learning Situation Analysis (LSA), and Present Situation Analysis (PSA). The main instruments for needs assessment are: questionnaires, interviews, observation, authentic text analysis, and discussions (Rahman, 2015).

Between the 1970s and the 1980s, two controversies arose to shape the progress of ESP studies. The first controversy centered on two different views of the scope of ESP content. The “wide-angle approach” argued that topics of English language should be taught beyond students’ specialty areas, while the “narrow approach” believed that the content should focus

on the students' specific areas of development. The second controversy encompassed the skill specificity. Some language specialists argued the usefulness of mono-skill in ESP teaching, especially in reading, but it might be thought that "concentration on one skill is limiting" (Johns & Dudley-Evans, 1991, p. 305), and multi-skill study would enhance learning qualities.

Another influential stage of ESP study from the 1980s has been genre analysis (i.e., there is no distinct difference between genre and register in the linguistic context). There are three distinct approaches to genre study: the Sydney School, the ESP School, and the New Rhetoric School (Cheng, 2011). They are marked in the differences of definitions, research focuses and the targeted audiences. Among these, genre studies in ESP research have received attention with the notable feature of linguistic-rhetorical relationship. For instance, Johns (2003) argues that ESP "is becoming increasingly context-driven, and the overlap between the New Rhetoric...and ESP research and theory...becomes greater every year" (p. 206). Genre is often influenced by social actions, communities of practice and other features. Reviewing Swales' work, Master (1992) proposes genre as "a class of communicative events which has a shared set of communicative purposes that are recognized by the parent discourse community...established constraints on contributions in terms of their content, positioning and form...and nomenclature for genres that is determined by the discourse community" (p. 287). That is, genres are viewed as schemas with particular linguistic features and rhetorical structure (which is also called rhetorical moves) for communication purpose and social action. Swales (1990) also claims that genres are the properties of discourse communities (defined as a group of individuals who share common public goals and specific genres). The analysis of

genre study aims to illustrate the salient textual features (i.e., linguistic features), to understand how contexts are set up (i.e., rhetorical functions), and to play a role as a means of communication in social environments. According to Rizzo and Llopis (2011), genres can be generally classified into four different groups. Workplace genres deal with the issues of everyday communications in the professional area; economic and legal genres encompass the fields of law and economics; media genres investigate the scope of social media and the written press; and academic and scientific genres include the realms of academic teaching and research. In the academic genre category, comparative genre analysis has been employed as a new but relatively untested strand of research. Its aim is to identify the distinct aspects of rhetorical moves and functions across different disciplines. As Basturkmen (2014) states, comparative genre studies (e.g., the comparison of rhetorical functions between legal English and tourism English) probe cross-disciplinary variation by means of quantitative research. As established by Swales in his work *Genre Analysis: English in Academic and Research Settings*, research articles exhibit three rhetorical moves (i.e., establish a territory, establish a niche, and occupy a niche). Based on this, Devitt (2015) proposes rhetorical-linguistic genre studies which combine linguistic patterns with rhetorical functions. He also refers to the concept of genre performance, that is, each genre manifests its own qualities such as historical occurrences, identities, degree of prototypicality, disciplinary membership, etc. As a consequence, the actual performance has been executed through specific utterances and texts, which can be distinct but still share a genre (e.g., both an argumentation of realism and a bill proposal belong to academic English genre). With the existence of genre performance, the term “genre competence” has also been presented as the inner variable (i.e., abstract

traits/awareness). Devitt also puts forward a three-column analysis of three genre pedagogies, which include the explicit teaching of particular genres (e.g., teaching how to write medical research papers), teaching genre antecedents (i.e., choosing and using a genre for the skills and strategies), and teaching genre awareness (i.e., the consciousness of genre analysis and critique). Although linguistic-rhetorical genre analysis is limited by inadequate knowledge of communities in specific genres, genre-based approach has been well developed and commonly used in ESP research. Therefore, the current author partially adopts linguistic-rhetorical approach to analyze four specific linguistic features of medical English to better delineate the genre of medical research article (RA).

Since English for Specific Purposes (ESP) originated with register analysis, corpus linguistics establishes functional registers on the basis of linguistic features used in each (Biber, 1988, 1992). According to Halliday (1988), register is “a cluster of associated features having a greater-than-random (or rather, greater than predicted by their unconditioned probabilities) tendency to co-occur” (p. 162). For instance, as a variety of ESP, medical English has been described as a register relating to medical science with its own linguistic features and rhetorical schemata for either academic or institutional (occupational) purposes.

2.2 Previous research on medical English in the EMP field

Lexical features

In the ESP context, the study of vocabulary in scientific/technical registers contributes to the quality of reading and understanding specialized texts. What Cowan (1974) calls *sub-technical words* (e.g., *investigate, analyze, demonstrate*) and Martin (1976) *academic*

vocabulary (e.g., *argue, research, found*) has aroused ESP specialists' interests. These words are used across scientific disciplines but less frequently in general textual counts. Farrell (1990) views this category as "formal, context-independent words with a high-frequency and/or wide range of occurrence across scientific disciplines, not usually found in basic general English courses; words with high-frequency across scientific disciplines" (p. 11). His study of sub-technical vocabulary reveals the fact that acquiring only technical terminology (i.e., specific jargons used only in particular fields such as the word *palpitation* in the medical field) is inadequate to understand technical texts. Baker (1988) describes the divergence of technical vocabulary and sub-technical/academic vocabulary as elusive and confusing because some specialized jargons are also used in daily life. It is also ambiguous when classifying words that are neither specialized nor generally used in everyday language. To better clarify the scope of sub-technical vocabulary, Coxhead developed an academic word list (AWL) in 1998 and revised it in 2000. Later, Nation (2001) divided words in academic writing into four main categories: high-frequency words, low-frequency words, technical words and academic words (also called sub-technical words). High-frequency words relate to generic English words occurring mostly in colloquial conversation and all running words in all types of writing. Low-frequency words are rarely used. Technical words are used in technical fields (e.g., medical terminology in the field of medical science), while academic words refer to those that constitute a high proportion of running words in all academic texts.

With regard to the technical vocabulary in the medical English register, the earliest medical terminology was written in Greek by Hippocrates. Although English becomes the leading language in the medical field, Greek, Latin, German, old French, and old English

have all left traces in medical terminology with the historical development of English language. From the fifteenth century, when Spain discovered and colonized the New World, traces of Spanish can also be found in the field of medical science (Ariza & Maria, 2005). In Ariza and Maria's (2005) research, two distinguished types of borrowings from Spanish in that field have been marked. One is the borrowing of proper names of Spanish anatomists, physicians, and surgeons, and that the other is the expressions of geography related to Spanish territory. Nevertheless, the specificity of the eponymic expressions is questioned by the current author, as no matter where people have been, they are able to name diseases or any other medical research by people's names or geography in their languages. For instance, Eisenmenger's syndrome was named by Dr. Paul Wood and after Dr. Victor Eisenmenger, who first described the condition in 1897. Therefore, counting this as a unique feature of medical terminology seems inappropriate.

Salager (1983) conducted a research on identifying the lexis of fundamental medical English (FME), in which he classified words into three categories based on word roots as follows:

“(1) Basic English (BE) defined as that part of the lexicon common to all kinds of writing, whether it be scientific, technical or literary; (2) fundamental medical English (FME)—the object of the present study—made up of those roots whose frequency of occurrence is homogeneously distributed over the whole ME corpus but in a significantly different way than in BE; and (3) specialized medical English (SME) composed of specialized terms particular to a given discipline and known only to the small group of workers in the field.”

(p. 56-57)

He determined to investigate the proportions of word roots distributed in these three categories and to obtain the core lexis of FME. However, the selected roots fell into the FME lexicon also contained words that are less relevant to medical English (e.g., *range*, *degree*, *to*

confirm). The corpus of his research included textbooks, journals, and reviews, which is different from the scope investigated in the current study (i.e., medical RAs). Moreover, Salager's research is distant from today's variation of medical vocabulary and is unable to keep pace with the rapid expansion of medical science.

Other remarkable research in the linguistic evolution of medical English has received attention. Salvager (1983) drew an illustration of three broad linguistic features of fundamental medical English as word-compounding processes (e.g., *weight loss*, *patient monitoring*, *heart failure*), affixation process (e.g., *relate*, *relationship*, *unrelated*), and doublet phenomenon in adjective vs. noun roots (e.g., *kidney function* vs. *renal function*, *heart failure* vs. *cardiac failure*). Based on the previous research, Dzuganova (2006) furthered the study of word affixation in the medical field. Since clinical branches of medicine deal with a huge number of signs, symptoms, defects, diseases, and all have to be named, the majority of these names take explicitly negative meaning expressed by means of negative affixes. Dzuganova (2006) called such type of negation "affixal negation." Among these, there are five basic negative prefixes (*a-*, *dis-*, *in-*, *non-*, *un-*) and only one negative suffix (*-less*). Other compound or combining affixes such as *anti-*, *counter-*, and *extra-* often occur in naming units that have a negative, reverse or antonymic meaning. It is well acknowledged that etymology plays an important part in medical vocabulary to trace the original source of words or partial affixes in order to investigate the lexical features of medical vocabulary. Dzuganova then published another article about partially negative prefixes in medical English in 2007. By analyzing the etymological sources and the meaning of these prefixes in both medical vocabulary and words in everyday use, none of the

following prefixes or combining words can be considered as negative: *anti-*, *contra-*, *counter-*, *de-*, *dys-*, *ex-*, *extra-*, and *mal-* can be considered as negative ones. These prefixes can only partially express some negation in a word they combine with (e.g., the word *antibody* conveys a positive meaning as a substance produced by one's body to fight disease). The affixal negation Dzuganova (2006) introduced before cannot be viewed as a distinct characteristic to differentiate medical terminology and ordinary words. Although he gave some hints to guess the meaning of medical vocabulary, there are still some affixes that have other connotations because of the original source. Nonetheless, his analysis continues to benefit the understanding of medical vocabulary.

Other than specialized terminology, some ESP researchers devote their attention to sub-technical/academic vocabulary. Chen and Ge's (2007) study of frequency and distribution of academic words in medical research articles further serves to verify the medical English register in the ESP/EAP field. The authors conducted a lexical study of the word frequency and the text coverage of 570 word families from Coxhead's Academic Word List (AWL) in medical research articles (RAs) based on a corpus of 50 medical RAs written in English with 190,425 running words. The findings show that the percentage of text coverage of the AWL word families in medical RAs is 10.073%, which corresponds to the AWL text coverage Coxhead and Nation (2001) found in cross-discipline academic writing (around 10%). The result repeatedly shows medical English as a sub-category of English for Academic Purposes (EAP), and it is further developed as English for Medical Purposes (EMP). The study also shows that some but not all academic words (292 out of the 570 AWL words) are distributed in medical RAs with high frequency. The significance of Chen and

Ge's study is to confirm that there is an overlap between academic vocabulary and vocabulary in the medical English register. It is supporting evidence that medical English is a sub-genre of ESP/EAP. On the other hand, the percentage of specialized medical vocabulary in medical RAs has not been examined. It could be an area for further exploration in lexical study. Wang, Liang, and Ge (2008) established a Medical Academic Word List (MAWL) with the coverage of 12.24% in the medical RAs. Later in 2016, Lei and Liu reestablished a new Medical Academic Word List (MAWL) with enhanced methodology (i.e., the improvement of computer software) and larger corpora.

Syntactic features and rhetorical analysis

With the advocacy for rhetorical and discourse analysis in the 1970s and 1980s, and the following genre analysis in late 1980s, linguistic features in ESP contexts such as tense, voice, and modality have been interpreted with functional meanings. The transition from pure linguistic/grammatical analysis to linguistic/grammatico-rhetorical studies has dominated ESP discourse. Genre analysis has thus flourished in the study of English for Medical Purposes. As Ard (1983) traces the origins of research articles (RAs), this “prestigious genre,” as Swales (2014) called RA, came out of letters between scientists exchanging information, with the first RA published in the first English-medium scientific journal: *The Philosophical Transactions of the Royal Society* of London in 1665. With this in mind, Swales (1990) established the Create a Research Space (CARS) model with three principal moves (i.e., establishing a territory, establishing a niche, and occupying a niche) to scaffold the structure of research articles on the foundation of four primary components (i.e., Introduction, Methods and materials, Results, and Discussion or IMRD). Although most medical practitioners follow

the CARS model and the traditional IMRD sections, they still encounter difficulties in highly technical and standard research paper writing. Adams-Smith (1983) exemplified some instances of poor medical report writing and summarized the main characteristics as excessive use of jargon, careless phrasing, poor flow of ideas and self-centered styles. In order to help medical professionals tackle the problem of ill-conceived research article writing, Gosden (1992, 1993) investigated the discourse functions of theme in research articles. Skelton (1994) furthered the research on the structure of research papers for the audience of general medical practitioners. Later, Nwogu (1997) conducted a study to describe scientific moves in research papers for linguistic specialists. Eleven moves were found in a conventional medical research paper as described below:

In a typical medical research paper, the author(s) is likely to:

1. Begin by providing some background details about the subject matter of the research investigation.
2. Conduct a brief review of previous studies.
3. Introduce the new research and state research objectives.
4. Identify the source of data and the method adopted in collecting them.
5. Discuss the methods adopted in the analysis of data.
6. Discuss the statistical procedures adopted in the analysis of data, if any.
7. Discuss results which are consistent with the objectives of the study.
8. Discuss results which deviate from expected outcomes.
9. Discuss the main results and their significance.
10. Explain and justify results and approaches adopted in achieving them.
11. Provide a conclusion and indicate implications of the research.

(Nwogu, 1997, p. 124)

Nwogu (1997) also classified these moves into two categories: obligatory moves as these moves occurred in all text samples; and optional moves, which occurred less frequently in texts. With the eleven moves and their discourse functions listed above, eight (move 2, 3, 4, 5, 7, 9, 10 and 11) are described as normally required and three (move 1, 6, and 8) are optional. Based on Nwogu's (1997) research, Li and Ge (2009) investigated the linguistic and

structural changes in medical research articles (RAs) in two decades from 1985 to 2004. According to the two corpora they collected, the status of some moves in the conventional structure change between the two categories. Move 1 and 6 changes from optional to obligatory. Move 8 tends to be avoided in today's medical RAs because of the difficulty of paper publication with negative results. Move 9 is likewise avoided in medical RA writing, which indicates that medical practitioners now use induction more than deduction to write the discussion section. With regard to the linguistic features analysis, the result shows the top three verb tenses used in medical RAs as the simple past tense, the simple present tense, and the present perfect tense. It also reveals that medical RAs no longer have singular first person pronouns, which may help to connect the relationship between researchers and readers. From another perspective, it indicates the effect of co-authors, that is, more than one researcher is dedicated to medical RA's writing and publication. Since Li and Ge's findings confirm some linguistic and rhetorical changes of medical RAs by analyzing the use of verb tense and first-person pronoun, these linguistic features (verb tense and first-person pronoun) are not included in my linguistic analysis of medical RAs. Nevertheless, Li and Ge's study provide evidence of some specific linguistic features and the implied rhetorical functions of medical English as supplementary.

From a systemic functional perspective, Yang, Zheng, and Ge (2015) examined the frequency and distribution of words that convey modality (e.g., *can*, *should*, *possible*) and interpreted the possible functions and the values (i.e., the degree of certainty) of epistemic modality in medical RAs. Lyons (1977) states that, "epistemic modality is concerned with matters of knowledge, belief, or opinion rather than fact" (p. 793). For instance, "it is *possible*

that..." indicates the writer's subjective speculation but not an actual fact. Halliday (2004) defines it as the degree of possibilities to propositions. In Yang et al.'s research, the corpus shows that low value (e.g., "He *may* be ill.") and median value (e.g., "They *should* be back by now.") are more often used in the construction of new knowledge because of the provisional nature of the outcomes being discussed. That low value and median value are common in medical RAs indicates that medical RA writers prefer to express uncertainty rather than certainty about their stated propositions. They are aware of the nature of medical science; that medical research will not always be precise and reliable. Also, Yang et al. interpret that the use of epistemic modality manifests some politeness strategies through which addressers can establish a relationship with their addressees and successfully communicate with them (i.e., showing modesty and humbleness). It often creates a sense of hedging for negotiation of appropriate representation in medical RAs (Salager-Meyer, 1994). In addition, epistemic modal expressions with implicitly subjective/objective or explicitly objective orientations are frequently utilized to persuade readers in order to avoid absolute or subjective statements or claims (e.g., "The results were similar for the population that could be evaluated"). Since Yang et al. have already examined the use of modal words in medical RAs and the corresponding rhetorical functions based on Halliday's explanation of epistemic modality, the current research does not contain this linguistic analysis of medical English through systemic functional perspective.

Overall, the genre of English-medium medical RAs is not linguistically homogeneous. Instead, the linguistic features contribute to rhetorical functions with the communicative purposes for social actions. Halliday (1988) discusses two prominent features of scientific

discourse. One is nominalization as “a nominalized world with high density and packed grammatical metaphors” (p. 149). Another is relational intensive processes with the pattern of *process* (nominal group)—*relation* (verbal group)—*process* (nominal group) (i.e., *doing A causes doing B*). He does not treat nominalization as a lexical or morphological feature, but implies it as a process or phenomenon with particular systemic functions. However, in the following research, nominalization is viewed as a lexical feature of medical English (i.e., nominalized forms). On the other hand, Salager-Meyer et al. (1989) are concerned with the principal component analysis and medical English discourse. They examine voice and verb tense, modals, and gerunds, and make a statement of research papers’ enquiry style, as this genre tends to ask questions, come up with hypotheses, conduct experiments and explain causes and possible answers with the frequent use of simple past tense, passive voice and *-ing* form in compound words (e.g., *energy-sparing neuroendocrine response*, *cholecystokinin-containing neurons*) to generalize and report.

Use of medical English in real life

With regard to the use of medical English in real life, the needs for medical English teaching and learning vary across the discourse communities in the ESP/EMP field. As medical English is both utilized in the academic discipline for education and applied to either doctor-patient diagnosis or doctor-doctor communication in the hospital, the medical English register can draw on both the English for Academic Purposes (EAP) and English for Occupational Purposes (EOP) categories. In the EAP context, medical English is mainly used for publishing research articles in international medical journals and presenting at international medical conferences. Maher (1986) stated that 72.2% of the worldwide medical

journals and magazines were written in English in 1980. Also, 372 out of 373 international conferences in the Americas, Europe, Asia, the Middle East, and Australia set English as one of the official languages. Although this inquiry is antiquated, it demonstrates that the English language, as a means of communication, plays an essential role in the medical field. In the institutional environment (i.e., hospital), medical terminology in English is a powerful instrument in doctor-patient communication (especially in an environment where English is a foreign language) when doctors frequently use complicated medical jargons with their patients to demonstrate that they are expert and professional who are well trained in the medical field, insofar patients should adopt doctors' suggestions and follow the instruction. Driven by needs analysis, Fang (1987) conducted a study of English language education in medical programs. It demonstrated that although students realized the importance of English studies in the medical field, and were highly motivated to improve English skills, their expectations and needs to facilitate their English language competence were not satisfied. On the other hand, Chia et al. (1999) carried out a study to examine the perceptions of the English language needs of college medical students and faculty in Taiwan. In 1993, the Ministry of Education had abolished English as a mandatory requirement in college. English programs, course design, and content focus differed between colleges and universities. Under these circumstances, the researchers picked Chung Shan Medical College in Taichung, Taiwan, as the object of study and made two questionnaires for both medical students and faculty in the Department of Medicine, in order to identify the perceptions of English language needs for both students and faculty and to suggest requisites for English language curricula. The results showed that most students and faculty believed English was important,

and that the faculty felt English was more important for the students than the students did. In terms of the extent to which English was used in class, 45% of the faculty indicated that 70% to 89% of the textbooks were English-written, yet 42% of them did not specify the language for students' reading assignments. The study also reported that most students considered reading English newspapers and magazines, having daily communication, and writing reports and research papers in English as their primary improvement strategies. With regard to the four Basic English language skills within freshmen English courses, the listening skill was the most in need of improvement by freshmen and writing skill the least needed. The skills ranking changed, however, in the perceived English language skill needed in medical studies. The majority of students believed reading to understand textbooks and medical research articles was most important, followed by listening, writing, and speaking skills. Regarding curriculum design, both students and faculty groups believed the best approach was to begin with Basic English, including some relevant materials in the medical field in the first college year, then specific elective English learning between sophomore and senior years. The English learning problems that students rated highest were limited vocabulary, slow reading speed, and unsatisfying listening comprehension, whereas inadequate grammatical knowledge was rated lowest.

In China's EFL (English as a foreign language) context, the study of medical English would benefit Chinese doctors and medical students in their effort to learn English in the field of medicine. However, there is not a unified teaching system to design and implement medical English curricula. Most Chinese universities do not recognize "medical English" as distinct from "basic English." While it is relatively straightforward to conceptualize medical

English as a genre of ESP and a sub-variety of EAP since medical science is also a part of the academic research community, what linguistically constitutes the concept of medical English for curricular and instructional purposes remains an area of empirical research. Based on the previous research, this study examines four specific linguistic features of medical research articles (RAs) that distinguish medical RAs as a register which can be studied by undergraduate medical students as part of their professional training.

Chapter 3 Methodology

This research adopts a corpus-based approach and requires collection of written data from medical journals to examine four specific lexical and syntactic features of medical English research articles (RAs). Linguistic-rhetorical analysis of these linguistic features is discussed in Chapter 5. Subsequent analysis of the corpus data is described as below.

3.1 Data collection

To answer the research questions about lexical and syntactic features of medical English, corpus linguistic data were collected from fifteen English-medium medical research articles (RAs) selected from five medical journals published in 2017. All textual samples selected for the corpus followed the Introduction-Method-Results-Discussion (IMRD) structure, with a length ranging from 3000 to 6000 \pm 500 words and a mean length of 4020 words.

Specific criteria for sample article selection should be followed, i.e., a) articles must be published in the United States or the United Kingdom; and b) articles should be selected from top journals with relatively high academic reputation and impact factors in the medical discipline. Therefore, target journals were selected from the Science Journal Rankings website (<http://www.scimagojr.com/journalrank.php>) with the impact factor above 15.000 and SJR index above 5.000. As a result, five criteria-meeting medical journals were confirmed as the following: *Lancet*, *New England Journal of Medicine*,

Journal of American Medical Association, Journal of Clinical Investigation, and British Medical Journal.

In order to obtain eligible data, Ge's (2015) collection method is adopted and a two-round stratified sampling is conducted to ascertain the representativeness and objectivity of the sample RAs (i.e., all candidate medical RAs have equal opportunities to be selected). In the first round, twenty criteria-meeting RAs were randomly selected from each journal (i.e., choosing two sample RAs from each medical category as described below). A total of 100 samples were thus randomly selected from all the issues of the five journals published in 2017. In the second round, 15 sample RAs were randomly selected from the 100 first-round samples (3 RAs from each journal). Therefore, the corpus consisted of fifteen English-medium medical RAs with a total of 60,297 running words classified into the following ten medical specialties:

-
- | | |
|---------------------|---------------|
| 1. Immunology | 2. Oncology |
| 3. Cardiology | 4. Gynecology |
| 5. Dermatology | 6. Pediatrics |
| 7. Endocrinology | 8. Surgery |
| 9. Gastroenterology | 10. Neurology |
-

3.2 Data analysis

The target data were collected manually, some of which were processed with the assistance of computer software *Activeperl*. Four specific lexical and syntactic features that are investigated get described below.

Frequency of occurrence of specialized medical vocabulary

The data (i.e., frequency of occurrence of specialized medical vocabulary in sample RAs) were used to determine the extent to which medical words occur in medical RAs. In order to investigate this lexical profile, the author collected words that relate to medical registers and are frequently used in the medical field with a general acknowledgement of medical words in the medical setting. Since Chen and Ge (2007) confirm medical English as part of EAP, and it is troublesome to conceptualize medical vocabulary separately as well, as the boundary of medical terminology, academic words, and high-frequent words blurs with ambiguity. The author collected words that are semantically medically-associated and are frequently utilized in the medical setting. For instance, the word *pregnancy* is counted as a medical word even if it is commonly used in daily life. Moreover, word-compounding process should be taken into consideration when collecting medical words. For instance, the compound words *liquid biopsy*, known as a medical term (i.e., the analysis of circulating and derived from cancer cells), should all be counted as medical words even though the word *liquid* is not medical-related in common use. Another example is the term *Attention-Deficit/Hyperactivity Disorder (ADHD)*. Neither *attention* nor *deficit* belongs to the medical category, but *Attention-Deficit/Hyperactivity Disorder (ADHD)* occurs as a medical

jargon to characterize a common neurodevelopmental disorder in children and adolescents. In this case, these words are counted separately (because the denominator of the frequency calculation is the number of all running words).

Nominalization

The data (i.e., frequency of occurrence of the nominalized forms ending with *-tion/-sion*, *-ment*, *-ness*, and *-ity*) were used to examine the degree of nominalized forms in medical RAs. According to Biber et al. (1998), nominalization is defined as “nouns that are related to verbs or adjectives morphologically” (p.59). In order to investigate this lexical feature, the author identified nouns that resulted from nominalizing affixes on words of other parts of speech (i.e., verbal and adjectival category) such as words that end with such suffixes as *-tion/-sion*, *-ment*, *-ness*, and *-ity* (and their plural forms). This collection method is based on Biber’s (1998) data processing as “a concordancing package can be used to identify nominalizations ending with these suffixes with a high degree of accuracy” (p. 59). Computer software *Activeperl* was used to search for words ending with *-tion/-sion*, *-ment*, *-ness*, and *-ity* (and their plural forms). Hand-editing is necessary after computer processing because automatic analysis would include some words that are not nominalizations (e.g., *element*, *section*).

Passivization

The data (i.e., frequency of each voice and percentages of the use of voice) were used to determine the extent to which passive voice occurs in medical RAs. In order to investigate this syntactic feature, the author identified and documented frequency of occurrence of each

voice and the percentages of passive, active, and “null” voice (linking verbs and *there be* structure) in sample RAs.

Sentence complexity

The data (i.e., the ratio of clause to sentence) were used to determine the extent to which complex sentences are used in medical RAs. A complex sentence is defined as any T-unit that consists of one or more embedded subordinate, complement, and relative clauses. Coordinate clauses are exempt as they do not constitute a complex sentence in conventional grammatical description. Two examples of complex sentences are shown as following:

(1) | The primary null hypothesis was | that clindamycin, TMP-SMX, and placebo **would have** equivalent rates of cure at the test-of-cure visit. | (sample RA 11)

(2) | Improvement was also observed in one patient | who **received** steroid pulses (patient 17), | whereas intramuscular adrenocorticotrophic hormone therapy for epileptic spasms **was reported** as not being of value in three patients (patient 2, 11, and 14). | (sample RA 13)

3.3 Rating procedure

Since most written data were collected and sorted manually, to ensure the inter-rater reliability, the author invited a second rater to recheck five out of fifteen sample RAs for consistency in the classification of the data. The invited rater was a native English speaker who took English grammar class and is familiar with English sentence structure. After the coding of five RAs was completed, both the researcher and the second rater discussed and negotiated the differences and disagreements based on the particular grammatical rules and patterns described above for the coding task. The

index of the inter-rater reliability is 91.23%. The same five RAs were re-coded by the researcher one week after the initial coding to achieve the intra-rater reliability (the index of intra-rater reliability is 98.04%). Computer software *Activeperl* was adopted as an assistive tool, with the program designed by a computer science technician who works in the Chinese Academy of Sciences.

Chapter 4 Results

In this chapter, the results of frequency of specialized medical words and particular nominalized forms, proportions of the use of voice, and ratios of passives to actives and clause to sentence are presented and interpreted as described below.

4.1 Frequency of occurrence of specialized medical words

To norm frequency counts from texts of different lengths, the raw frequency count of specialized medical words was divided by the total running words of one text and multiplied by 1,000 to yield a frequency ratio of occurrences per thousand running words. An example is shown as follows:

Sample RA 1

$$(526 \text{ medical words} / 4534 \text{ words}) \times 1,000 = 116 \text{ medical words per } 1,000 \text{ words}$$

To make a comparison with the percentage in Chen and Ge's (2007) study in the following chapter, each percentage of specialized medical words was also listed in the parenthesis. As Table 1 shows, regardless of the low frequency of specialized medical words in sample RA 9 (53), specialized medical words were used in all other sample RAs with a frequency ranging from 127 to 210. The Standard Deviation (SD) is 39.99. The Mean is 162.73, which demonstrates that average 162.73 specialized medical words per 1,000 words (16.27%) occur in medical RAs (i.e., to make a comparison with the percentage in Chen and Ge's 2007 study in the following chapter, the percentage of average specialized medical words was provided here).

Table 1

Frequency of occurrence of specialized medical words in medical RAs (normed per 1,000 words)

Sample RA	Normed Frequency
1	116
2	196
3	166
4	155
5	208
6	172
7	210
8	192
9	53
10	159
11	164
12	181
13	170
14	172
15	127

4.2 Frequency of occurrence of nominalized words

To norm frequency counts from texts of different lengths, the raw frequency count of the selected nominalized words was divided by the total running words of one text and

multiplied by 1,000 to yield a frequency ratio of occurrences per thousand running words. An example is shown as follows:

Sample RA 3

$$(244 \text{ nominalized words} / 4185 \text{ words}) \times 1,000 = 58 \text{ nominalized words per 1,000 words}$$

As shown in Table 2, nominalized words ending with *-tion/sion*, *-ment*, *-ness*, and *-ity* (and their plural forms) were used in the fifteen sample RAs with a frequency ranging from 23 to 58. The SD is 12.24. The Mean is 41.13, which indicates that average 41.13 nominalized words with the selected endings per 1,000 words occur in medical RAs.

Table 2

Frequency of occurrence of nominalized words ending with *-tion/sion*, *-ment*, *-ness*, and *-ity* (and their plural forms) in medical RAs (normed per 1,000 words)

Sample RA	Normed Frequency
1	38
2	43
3	58
4	23
5	37
6	27
7	52
8	55
9	23
10	46
11	45

12	34
13	35
14	38
15	63

4.3 Proportions of voices

To norm frequency counts from texts of different lengths, the raw frequency count of the active/passive/“null” voice (i.e., linking verbs and *there be* structure) was divided by the total running words of one text and multiplied by 1,000 to yield a frequency ratio of occurrences per thousand running words. An example is shown as follows:

Sample RA 6

$$(125 \text{ active voice} / 4636 \text{ words}) \times 1,000 = 26.96 \text{ active voice per } 1,000 \text{ words}$$

As Table 3 shows, active voice was used in the fifteen sample RAs with a frequency ranging from 17.74 to 53.50 (per 1,000 words). The SD is 7.97. The Mean is 32.62, which presents that average 32.62 active sentences per thousand words occur in medical RAs. Passive voice contained a frequency ranging from 7.18 to 25.64 (per 1,000 words). The SD is 5.18. The Mean is 17.84, which demonstrates that average 17.84 passive sentences per thousand words occur in medical RAs. Linking verbs and *there be* structure were used with a frequency ranging from 7.76 to 20.94 (per 1,000 words). The SD is 3.79. The Mean is 14.39, which indicates that average 14.39 sentences without voice (i.e., the use of linking verbs and *there be* structure) per thousand words occur in medical RAs.

Table 3

Frequency of occurrence of active/passive/“null” voice in medical RAs (normed per 1,000 words)

Sample RA	Normed frequency of active voice	Normed frequency of passive voice	Normed frequency of “null” voice
1	41.46	15.43	17.20
2	28.63	12.64	13.12
3	34.88	7.88	12.42
4	34.05	19.06	17.56
5	38.49	21.35	7.76
6	26.96	20.70	7.76
7	28.35	18.96	14.35
8	28.46	20.64	12.42
9	53.50	7.18	17.03
10	28.60	21.29	10.48
11	17.74	21.81	20.94
12	29.53	25.64	18.50
13	33.25	20.82	15.45
14	31.46	17.85	14.88
15	34.07	16.48	16.12

Moreover, since both “frequency of occurrence” and “percentage” appear in corpus linguistics, to further ascertain the proportion of each voice in medical RAs, the percentage of each voice in texts of different lengths was also calculated. The raw number of each voice

was divided by the total number of voices (active + passive + null), and then multiplied by 100%. An example is shown as below:

Sample RA 12

$$(91 \text{ actives} / 91 \text{ actives} + 79 \text{ passives} + 57 \text{ "null" voice}) \times 100\% = 40.08\%$$

It demonstrates that of all sentences in Sample RA 12, a little over 40%—close to half—are active sentences.

The percentages of active, passive, and “null” voice (i.e., linking verbs and *there be* structure) were listed as below. As Table 4 reveals, regardless of the differences of the percentages of the voices used in different texts, active voice is most frequently used in all sample RAs (except for sample RA 11 with 29.33% of active voice). In sample RA 1, 2, 3, and 9, the percentage of “null” voice is in excess of that of passive voice, whereas in other sample RAs, the data shows passive voice outnumber “null” voice. Overall, active voice occurs most frequently in medical RAs, with the average percentage of 53.75%, followed by passive voice (average 29.20%), and then the linking verbs and *there be* structure (average 22.85%), which do not have any voice.

Table 4

Percentages of voices (active, passive, and null) in medical RAs

Sample RA	Percentage of active voice (%)	Percentage of passive voice (%)	Percentage of “null” voice (%)
1	55.95	20.83	23.21
2	52.63	23.24	24.12
3	63.20	14.28	22.51

4	48.18	26.97	24.85
5	56.93	31.57	11.48
6	48.63	37.35	14.01
7	45.97	30.74	23.28
8	46.25	33.55	20.19
9	68.83	9.25	21.92
10	47.37	35.26	17.37
11	29.33	36.05	34.62
12	40.08	34.80	25.11
13	47.83	29.95	22.23
14	49.00	27.81	23.17
15	53.75	29.20	22.85

4.4 Ratio of clause to sentence

The ratio of clause to sentence (i.e., number of clause was divided by number of sentence) is listed in Table 5 below. The result shows that all sample RAs carry the ratio over 0.00 (i.e., the ratio of clause to sentence is 0:1, which means there is no complex clause in each sentence), with the average of 0.51. It demonstrates that average 0.51 clause per sentence occurs in medical RAs, and further confirms sentence complexity in medical RAs.

Table 5

Ratio of clause to sentence

Sample RA	Ratio
-----------	-------

1	0.68
2	0.50
3	0.44
4	0.64
5	0.55
6	0.24
7	0.72
8	0.59
9	1.13
10	0.42
11	0.26
12	0.50
13	0.23
14	0.29
15	0.35

In general, the research results show that average 162.73 specialized medical words per 1,000 words occur in medical RAs,; when searching for nominalized forms with the end of *-tion/sion, -ment, -ness, and -ity*, average 41.13 nominalized words per 1,000 words occur in medical RAs; active voice occurs most frequently in medical RAs (32.63 per thousand words, with the average percentage of 53.75%), followed by passive voice (17.84 per thousand words, with the average percentage of 29.20%), and then the linking verbs and *there be*

structure (14.39 per thousand words, with the average percentage of 22.85%), which do not have any voice; and medical RAs are characterized by sentence complexity with the average ratio of 0.51 (clause to sentence). With regard to the analysis of the results, interpretations of these linguistic features and the corresponding rhetorical functions are further discussed in the next chapter.

Chapter 5 Discussion and Conclusion

In this chapter, analysis of the findings is further discussed, compared with the previous research that is associated with this topic. Pedagogical implications are provided to assist English instructors in curriculum and assessment design. Limitations of this research are also considered.

5.1 Discussion

Lexical features

The first research question, “What lexical features of medical English are characteristically deployed in medical RAs that instructors of English should attend to when teaching English to medical students?” is answered using a statistical frequency-count analysis to examine the frequency of occurrence of both specialized medical words and nominalized words ending with *-tion/sion*, *-ment*, *-ness*, and *-ity* (and their plural forms). The results reveal that average 162.73 specialized medical words per thousand words occur in medical RAs, with the average percentage of 16.27%, which surpasses the proportion of academic vocabulary (10.07%) (Chen & Ge, 2007). In the current English language teaching (ELT) context, vocabulary may not be considered as an important area to be worthy of teaching in class. However, as Elliman (1980, p. 40) proposes, reading comprehension will be interrupted with confusion and frustration when the proportion of unknown words is up to 10%. Therefore, specialized medical vocabulary remains a significant lexical feature to characterize “medical English” that instructors of English should consider.

Compared to Salager's (1983) research, this study has a different operational definition to identify specialized medical words. That is, words that are semantically associated with medical registers and that are frequently used in the medical field with a general acknowledgement of medical words in the medical setting are defined as specialized medical words (word compounding is included), which range from high-frequency words to academic words to technical words to low-frequency words based on Nation's (2001) classification of vocabulary. Since there is an overlap of word use across vocabulary categories in the medical setting (i.e., some high-frequency words are also often used in medical RAs by doctors and undergraduate medical students such as *drugs*, *pregnancy*, *lesions*, *neurodevelopmental*), it cannot be denied that these semantically medically-associated words are part of medical vocabulary. In addition, the fact is that there is not an established operational definition of medical vocabulary to regulate what is and is not included. The decision made to select specialized medical words as described in research methodology is well-founded.

With regard to the nominalized forms, the mean frequency is 41.13, which demonstrates that average 41.13 nominalized words are used with the selected endings per thousand words in medical RAs. The proportion is small because the data of nominalized words were selected with limits (i.e., ending with four specific suffixes). Other nominalizations such as conversion (zero derivation) and gerund (e.g., *monitoring*) were not considered in this study.

Nevertheless, nominalization still exists as a lexical feature in medical RAs that English teachers may need to explain when students ask.

Overall, medical RA, characterized by proportional specialized medical words and nominalization, carries a formal, academic, and professional writing register that individuals

are aware of the power of discourse. When reading a medical RA full of low-frequency words and recondite terminologies such as *bacteraemia*, *rifampicinl*, and *placebo*, individuals are likely to think that the writer must receive higher education and high-standard professional training, treating the writer as the authority of this discourse. Moreover, nominalization has the rhetorical function of making different stresses and emphases, which not only transforms the speaking tone to a more formal and somewhat solemn aspect, but also manifests the characteristic of information condensation as nominalized words “can be qualified and related to other activities in a highly condensed manner” (Lemke, 1990, p. 440). For instance, *antibiotics used at randomization* (sample RA 8), with the emphasis on the noun (*antibiotics*), delivers a different meaning from *use antibiotics randomly*, which addresses the action. The nominalized word *randomization* thus makes the expression highly condensed. It also sounds academic as *randomization* is usually seen in academic papers but not commonly used in everyday language.

Syntactic features

The second research question, “What syntactic features of medical English are characteristically deployed in medical RAs that instructors of English should attend to when teaching English to medical students?” required passivization and sentence complexity to be examined through the frequency of voices and the ratio of clause to sentence in medical RAs. The findings demonstrate that active voice occurs most frequently in medical RAs (32.63 per thousand words, with the average percentage of 53.75%), followed by passive voice (17.84 per thousand words, with the average percentage of 29.20%), and then the linking verbs and *there be* structure (14.39 per thousand words, with the average percentage of 22.85%). With

regard to the rhetorical function of passivization, passive voice is used with the emphasis on the entities being studied, while active voice mainly reflects on the researcher's actions and research processes. Linking verbs and *there be* structure to some extent manifest a kind of status or condition (e.g., "...when energy stores are low"). In this case, passivization carries the significance of making claims, illustrating facts, and describing what have already been done with the emphasis on the patient. In addition, passivization in medical RAs act as a means of objective statements and descriptions based on established facts, instead of subjective argumentation with opinions, inferences, and rebuttals.

Medical research articles are also characterized by complex sentence structure including relative, complement, and subordinate clauses, with the average ratio (clause to sentence) of 0.51. The data shows that average 0.51 clause was used per sentence in medical RAs. It also indicates that medical RAs carry a great amount of information with a precise logical framework (i.e., frequent use of loop-de-looping clauses). It further indicates that those who are dedicated to medical RA writing should be intellectuals who have both received professional training in medicine and obtained academic achievements with high English proficiency.

5.2 Pedagogical implications

This research provides theoretical support to instructors of English when they teach medical English to medical students. The research also raises some pedagogical issues such as what should be taught when instructors design the curriculum and assessment of a "medical English" class. First, specialized medical words should be given more attention by

instructors when they teach medical English in class, especially when some low-frequency medical vocabulary are frequently used in medical RAs. Vocabulary knowledge such as word formation, affixation, and word compounding are also recommended in classroom teaching. Second, it is advisable to introduce the concept of nominalization in class to raise students' awareness of the use of nominalized words and other linguistic features of medical RAs. Third, the use of grammatical voice and the corresponding implications should be targeted to explain the researcher's intentions when students are taught to read medical RAs. In addition, sentence structure, especially clause varieties, should be considered when instructors teach grammatical rules, to ensure that medical students are able to recognize complex sentence patterns and to find important information precisely when they read medical RAs.

5.3 Limitations

This research has limitations. Firstly, the database contains a small-scaled corpus, including only fifteen medical RAs with a total 60, 297 running words. A study on a larger scale would provide more solid evidence of linguistic features in medical RAs. The second limitation may result from the lack of comparative analysis. Apparently, a comparison between the corpora from two registers (e.g., the normed frequency of nominalized forms in medical RAs and in legal documents) would present distinct differences of medical English and English from other specific language use domains, thereby, to better delineate the specificity of medical English. Nevertheless, since corpus-based study is time-consuming and requires technical assistance, a better-established study with technical improvement will be processed in future research.

5.4 Conclusion

The research was conducted to examine specific linguistic features of medical English in medical research articles (RAs), contributing to explaining “medical English” in medical RAs for instructors of English when they purport to teach medical English to undergraduate medical students. The findings of the lexical and syntactic features characterize medical English by examining proportional specialized medical words, nominalization, passivization, and complex sentence structure. Based on this study, pedagogical implications are offered to instructors when they teach English to medical students. In spite of limitations described above, this research hopefully contributes to the understanding of linguistic features of medical English for curricular and instructional purposes.

Bibliography

Adams-Smith, D. E. (1983). *Prescribing and general practice: Style in medical journals*.

Paper presented at TESOL Conference, February 1983.

Ard, J. (1983). The role of the author in scientific discourse. *Paper presented at the annual*

American applied linguistics meeting, Minneapolis, MN.

Ariza, A. & Maria, A. (2005). Medical English: A short study of Spanish borrowings.

Lebende Sprachen. 50 (2), 58-61.

Baker, M. (1988). Sub-technical vocabulary and the ESP teacher: An analysis of some

rhetorical items in medical journal articles. *Reading in a Foreign Language*. 4 (2),

91-105.

Basturkmen, H. (2014). Replication research in comparative genre analysis in English for

academic purposes. *Language Teaching*. 47 (3), 377-386.

Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University

Press.

Biber, D. (1992). On the complexity of discourse-complexity: A multi-dimensional analysis.

Discourse Processes. 15, 133-163.

Biber, D., Conrad, S. & Reppen, R. (1998). *Corpus linguistics: Investigating language*

structure and use. Cambridge: Cambridge University Press.

Chia, H., Johnson, R., Chia, H. & Olive, F. (1999). English for college students in Taiwan: A

study on perceptions of English needs in a medical context. *English for Specific*

- Purposes*. 18 (2), 107-119.
- Chen, Q. & Ge, G. (2007). A corpus-based lexical study on frequency and distribution of Coxhead's AWL word families in medical research articles (RAs). *English for Special Purposes*. 26, 502-514.
- Cheng, A. (2011). Language features as the pathways to genre: Students' attention to non-prototypical features and its implications. *Journal of Second Language Writing*. 20, 69-82.
- Cowan, J. R. (1974). Lexical and syntactic research for the design of reading material. *TESOL Quarterly*. 8 (4), 389-399.
- Coxhead, A. (1998). *An academic word list*. E.. L. I. Occasional publication number 18, LALS, New Zealand: Victoria University of Wellington.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*. 34 (2), 213-238.
- Coxhead, A. & Nation, P. (2001). The specialised vocabulary of English for academic purposes. In J. Flowerdew, & M. Peacock (Eds.), *Research Perspectives on English for Academic Purposes*. Cambridge: Cambridge University Press.
- Devitt, A. J. (2009). Teaching critical genre awareness. In C. Bazerman., A. Bonini. & D. de C. Figueiredo (Eds.), *Genre in a Changing World*. Fort Collins, Colo.: West Lafayette, Ind: WAC Clearinghouse: Parlor Press.
- Devitt, A. J. (2015). Genre performances: John Swales' *Genre Analysis* and rhetorical-linguistic genre studies. *Journal of English for Academic Purposes*. 19, 44-51.
- Dudley-Evans, T. & ST John, M. J. (1998). *Developments in English for specific purposes*.

- Cambridge: Cambridge University Press.
- Dudley-Evans, T. (2001). English for specific purposes. In R. Carter & D. Nunan (Eds.). *Teaching English to Speakers of Other Languages*. 131-136. Cambridge University Press.
- Dzuganova, B. (2006). Negative affixes in medical English. *Bratisl Lek Listy*. 107 (8), 332-335.
- Dzuganova, B. (2007). Seemingly or partially negative prefixes in medical English. *Bratisl Lek Listy*. 108 (4-5), 233-236.
- Elliman, J. (1980). EST TTI Corvallis 1979. *Al Manakh*. 4 (1), 33-47.
- Fang, F. (1987). An evaluation of the English language curriculum for medical students. *Papers from the fourth conference on English teaching and learning in the Republic of China*, p.290.
- Farrell, P. (1990). *A lexical analysis of the English of electronics and a study of semi-technical vocabulary* (ERIC Document Reproduction Service ED 332 551). CLCS Occasional Paper No. 25. Dublin: Trinity College.
- Halliday, M. A. K. (1988). On the language of physical science. In M. Ghades, *Registers of written English: Situational factors and linguistic features*. 161-178. London: Frances Pinter.
- Halliday, M. A. K. (1994). *An introduction to functional grammar* (2nd ed.). London: Edward Arnold.
- Halliday, M. A. K. (2004). *An introduction to functional grammar* (3rd ed.). London: Edward Arnold (revised by Matthiessen, C.).

- Hutchinson, T. & Waters, A. (1987). *English for specific purposes*. New York: Cambridge University Press.
- Johns, A. & Dudley-Evans, T. (1991). English for specific purposes: International in scope, specific in purpose. *TESOL Quarterly*. 25 (2), 297-314.
- Johns, A. M. (2003). Genre and ESL/EFL composition instruction. In B. Kroll (Ed.), *Exploring the Dynamics of Second Language Writing*. 195-217. Cambridge, UK: Cambridge University Press.
- Khuwaileh, A. (1994). ESP after thirty years: An overview of the position of ESP in the 1990s. *Language Research*. 30 (3), 657-665.
- Khuwaileh, A. (2006). Medical rhetoric: A contrastive study of Arabic and English in the UAE. *English Today* 86. 22 (2), 38-44.
- Lemke, J. L. (1990). Technical discourse and technocratic ideology. In M. A. K. Halliday, J. Gibbons, & H. Nicholas, *Learning, keeping and using language. Vol. II. Selected Papers from the 8th World Congress of Applied Linguistics*. 453-460. Amsterdam, Philadelphia: John Benjamins.
- Li, L. & Ge, G. (2009). Genre analysis: Structural and linguistic evolution of the English-medium medical research article (1985-2004). *English for Special Purposes*. 28, 93-104.
- Lei, L. & Liu, D. (2016). A new medical academic word list: A corpus-based study with enhanced methodology. *Journal of English for Academic Purposes*. 23, 42-53.
- Lyons, J. (1977). *Semantics*. v.2. Cambridge: Cambridge University Press.
- Maher, J. (1986). The development of English as an international language of medicine.

Applied Linguistics. 7 (2), 206-217.

Maher, J. (1992). *International Medical Communication*. Ann Arbor, Mich.: University of Michigan Press.

Marco, M. J. L. (2000). Collocational frameworks in medical research papers: A genre-based study. *English for Specific Purposes*. 19, 63-86.

Martin, A. V. (1976). Teaching academic vocabulary to foreign graduate students. *TESOL Quarterly*. 10 (1), 91-99.

Master, P. (1992). Genre analysis: English in academic and research settings: John Swales, Cambridge: Cambridge University Press, 1990. 260 pp. (pb). *Book notices. Journal of Pragmatics*. 17(3), 286-289.

Mauranen, A. (1993). *Cultural differences in academic rhetoric*. Frankfurt: Peter Lang.

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

Nwogu, K. N. (1997). The medical research paper: Structure and functions. *English for Specific Purposes*. 16 (2), 119-138.

Ramirez, C. G. (2015). English for specific purposes: Brief history and definitions. *Revista de Lenguas Modernas*. 23, 379-386.

Rahman, M. (2015). English for specific purposes (ESP): A holistic review. *Universal Journal of Educational Research*. 3 (1), 24-31.

Rizzo, C. R. & Llopis, M. A. (2011). New and further approaches to ESP discourse: Genre study in focus. *International Journal of English Studies*. 11 (1), i-vii.

Robinson, P. (1991). *ESP today: A practitioner's guide*. New York: Prentice Hall.

- Salvager, F. (1983). The lexis of fundamental medical English: Classificatory framework and rhetorical function (a statistical approach). *Reading in a Foreign Language*. 1 (1), 54-64.
- Salvager-Meyer, F., Defives, G., Jensen, C. & Filipis, M. (1989). Principal component analysis and medical English discourse: An investigation into genre analysis. *System*. 17 (1), 21-34.
- Salvager-Meyer, F. (1994). Hedges and textual communicative function in medical English written discourse. *English for Specific Purposes*. 13 (2), 149-170.
- Skelton, J. (1994). Analysis of the structure of original research papers: An aid to writing original papers for publication. *British Journal of General Practice*. 44, 455-459.
- Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Swales, J. M. (2004). *Research genres: Exploration and applications*. Cambridge: Cambridge University Press.
- Thomas, S. & Hawes, T. P. (1994). Reporting verbs in medical journal articles. *English for Special Purposes*. 13 (2), 129-148.
- Wang, J., Liang, S. & Ge, G. (2008). Establishment of a medical academic word list. *English for Specific Purposes*. 27, 442-458.
- Wood, A. (2001). International scientific English: The language of research scientists around the world. In J. Flowerdew, & M. Peacock (Eds.), *Research Perspectives on English for Academic Purposes*, 71-83. Cambridge: Cambridge University Press.
- Yang, A., Zheng, S. & Ge, G. (2015). Epistemic modality in English-medium medical

research articles: A systemic functional perspective. *English for Special Purposes*. 38,
1-10.

Zheng, S., Yang, A. & Ge, G. (2014). Functional stylistic analysis: Transitivity in
English-medium medical research articles. *International Journal of English Linguistics*.
4 (2), 12-25.