Characteristics of Independent Claim:
A Corpus-Linguistic Approach to Contemporary
English Patents

Darren Hsin-Hung Lin[*,+] and Shelley Ching-Yu Hsieh[+]

Abstract

This paper presents a corpus-driven linguistic approach to embodiment in modern patent language as a contribution to the growing needs in intellectual property rights. While there is work that appears to fill a niche in English for Specific Purposes (ESP), the present study suggests that a statistical retrieval approach is necessary for compiling a patent technical word list to expand learner vocabulary size. Since a significant percentage of technical vocabulary appears within the range of independent claim among claim lexis, this study examines the essential features to show how it was characterized with respect to the linguistic specificity of patent style. It is further demonstrated how the proposed approach to the term independent claim contained in the patent specification is reliable for patent application on an international level. For example, clausal types that specify how clauses are used in U.S. patent documents under co-occurrence relations are potential for patent writing, while verb-noun collocations allow learners to grip hidden semantic prosodic associations. In short, the research content and statistical investigations of our approach highlight the pedagogical value of Patent English for ESP teachers, applied linguists, and the development of interdisciplinary research.

Keywords: Intellectual Property Rights, Patent Document Processing, Corpus, Systemic Functional Linguistics, Co-Ocurrence.

[*] Graduate School of Decision Science and Technology, Tokyo Institute of Technology, Japan

 E-mail: darryanlin@gmail.com

 The author for correspondence is Darren Hsin-Hung Lin.

[+]Department of Foreign Languages and Literature, National Cheng Kung University, Taiwan

 E-mail: shelley@mail.ncku.edu.tw

1. Introduction

In the knowledge economy age, intellectual property rights (IPR) are important assets. Especially to the knowledge industry, IPR is the key measure of a company competing with others.

As globalization has resulted in rapid greater economic growth, the challenges of interdisciplinary communication concerned with intellectual property and other significant sector encounters have increased. The recognition of this importance has brought intellectual property into the limelight. Resulting from such recognition, as well as the recent emphasis on using English as the lingua franca to apply patents on an international level, the application of technical vocabulary for the writing of professional patents has become an essential issue in applied linguistic research.

2. Literature Review

Law is a system of rules, carried out by lawyers, attorneys, jury, paralegals, and related legal institutes. It is not just a tool that shapes politics, economy, and society, but also it is a socially prominent medium applied to maintain social order. A large number of recently specialized areas, such as international trade, economics, finance, accounting, and electronic commerce, recently have been recruiting interdisciplinary specialists with expertise in both law and English proficiency to engage in legal workplace practice.

While the widespread use of law has naturally had impact on almost all fields of discipline, the increasing use of English has radically changed the way in which we perceive this language's international function (Modiano, 2001). English for Legal Purposes (ELP), a growing trend in the field of English for Specific Purposes (ESP), therefore, has become a research topic (Dudley-Evans & St. John, 1998:7) and is used in either professional or legislative settings.

As the Internet shortens the distance between countries, patent information is readily available via online access. To protect novel inventions, intellectual property law is a developing domain across legal professions. The area of intellectual property law includes patent law, copyright law, trademark law, and trade secret law, together with some aspects of other branches of the law, such as licensing and unfair competition (American Bar Association, 2010). Intellectual property lawyers are required to have command of interdisciplinary knowledge as new developments in law generate the need for lawyers with specific backgrounds-patent law, technology law, business law, and economy economic law. It is worth mentioning that the demand for intellectual property lawyers has remained unusually high even though the global markets were affected by economic recession in the end of 2007 (World Intellectual Property Organization, 2009). As long as novel inventions continue to be

created, there is a need for intellectual property law to be enforced to protect human rights and their invisible property for specific purposes.

Patent, known as interdisciplinary innovation, has drawn the attention of most lawyers. Tsai (2008) reported that patents are granted for innovations as they reflect economic growth of a country by illustrating creative activities and displaying the knowledge power of that particular country or region. The diversity of languages used in patent applications has boosted translation demand for patent right protection. Besides, many paralegals, such as patent attorneys, lawyer assistants, or translators, participate in legal circles for a living nowadays. It is important to equip them not only with background knowledge, but competency of professional writing for the job market. Accordingly, in the present study, the researchers look at the role of patent writing for research purposes.

2.1 Corpus-based Studies on Law

Corpus linguistics is often concerned with the study of natural language, which explores real and authentic language use by means of a corpus (McEnery & Wilson, 2001). At the present day, a corpus represents a wide variety of language use, both spoken and written language, by a collection of texts stored in a computer (Mudraya, 2006).

Biber, Conrad, & Reppen (1998) claimed a corpus-based analysis is characterized by four primary features. First, a corpus-based study is empirical, for it uncovers the natural patterns of real language use. A corpus-based study, however, relies heavily on computer-assisted tools. Computer-assisted tools, such as concordancers, enable researchers and practitioners to tag linguistic features, to code grammatical variants, and to carry out data capture and mark-up. Third, research data are analyzed either quantitatively or qualitatively in a corpus-based study. For example, the total frequency of the term independent claim is shown in a quantitative way. The concordancer can show the frequency of coded articles and average words per article. Analysis probing into observing linguistic phenomena of the term, such as polysemy or near-synonym, in turn, is qualitative. Finally, a corpus-based analysis is meaningful once research questions have been proposed. A corpus may be designed to characterize the use of an independent claim adopting a functional approach. Since the investigation is prompted to answer the research questions concerned with such design, the corpus-based analysis becomes meaningful. As corpus-based study is widely accepted and has become the norm in interdisciplinary social sciences (Ball, 1996; Chen, 2001; Lee & Swales, 2006), it further represents how language has been evaluated in prescriptive and descriptive ways in academic research (Dudley-Evans & St. John, 1998; Hyland & Tse, 2005; Nelson, 2006; Hyland, 2008).

In sum, a corpus-based study is insightful in that it is not only representative in social science research, but also it contributes to characterizing the legal language people associate

with (Hsieh, 1998). Over the years, there has been corpus-based research on law in social science research (Feak, Reinhart, & Sinsheimer, 2000; Candlin, Bhatia, & Jensen, 2002; Badger, 2003; Chiu, 2008). Nevertheless, few works concerned with patents can be found. As corpus-based studies have been conducted widely in social science research, the application of corpus tools has been noticed in recent development. The present study is warranted by such trends for investigation into contemporary patents.

2.2 ESP Studies on Law

ESP is now well established as an important and distinct part of English teaching (Cheng, Sin, & Li, 2008:16). As English has acquired the status of lingua franca in almost any field of research, the teaching of ESP generally has been seen as a separate activity within English language teaching and ESP research has been seen as an identifiable component of applied linguistic research (Dudley-Evans & St. John, 1998).

The origins of ESP can be traced back to the 1960s, when there was a growing need for technological and business industries (Swales, 2000:59-61). ESP, the prime realization of applied discourse analysis, later evolved for every specialized area needing appropriate teaching materials. Recently, ESP has been utilized as an umbrella term with a multitude of acronyms denoting the various sub-fields (Dudley-Evans & St. John, 1998).

Under the ESP framework, there are two major sub-fields, English for Academic Purposes (EAP) and English for Occupational Purposes (EOP), which are distinguished by their research nature and pedagogical tradition (Robinson, 1991; Dudley-Evans & St. John, 1998). EAP is concerned with students' needs to learn academic language, which constitutes the majority of ESP, whereas EOP comprises professional purposes in administration, medicine, law and business, and vocational purposes for non-professionals in work or pre-work situations (Dudley-Evans & St. John, 1998:7).

In the ESP domain, ELP is an important but comparatively uncultivated corner (Dudley-Evans & St. John, 1998:51). González and Vyushkina (2009) characterize English for Academic Legal Purposes (EALP) as being used in university degree programs, while English for Occupational Legal Purposes (EOLP) is used in training for practical skills in the workplace. Over the years, there has been continuing interest in the research of EALP (Bhatia, 1993; Bowles, 1995; Harris, 1997; Feak, Reinhart & Sinsheimer, 2000; Candlin, Bhatia, & Jensen, 2002; Badger, 2003; Du, 2009). Nevertheless, studies have been concerned mostly with material development, genre analysis, and curriculum design. Corpus-based studies on EOLP, in contrast, are relatively undeveloped. Badger (2003) once conducted a corpus-based study on law in the genre of newspaper law reports. He found that newspaper law reports serve the same function as law cases do, which facilitates law school students in identifying the reasoning of the legal decision of the case. His corpus-based study is innovative, but it is

EALP and is solely for reading. To be specific, corpus-based applications on EOLP are comparatively unseen and the voice that professional writing gathers in the workplace entails the directions for future research.

Accordingly, it is confirmed that while EALP is widely developed for law school students and academic purposes, there is an underlying need to build up EAOP, in particular, Patent English, for workplace needs.

2.3 Vocabulary Studies on Law

Writing for specific purposes requires familiarity with not only the content but also the language. Unfamiliarity with vocabulary in writing is perceived to be a challenging task for language learners. As the importance of teaching vocabulary in ESP has gained recognition (Swales, 1983), Coxhead & Nation (2001) have categorized vocabulary in ESP into four groups: high-frequency words, academic vocabulary, technical vocabulary, and low-frequency vocabulary.

Nation (2001) defines those words in the use of writing. High-frequency words refer to the most frequently used 2000 words of English used in all types of writing. Low-frequency words are the rarely used terms and cover only 5% of all words. Academic words, namely semi-technical or sub-technical vocabulary, are for academic purposes. This kind of vocabulary is common to a wide range of academic fields but is not what is known as high-frequency vocabulary and is not technical in that it is not typically associated with just one field (Chung & Nation, 2003:104). In contrast, technical words are the ones used in a specialized field that are considerably different from subject to subject. As Chung & Nation (2003:104) point out, technical vocabulary is largely of interest to and used by people working in a specialized field. In the genre of law, Mellinkoff (1963) suggests legal vocabulary highlighting those common words with uncommon meanings. For example, merger and acquisition bear the same literal meaning as 'combination' in general English. Nevertheless, in economic and financial law, *merger* depicts the acquisition of one company by another. This combination into a single legal entity will increase the benefits to each other and is semantically positive. As to *acquisition*, the combination often bears unequal treatment and is considered negative.

Since there is very little research on technical vocabulary in legal disciplines, Harris (1997) analyzed procedural vocabulary extracted from the area of English contract law. His research shows that technical words enhance legal reading and also strengthen text analysis skills. Denton (2009:5) covered frequently used legal vocabulary in his teaching. Specific meaning of vocabulary, such as merger and acquisition in economic law, is viewed as concept for him to teach. His research concludes that the learning of terminology for Legal English is the priority for participants to foster when they are learning vocabulary conceptually. In other

words, learning legal vocabulary with concepts of the target context is essential in vocabulary development. Haberstroh (2009) developed the legal academic word list. His research enriches the well-established area of EALP at the present day; however, the rapidly growing trend of EOLP remains comparatively undeveloped.

In brief, a general conclusion can be drawn in that there is a need to prepare inter-disciplinary patent writing, but exploring technical vocabulary with corpus-driven approaches into such development has the higher priority.

3. Methodology for Corpus Creation

The present study adopted a corpus-based research approach to study patent technical words from the USPTO (United States Patent and Trademark Office) glossary[1] in the field of intellectual property, with an emphasis on their frequency and word associations in contemporary patents.

In assessing the proper coverage needed for a lexical study, the distribution of each IPR domain is taken into consideration beforehand. Figure 1 shows the results of the coverage of technical words of IPR from the USPTO glossary.
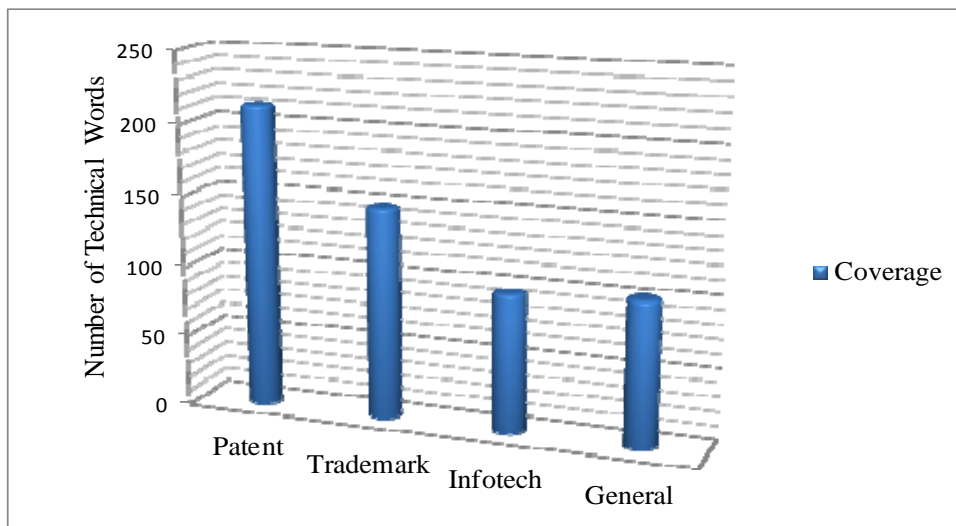
Figure 1. The coverage of technical words in intellectual property

The coverage was confined to the domains. As the USPTO glossary surveyed, four primary domains were outlined-patent, trademark, infotech (information technology), and general domain. Among the total 558 words of the glossary, 212 words are word items

---

[1] USPTO glossary is available at http://www.uspto.gov/main/glossary

included in the patent on a domain level, making up 38% of the total. In other words, the coverage of patent technical words was 38%, which is much higher than the 18.3% of the 102 words used in a general domain, as shown in Figure 1. Compared with the coverage of trademark (26.3%) and infotech (17.4%) domains, patent technical words are more widely covered in intellectual property. This suggests that there is a growing need in the area of patents and is consistent with the literature review, which suggests that patent plays a significant role in the genre of intellectual property.

3.1 Purpose

One of the major objectives in this section was to find the most frequently used technical words in patents. This aim was achieved by calculating the frequencies of each patent word in Figure 1. The frequency of the patent technical words has been listed according to the frequency of their occurrence in the USPTO Patent Full-Text and Image Database (PatFT)[2], and the distribution is presented in Table 1.

Table 1. Distribution of patent technical words in PatFT

Times of Occurrence	Number of Words	Percentage	Accumulative Percentage
≧ 1,000,000	7	53.20	53.20
1,000,000 ~ 999,999	23	42.56	95.76
10,000 ~ 99,999	21	3.62	99.38
1,000 ~ 9,999	39	0.53	99.91
100 ~ 999	40	0.07	99.98
1 ~ 99	53	0.02	100.00
0	29	0.00	100.00
TOTAL	212	100	100.00

Among the 212 patent technical words, 90 words (99.91%) occurred more than 1000 times in PatFT and are considered frequently used patent technical words. There were only 53 words (0.02%) that appeared less than 100 times and 29 words that did not appear at all, both of which are viewed as not frequently used technical words in patents. The other 40 words occurred less than 1000 times but more than 100 times in PatFT.

As can be seen, there were 7 words that occurred more than one million times. Among them, the most frequently used technical word was the verb 'comprising,' which appeared 3,785,213 times. Other technical word items, such as scope, patent, group and element,

---

[2] USPTO Patent Full-Text and Image Database (PatFT) is available at

   http://patft.uspto.gov/netahtml/PTO/search-adv.htm

consisting of, and drawing, occurred over one million times. The high-frequency of these words reflects the important role of technical vocabulary in patent texts.

With regard to word associations, Nattinger (1988) suggests that grouping of the words according to their meanings enhances vocabulary learning. He once mentioned that word grouping can be presented in the form of topic (situational sets). With a library, such words as book, shelf, borrow, loan, and so on can be taught together for teaching and learning. In order to get a clearer picture of the patent technical words for better use, the researchers here made a detailed analysis of the word associations based on topic.

The 212 patent technical words are considered to be statistically unusually frequent in their occurrence, but it was then noted that they seemed to fall into a limited number of recurring topic sets; therefore, six sections were proposed based on words in the same semantic network or field that share similar meanings or semantic features in PatFT: 'patent activity (99),' 'patent aid (25),' 'patent community (23),' 'patent claim (17),' 'patent description (30),' and 'people of the patent community (18)'. This was made not only on an intuitive basis, but also on the criteria of the produced data. The following illustrates the criteria the researchers set up for each section.

(1) What do patent-specific activities usually consist of? (Patent activity)

(2) What tools can be applied in a patent-specific context? (Patent aid)

(3) Where are patent-specific places in United States? (Patent community)

(4) What entities do patent applicants need for specific requests? (Patent claim)

(5) What specific entities can usually be found in patents? (Patent description)

(6) Who are in the patent-specific contexts? (People of the patent community)

Table 2 presents the top ten frequently used technical words in each of the six sections, which are arrayed according to their frequency of occurrence in descending order based on PatFT.

Table 2. Top 10 technical word items of six topic-based sections

Rank	Activity	Aid	Community	Claim	Description	People
1	patent	concept	Group	comprising	specification	representative
2	disclosure	doctrine of equivalents	Pubs	scope	sequence listing	person
3	application	file wrapper	TC	element	filing date	assignee
4	patent application	ADS	Technology Center	consisting of	serial number	applicant
5	continuation	mask work	ISA	drawing	application number	inventor
6	interference	EFS	IB	dependent claim	PLT	practitioner
7	demand	PAIR	RO	composed of	Control No.	attorney
8	restriction	OG	IPEA	independent claim	publication number	disclaimer
9	designation	PSIPS	GAU	benefit claim	issue date	CSR
10	divisional application	PALM	Group Art Unit	priority claim	patent number	lawyer

The keyword analysis made on a large number of words in the present study was not intended solely to keep interdisciplinary learners informed of the frequency of some word items, but also to awaken the learners to the influence of intellectual property and patent on lexical units, which might vary in accordance with different topics.

In addition to the top ten word items, the researchers calculated the total frequency and total words of each section. Table 3 shows the total frequency and total words of each topic section.

In the patent technical word list, patent claim accounts for 54%, followed by patent activity (making up 28%), patent community and people of the patent come next at 6%, and finally patent description (represented by 4.5%). Patent aid only constitutes 1.5% of all.

As patent law 35 U.S.C.§112 Paragraph 1 reads, "patent claim" is viewed as the specifications, containing a written description of the invention and the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same, and shall set forth the best mode contemplated by the inventor of carrying out his invention. That is to say, the patent claim of a published patent informs the public of the scope of rights that distinguish the invention. As it is technically dealt with, specific terms are used, allowing users to become familiarized with the invention an applicant owns.

Table 3. Frequency of the patent technical world list

Topic	Technical Words	Total Frequency	Percentage	Rank
Patent Activity	99	6,622,873	28	2
Patent Claim	17	12,695,484	54	1
Patent Community	23	1,455,693	6	3
People of the Patent	18	1,468,215	6	3
Patent Description	30	1,060,782	4.5	5
Patent Aid	25	342,988	1.5	6
TOTAL	212	23,646,035	100	-

As shown in Table 3, "patent claim," which has high priority, is valuable for corpus-based research. Besides, to build up a small-scale corpus for the present study, the researchers analyzed 'patent claim' based on parts of speech for further investigation. Table 4 shows the results.

Table 4. Distribution of patent claim in the patent technical word list

Group	Patent Technical Word	Total Frequency	Percentage
Noun	scope	2,459,656	55.24
	element	1,245,265	
	drawing	1,015,261	
	dependent claim	625,886	
	independent claim	587,926	
	benefit claim	437,599	
	priority claim	381,352	
	withdrawn claim	227,433	
	canceled claim	32,306	
	multiple dependent claim	494	
	rejoinder	80	
	claims	6	
Verb	comprising	3,785,213	44.76
	consisting of	1,165,427	
	composed of	617,353	
	consisting essentially of	114,211	
	having	16	
TOTAL		12,695,484	100

As can be seen, there are only two syntactic categories that can be found and noun words outperform verb words, making up a 55.24%. Accordingly, the researchers lay their attention on noun words of patent claim, and compile a patent technical word corpus.

As a first step motivated toward the establishment of a patent corpus for investigation, the present study was based on a corpus of U.S. patents, European patents, Patent abstracts of Japan, PCT[3] patents, and U.K. patents over a decade (2000 to 2009) gathered from LexisNexis[4], a corpus of a multitude of information for professionals in legal fields, in forms of case, newspaper, journal, and magazine reporting.

The Patent Technical Word Corpus (PTWC), made up of patent claim texts retrieved from LexisNexis, was created consisting of 16,101,256 word tokens.

Although LexisNexis does not have built-in patent claim subcorpora, the self-compiled PTWC adds significant strength to the development of claim language. Although an available specialized corpus contains an infinite amount data, constructing a small scale one would be needed for a profound linguistic study (Hsieh, 1998:48). Therefore, the PTWC was established for the present study.

3.2 Lexical Analysis Software

Owing to the size of the text collection, the quantitative analysis was computer-assisted, using WordSmith Tools 5.0 (Scott, 2008) to search for the word item as a string of letters to ascertain the absolute and relative frequency. The concordancer-tagged function of WordSmith 5.0 allowed us to calculate collocations and clusters around the search or node word.

With the help of such tools, we can find more discriminative linguistics patterns and structures of patents. Table 5 shows the result of citations from each level.

As can be seen, element, drawing, and scope were noun words that occurred over 10,000 times. In turn, claims occurred less than 10,000 times but more than 1,000 times on in-between levels. There were noun words that occurred less than 1,000 times-independent claim, dependent claim, multiple dependent claim, benefit claim, rejoinder, priority claim, withdrawn claim, and canceled claim. Compared with the other two levels, the current level covers noun words that were comparatively less used but more precisely employed. Dependent claim, for example, contains a reference to a claim previously set forth. Multiple dependent claim, in contrast, contains a reference to more than one claim previously set forth.

---

[3] PCT (Patent Cooperation Treaty).

[4] LexisNexis is available at https://www.lexisnexis.com

Table 5. Citations of patent claim

Patent Claim	Citation	Level
element	54,151	$\geqq$ 10,000 times
drawing	40,634	
scope	28,236	
claims	5,800	$\leqq$ 10,000 times
independent claim	249	$\leqq$ 1,000 times
dependent claim	59	
multiple dependent claim	58	
benefit claim	57	
rejoinder	22	
priority claim	3	
withdrawn claim	0	
canceled claim	0	
TOTAL	129,268	

As WordSmith tools provide a comprehensive view of noun words in patent claim texts, it is noted that more effort should be made to explore the possibilities on those that occurred less than 1,000 times. Therefore, the present study lays its focus on such words.

3.3 Data Selection Criteria

The researchers observed citation of each level from Table 5 and found those that appeared less than 1,000 times were more specific word items. Among them, 'independent claim' outperforms others with 249 occurrences.

Technically, an "independent claim" is a proper noun of patent that formally describes the invention in adding the essential features. In the patent application for a pencil, the independent claim might begin with "a device comprising a cylindrical piece of wood with a piece of graphite inserted into the center of the wood." In such a case, the pencil was distinguished with regard to the shape (cylindrical) and the materials it was made of (wood and graphite). For the same pencil with another shape, it will not be taken into consideration for such an invention.

Patent claim is the precise legal definition of the invention, identifying the specific elements of the invention for which the inventor is claiming rights and seeking protection. A patent claim shares technical terminology with the rest of a patent but differs greatly in its

contents and syntax (Sheremetyeva, 2003). Of patent claims, *independent claim* best describes the invention in adding essential features. Since the independent claim is specific in that it stands on its own and does not rely upon or refer to any other claims in a patent, the researchers chose "independent claim" as the target word for data analysis.

The corpus of the present study contains 98 English patents with 'independent claim' tagged in the patent specification[5], and is made up of 4,887,084 word tokens. The researchers use the concordance function to find the technical terminology 'independent claim,' with 249 citations generated. There is a list of the 249 examples of 'independent claim' with the words that preceded and followed. Figure 2 shows part of the concordance.

Figure 2. Concordance of independent claim tagged in the patent specification

Out of the 249 examples of 'independent claim,' 5 were irrelevant to the researchers' analysis because 'independent claim' was used without a subject in the present progressive tense-"identifying at least one independent claim of the patent." Of the remaining 244 examples, all concordance entries for each were stored. Then, the source texts for each concordance line were manually analyzed one by one for further investigation. The authors provide a screenshot of the source text tool interface with technical terminology *independent claim* contained in the patent specification documentation in Figure 3.
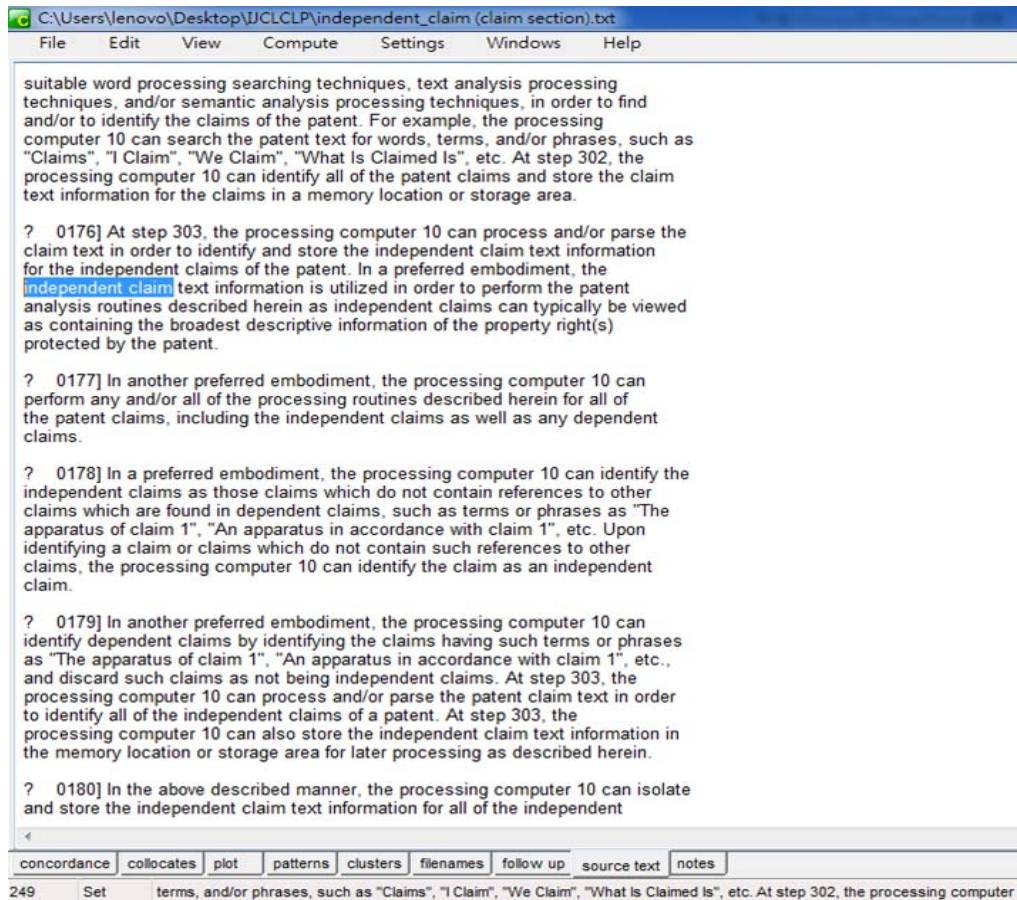
Figure 3. Source text tool interface with independent claim taggings

4. Transitivity Development of Independent Claim

To explore the embodiment, the researchers first looked at the transitivity in the corpus. Analyzing the transitivity patterns of independent claim, in turn, contributes to the understanding of how verbs measure the clausal function through Halliday's systemic-functional view of language.

4.1 Transitivity

In Halliday's (2004:168) study, he distinguished six central processes that elicit transitivity to describe a whole clause, rather than the verb and its object. These six central processes, in turn, were material, mental, relational, behavioural, verbal, and existential processes. Each process type, however, constructs a particular experience that distinguished distinguishes clausal functions. Li (2010:3447) suggests transitivity analysis is a semantic perspective on the ideas expressed by a clause, a proposition about the world in which an event, situation, relation or attribute is predicted of some participants. Toward a functional view of language, the total set of functions used in interpreting the clause as representation can include the process types summarized in Table 6.

Table 6. Process types, their meanings and participants (Halliday, 2004:260)

Process type	Category meaning	Participants, directly involved	Participants, obliquely involved

Material
  action
  event
	Doing
	  doing
	  happening
		Actor, Goal	Recipient, Client; Scope; Initiator; Attribute

Mental
  perception
  cognition
  desideration
  emotion
	Sensing
	  seeing
	  thinking
	  wanting
	  feeling
		Senser, Phenomenon	-

Relational
  attribution
  identification
	Being
	  attributing
	  identifying
		Carrier, Attribute
		Identified, Identifier;
		Token Value
			Attributor, Beneficiary
			Assigner

Verbal	saying	Sayer, Target	Receiver; Verbiage
Behavioural	behaving	Behaver	Behaviour
Existential	existing	Existent	-

As for the present study, transitivity analysis is applied as the 244 citations of independent claim were examined. The researchers first singled out each citation as a constructed clause. In this regard, the researchers conducted analysis at the clausal level to better reflect the actual process an independent claim was associated with. In this manner, the researchers elicited the verbs that distinguished each process type. For verification, the researchers derived nominals that represent participants in each clause. The researchers give an instance in (1).

(1) <u>The processing computer</u>   <u>can store</u>   <u>the independent claim text information</u>
            Actor   Process   Goal

As shown in (1), 'store' outlines a material process in which the *processing computer* (Actor) accumulates *independent claim text information* (Goal). In such a case, *processing computer* which occurs with 'store' might provide selectional features[6] (Chomsky, 1965:111) of the knowledge of independent claim. It is noted that the verb-noun collocation 'store + independent claim' followed by *processing computer* is a subtle distinctive feature of independent claim that is expected to be known for such a grammatical pattern, which makes up knowledge of the grammar of patents in use. The investigation of this collocationally-fixed relationship will, in turn give insights to learners on how independent claim is used on a lexical level and prepare them for the actual business world they may need to work in or offer them the information regarding modern patent language where they already work.

4.2 Transitivity Structures of Independent Claim

Since independent claim describes the invention in adding the essential features, in this section, independent claim is annotated by three primary clauses of the total four clausal types found in the data. They are material, relational, and verbal clauses. The concept of clause as representation (Halliday, 2004) is applied to remind language users where to locate independent claim to produce correct sentences.

There are a total of four clausal types found in the data (see Table 7).

Table 7. Clausal types of patent specification tagged with independent claim

Clausal Type	Total Frequency	Percentage
Material Clauses	127	52.0
Relational Clause	65	26.6
Verbal Clause	48	19.7
Existential Clause	4	1.7
TOTAL	244	100

As Table 7 shows, material clauses make up the largest proportion of the total, accounting for 52%, with relational clauses coming next at 26.6%, followed by verbal clauses, making up 19.7%, and existential clauses at 1.7%. Nevertheless, behavioral clauses were not found as legal discourse in the Republic of China to address such phenomena. Tsai (2006:108) explains that law is essential as it elaborates the obligations of human beings. Behaviors such

---

[6] Chomsky (1965:111) defined selectional features as vocabulary knowledge requiring a number of qualified rules in which lexical items in question cannot appear. For example, *admire* only occurs with subject nouns that are human, such as man, not abstract ones, like faith.

as dream, cough, and cry, however, are basic instincts that human beings embrace. There is less importance to further develop such behaviors in the discourse of law. Although patent language and legislative language differ in their rationale, the fact that declarative sentences were favored in the examined clauses of the present study is in accordance with Tsai's (2006:109) research on legislative language.

In sum, it can be concluded from Table 7 that material clauses are the most frequent experience that independent claim shares, while existential clauses are the least. These clausal types of independent claim provide direction for the novice. They should learn material clauses first. As the distribution of independent claim involves different transitive processes, we make a further step to delineate how lexical items were generated with reference to the co-occurrence relations.

5. Lexical and Clausal-Specific Features of Independent Claim

In an attempt to characterize language-specific entities that could serve as a pedagogical base to help language awareness for patent writing, we explicate lexical and clausal-specific features of independent claim to promote discourse-level proficiency in modern patent language learning contexts.

5.1 Verb-Noun Collocation

The researchers investigated the verb-noun collocations in three primary clauses and introduce verb-noun collocations that make up the knowledge that learners need to be aware of in their learning. Verb-noun collocation here is defined as verbs with specific meaning that collocate with independent claim. The frequency of the verb-noun collocations then is annotated.

Technically, frequently used verbs in patents can be seen as concepts that carry meanings to specify the clauses for communication. Among the 244 examined clauses, the researchers found 23 verb-noun collocations from the data. Meanings of each collocating verb from the verb-noun collocations were carefully analyzed. Table 8 illustrates the results.

As Table 8 indicates, the auxiliary 'be' made up nearly 8.2%, while the rest constitutes 91.8%. 'Identify' and 'direct' were frequently used with independent claim, accounting for approximately 46%. 'Contain,' in contrast, was the third most remarkable (17.2%). These three verbs represent over 63% of the verb-noun collocations. There were five verb-noun collocations (identify, direct, contain, be, and correspond) that appeared over 10 times, making up 76.2%.

Table 8. Collocating verbs of patent specification tagged with independent claim

Verb	Verb Meaning	Total Frequency	Percentage
identify	to extract, recognize, discover, or find	61	25.00
direct	to request or enjoin with authority	51	20.90
contain	to have within	42	17.20
be	to state of having existence	20	8.19
correspond	to be in conformity or agreement	11	4.50
infringe	to encroach upon in a way that violates law or the rights of another	7	2.90
analyze	to determine the nature and relationship of the parts of by analysis	6	2.50
isolate	to set apart from others	6	2.50
perform	to carry out an action or pattern of behavior	6	2.50
generate	to bring into existence	5	2.00
process	to a series of actions or operations conducing to an end	4	1.64
store	to place or leave in a location	4	1.64
regard	to an aspect to be taken into consideration	4	1.64
exist	to have the functions of vitality	4	1.64
break up	to do away with	2	0.80
formulate	to develop a formula for the preparation	2	0.80
permit	to consent to expressly or formally	2	0.80
fall	to come within the limits	2	0.80
illustrate	to make clear	1	0.41
provide	to take precautionary measures	1	0.41
utilize	to turn to practical use or account	1	0.41
associate	to bring together or into relationship	1	0.41
exhibit	to show or display outwardly, especially by visible signs or actions	1	0.41
TOTAL		244	100

In most cases, 'identify' (to extract, recognize, discover, or find) collocates with an independent claim, making up 25% of the verb-noun collocations. Examples (2) to (4)

demonstrate such collocations.

(2) The database can also contain any one or more of software programs and/or algorithms for parsing patent language in order to <u>identify</u> a claim or claims of a patent, software programs, and/or algorithms for parsing patent language in order to identify an independent claim or independent claims of a patent.

(3) Parsing claim information of the patent in order to <u>identify</u> at least one independent claim.

(4) The processing computer can <u>identify</u> and store the preamble text information for the independent claim.

As can be seen, in (2) to (4), "independent claim" is viewed as the Goal. For instance, Example (2) points out that database will parse the patent language to be identical in independent claim. Example (3) elaborates the behavior to parse information regarding patent claim to recognize independent claim. In (4), the processing computer enables the preamble textual information to be extracted with independent claim as the Goal. In (2) to (4), 'identify' is with the precise meaning "to cause something to become identical," implying that patent is a specific genre where the fixed verb meaning is embodied.

While vocabulary knowledge may involve a number of qualified rules of the kind Chomsky (1965) calls "selectional feature," a collocating verb has a selectional feature of its own. In other words, a collocating verb is a collocation-based feature of verb-noun collocation that maps the detailed contour of knowledge on clausal types. For each clausal type, the verb-noun collocations involved explain the grammar of words, the interaction between two associated participants, and the experience a particular clausal type has embraced. In this regard, verb-noun collocations elicited from the present study can equip learners with a better sense of the firmness of this collocational relationship.

5.2 Clausal Nominalization

As the verb-noun collocation 'independent claim + direct' shows a strong tendency in characterizing the passive structure of verbal clauses, the researchers found the nominalized *to which the independent claim is directed* functions as an adverbial constituent of the clauses and is unusually positioned clause-final. Based on this, 'independent claim + direct' is a selectional feature of clausal nominalization in verbal clauses as transitivity analysis is applied. Clausal nominalization, in turn, is a functional feature that elucidates mutual information shared in verbal clauses of the modern patent language. The following elaborates our finding.

Theme is a single constituent that happens to come at the beginning of a given clause that will label the function of the clause, while everything else in the clause is known as rheme. Example (5) illustrates the theme-rheme structure of the clause.

(5)   <u>What the duke gave to my aunt</u>   <u>was this teapot</u>

             Theme   Rheme

As Halliday (2004) elaborates, this kind of clause is known as a "thematic equative" because it sets up the theme-rheme structure in the form of an equation, where theme=rheme. According to Halliday, a form, such as what the duke gave my aunt, is an instance of a structural feature known as nominalization. In this case, theme is the primary element, while nominalization serves a thematic purpose for communication. Nevertheless, once the normal relationship is reversed, the nominalization becomes marked. In this fashion, it is called 'marked thematic equative,' as presented in Example (6).

(6)   <u>This teapot</u>   <u>was what the duke gave my aunt</u>

      Theme   Rheme

Syntactically, the theme-rheme structure constructs the topic of a clause and further helps learners identify the linguistic elements within, such as Goal and Actor of material clauses, Say and Verbiage of verbal clauses, or Identified and Identifier of relational clauses. In this regard, the researchers found verbal clauses in the data displayed marked thematic equative followed Halliday's research. Such kinds of nominalization of clausal or clause-like structures into a nominal one conform to Heyvaert's (2003) nominalization as functional reclassification. Based on Lehrmann (1988), such nominalization is the process wherein a clause is reduced so that it loses the properties of being a clause but acquires nominal properties that allows it to become a nominal or adverbial constituent of a matrix clause. In Halliday's (2004) term, such nominalization is known as a structural feature in which theme-rheme structure in the form of an equation occurred. In the following, the researchers examine clausal nominalization of verbal clauses and specify the syntactic environment where nominalized units were found.

Of the 48 verbal clauses, the researchers found that 48 (100%) were nominalized. Table 9 shows the findings.

Table 9. Clausal nominalization of verbal clauses

Item	Total Frequency	Percentage
Product	18	37.50
Product/service	15	31.25
Service	15	31.25
TOTAL	48	100

In the verbal clausal nominalization the researchers investigated, "to which an independent claim is directed" appears to be the adverbial constituent of the main clause nominalization. In this manner, product/service and service make up a similar proportion at

31.25%, whilst product represents 37.5%. Examples (7) to (9) illustrate such findings.

(7) A product <u>to which the independent claim is directed</u>.

(8) The product(s) and/or service(s) <u>to which the independent claim is directed</u>.

(9) A service <u>to which the independent claim is directed</u>.

As can be seen, these examples demonstrate not only 'marked thematic equatives' but also wh-cleft[7]. Based on this observation, the researchers found that rheme in verbal clauses of modern patent language states an authority to its target of product and/or service. In (7), for example, to which the independent claim is directed as rheme and the independent claim located requests for a particular product, a particular product is addressed by "to which the independent claim is directed" where the independent claim is within.

In short, the emergence of nominalization underlines the psychological phenomenon that human beings' verbal behavior (independent claim) is embodied in modern patent language. Further, since the verb-noun collocation 'independent claim + direct' has no other similar collocation in verbal clauses, "to which an independent claim is directed" was of mutual information value with the same rheme but alternative themes.

5.3 Semantic Prosody

As mentioned earlier, a verb-noun collocation has selectional features that associate it with a particular set of semantic contexts. A verbal clause, for example, shows a tendency to occur when a product collocates with 'independent claim + direct.' Based on this, it shows how a verbal clause is regularly found collocated with 'independent claim + direct' that share a semantic similarity-product. In this regard, the semantic context that attracts such a verb-noun collocation is considered 'semantic prosody'. Since the function of semantic prosody is to transfer communicative purposes (Stubbs, 2009:125), the researchers lay their attention on semantic prosody of the verb-noun collocations to further elucidate semantic associations in patent environment of independent claim.

Based on the verb-noun collocations the researchers examined, semantic prosodic associations of the technical terminology independent claim contained in the patent specification are elaborated in Table 10 below.

---

[7] 'Wh-cleft' involves the division and repacking of the information in a clause in two parts (Locks, 1996:238).

Table 10. Semantic prosodic relation of independent claim taggings

Prosodic Type	Semantic Prosody	Total Frequency	Percentage
Innovation	product, present invention	63	25.8
Technology	processing computer, processing device	59	24.2
Service	service	39	16.0
Knowledge	information	34	13.9
Tool	apparatus, database, vehicle	29	11.9
Function	search query, claim	16	6.60
Violation	infringement	4	1.60
TOTAL		244	100

From the corpus-based analysis, verb-noun collocations of independent claim were found to collocate mostly with prosodic type 'innovation' (25.8%), followed by 'technology' (24.2%), 'service' (16%), and 'knowledge' (13.9%), making up nearly 80% of the total. All of these prosodic types imply a positive semantic prosody-patents are important assets of human beings. Based on this, the researchers argue that semantic prosody is the exponent of a special correlation between the semantic structure and syntactic form they were put into. The distribution of the prosodic items, in turn, shows the extent of the syntactic forms expressed by semantic links of the grammar of words. The present study rated those over 20% as high frequency; less than 20% but more 10% as mid frequency; less than 10% as low frequency. It is noted that 1.6% were concerned with infringement. This is of lower percentage but of importance in that the public should draw their attention to the rise of potential perils as 'violation' (infringement), which bring about torts and plagiarism, were overlooked.

The researchers turned their focus on the low frequency level for an instance. In their opinion, aside from prosodic type "violation" which is on the low frequency level discussed earlier, there is a rate of 6.6% verb-noun collocations that co-occur with prosodic type "function" that might elicit the underlying mechanisms of independent claim. The researchers give examples in (10) to (11).

(10) A search query containing information corresponding to the at least one independent claim.

(11) An example of a search or search query, associated with the independent claim directed to the exemplary vehicle locating apparatus, can include the following search words terms and/or connectors.

From the above examples, search query is viewed as the semantic prosody that co-occurs with 'correspond + independent claim' and 'associate + independent claim,' respectively. In (10) and (11), search query is the shared prosody embraced by different collocating verbs. Since semantic prosody is a powerful linguistic device in that it stands for language universality (Lewandowska-Tomaszczyk, 1996:159), the result obtained from the low frequency level further highlights how it is shared by a particular syntactic category of collocating verbs, which may motivate the investigation into different frequency levels for future research.

5.4 Discourse Thematic Referentiality

Chen (2009:1666) proposed a discourse-functional approach "discourse thematic referentiality" to the referential use of NP. He points out such context-dependent referentiality is viewed as thematicity of referents or referentiality in terms of thematic importance of objects in discourse. Based on this, he holds the view that grammatical categories, such as nouns and verbs, are potential functional features to perform the referring function. He lays his attention on the noun group as we lay the focus on how semantic prosody associated with verb-noun collocations. He further emphasizes that the noun group is of genuine importance in that it highly represents thematic referentiality in the context of language use.

In the previous section, semantic prosody is considered referential of thematic importance in the discourse of independent claim. As for the present study, semantic prosody, however, only collocates with certain verbs unusually. Some share the same verbs; some share a unique verb on their own; some have both tendencies. In this section, semantically, we state the intimacy between semantic prosodies and independent claim taggings. Pragmatically, we address semantic prosodies that are referential when they were structured with collocating verbs that highlight the referring functions.

As Table 11 shows, discourse thematic referentiality shows a strong tendency of language specificity. It can be said of true condition in which conditions that must be satisfied by the world if an utterance of a declarative sentence is true. For example, the utterance "There is a cat on the table" is only true if there actually is a table with a cat on it at that time of the utterance (Hurford, Heasley, & Smith, 2007:252). Based on this, discourse thematic referentiality of material clauses can be realized only when the processing computer, processing device, present invention, product/service, search query, information, apparatus, database, or claim is associated with specific verb-noun collocations of independent claim. Once inappropriate elements, such as boy toy or gossip girl appear, it violates the truth condition because it goes with the wrong semantic prosody so as to hinder semantic

presupposition (Levinson, 1983:201).[8] Further, once an inappropriate verb works with semantic prosody, it no longer satisfies the truth condition and infringes on semantic presupposition. For example, processing device only works with 'identify' and once either 'analyze' or 'fall' is adopted, the principle is not cooperated with; discourse thematic referentiality then is cancelled.

Table 11. Discourse thematic referentiality of material clauses

Theme (Semantic Prosody)	Referentiality (Verb)	Discourse (Genre)
processing computer	[+identify], [+be], [+break up], [+contain], [+formulate], [+generate], [+infringe], [+isolate], [+perform], [+process], [+regard], [+store]	independent claim
processing device	[+identify]	independent claim
present invention	[+identify]	independent claim
product/service	[+fall]	independent claim
search query	[+infringe]	independent claim
information	[+identify], [+correspond], [+provide]	independent claim
apparatus	[+identify], [+be], [+utilize], [+store]	independent claim
database	[+identify]	independent claim
claim	[+analyze], [+permit]	independent claim

Of the relational clauses, 'contain' addresses the function mostly as a product/service, information, and service, in turn, becoming thematically referential, as described in Table 12.

Table 12. Discourse thematic referentiality of relational clauses

Theme (Semantic Prosody)	Referentiality (Verb)	Discourse (Genre)
product/service	[+be], [+exhibit], [+contain]	independent claim
search query	[+correspond]	independent claim
information	[+contain], [+regard], [+correspond]	independent claim
service	[+be], [+contain], [+regard]	independent claim

Of verbal clauses, discourse thematic referentiality is maintained when semantic prosodies work with 'direct.'

---

[8] Semantic presupposition is presupposition based on either truth conditional theory or semantic relations, which were defined in terms of semantic feature or atomic concepts.

Table 13. Discourse thematic referentiality of verbal clauses

Theme (Semantic Prosody)	Referentiality (Verb)	Discourse (Genre)
product/service	[+direct]	independent claim
product	[+direct]	independent claim
service	[+direct]	independent claim

As shown in Table 13, product/service, product, and service were referential once they were functioned with 'direct.' Further, 'direct' is specifically used in that it appears in only verbal clauses. The degree of discourse thematic referentiality, therefore, is therefore comparatively stronger than that of other clauses. It appears that product and service are basic prosodies that, when interacting with a semantic trigger, 'direct,' brings about discourse thematic referentiality. Based on clausal nominalization mentioned earlier, in Example (7) ("A product to which the independent claim is directed"), product and 'direct' were essential linguistic components that represent the relatively compositionality fixed relationship of verbal clauses.

In sum, discourse thematic referentiality accounts for how collocating verb, semantic prosody, and independent claim are constructed linguistically. Before closing, it is important to accentuate discourse thematic referentiality, which addresses how lexical units build up modern patent language, providing empirical evidence for the overall characterization of independent claim.

6. Conclusion and Future Work

There has been little investigation into modern patent language in applied linguistics research. Therefore, the present study fills the gap by compiling a patent technical word corpus. The researchers create a patent technical word list regarding frequently used word items of six primary patent areas. Such a word list is significant in that it can help learners expand their vocabulary by displaying the words they should learn. Further, since learners are especially deficient in verb-noun collocations (Chen & Tang, 2004; Liu, 1999), collocational patterns identified in the present study can equip learners with a better sense of verb-noun collocational relationships. For practitioners and researchers, the results of the present study are essential to be incorporated into the English for Occupational Purposes curriculum development.

On the one hand, practitioners and researchers can encourage the application of independent claim as a primer or beginner guide of English patent language. Based on the functional account of independent claim, teachers can show examples by means of clauses as the hidden context. Students can learn how clauses are used in patents under different situations. For example, a product to which an independent claim is directed in Example (7) is

a verbal clause constructed by virtue of collocating verb 'direct' and semantic prosody product in which clausal nominalization occurred. Based on this, teachers can integrate verb-noun collocation 'independent claim + direct' to guide students to notice the overlooked prosodic relations. Moreover, embedded clausal nominalization can be taught for a better understanding of the rhetorical function. Finally, teachers can encourage students to apply and learn other technical vocabulary for the writing of professional patents.

On the other hand, vocabulary teaching needs to take account of semantic prosody (Hunston, 2002:142) because ESL/EFL textbooks or bilingual dictionaries do not explicitly represent the feature of semantic prosody or may provide inappropriate semantic prosodic information that can mislead language learners (Zhang, 2009:10). In this regard, teachers can choose a particular area that students familiar with or feel interested in to encourage the application of semantic prosody to further develop technical vocabulary for the writing of patents. Consequently, functional accounts of independent claim add relatively importance in the teaching of technical vocabulary for the writing of professional patents.

Although the 'independent claim' corpus in this study contains over 4.8 million running words, it is relatively small compared to the PTWC corpus (16 million running words) for Patent English. It is suggested that future works can examine other technical words such as 'dependent claim' or 'beneficial claim' in order to generalize the results.

Due to restricted time, the present study examines contemporary patents over a decade, 2000 to 2009. It is suggested that future work can further probe into different temporal periods so as to provide a more comprehensive point of view for this field.

Further, since the present study aims at exploring language-specific characteristics of independent claim, teachers can measure students' familiarity from functional perspectives. It is suggested that future work can collect students' writing and compare their use of verb-noun collocations, semantic prosody, and other linguistic features. The results may provide in-depth insights into how teachers can help students learn technical vocabulary in the EOLP-based courses.

As the extensive use of generic terms and vague expressions poses a great challenge in patent retrieval (Sarasúa, 2000), it would be essential to research on linguistic specificity of patent lexis for a better understanding of relational lexical semantics in modern patent language. In considering Sheremetyeva's (2003) approach to analyzing patent claim texts with natural language processing (NLP) methodology which improved analyses robustness, our work, in contrast, pinpoints the preliminaries and peculiar associations in patent documentation. Aside from playing a role in modern patent language, the proposed approach and genre-based characteristic analysis is considered influential in bridging ELP to NLP for future research.

References

American Bar Association. (2010). *ABA section of intellectual property law: Introduction*. Retrieved from http://www.abanet.org/intelprop/intro.html

Badger, R. (2003). Legal and general: Toward a genre analysis of newspaper law reports. *English for Specific Purposes*, 22, 249-263.

Ball, C. (1996). *Tutorial notes: Concordances and corpora*. Retrieved from http://www.georgetown.edu/cball/corpora/tutorial.html

Bhatia, V. K. (1993). *Analyzing genre: Language use in professional settings*. New York: Longman.

Biber, D, Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.

Bowles, H. (1995). Why are newspaper law reports so hard to understand? *English for Specific Purposes*, 14, 201-222.

Candlin, C. N., Bhatia, V. K., & Jensen, C. H. (2002). Developing legal writing materials for English second language learners: Problems and perspectives. *English for Specific Purposes*, 21, 299-320.

Chen, C. Y., & Tang, Y. I. (2004). Collocation Errors of Taiwanese College Students: Oral and Written Production. In *Proceedings of the Tenth International Symposium on English Teaching* (pp. 278-286). Taipei: Crane.

Chen, H. J. (2001). Taiwanese EFL learner corpus and interlanguage analysis. In *Proceedings of the Tenth International Symposium on English Teaching* (pp. 288-299). Taipei: Crane.

Chen, P. (2009). Aspects of referentiality. *Journal of Pragmatics*, 41, 1657-1674.

Cheng, L., Sin, K. K., & Li, J. (2008). A discursive approach to legal texts: Court judgement as an example. *The Asian ESP Journal*, 4(1), 14-28.

Chiu, S. H. (2008). WAR metaphor in legal discourse: A reminder of their perils. *The Seventh International Conference on Research and Applying Metaphor*, Cáceres, Spain. May 29-31.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: The MIT Press.

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

Coxhead, A., & Nation, P. (2001). The specialized vocabulary of English for academic purposes. In J. Flowerdew & M. Peacock (Eds.), *Research perspectives on English for Academic Purposes* (pp. 252-267). Cambridge: Cambridge University Press.

Denton, J. (2009). Content vs. concept: Two different focuses in the teaching of Legal English. In *Proceedings of ESP Seminar: English for Legal Purposes* (pp. 4-11). National University of Kaohsiung, Taiwan.

Du, J. (2009). Content and language integration in tertiary education in China: A case study in Wuhan Law College. *The Asian ESP Journal*, 5(1), 61-77.

Dudley-Evans, T., & St. John, M. (1998). *Developments in English for Specific Purposes: A multi-disciplinary approach*. Cambridge: Cambridge University Press.

Feak, C. B., Reinhart, S. M., & Shinshimer, A. (2000). A preliminary analysis of law review notes. *English for Specific Purposes*, 19, 197-220.

González, M., & Vyushkina, E. G. (2009). International cooperation in designing effective methods to prepare non-native EFL teachers for training and assessing Legal English skills. *Georgetown Law Global Legal Skills Conference IV*, Washington D.C., United States. June 4-6.

Haberstroh, J. (2009). The LAW List (The Legal Academic Word List). *Georgetown Law Global Legal Skills Conference IV*, Washington D.C., United States. June 4-6.

Halliday, M. (2004). *An introduction to functional grammar* (3rd Ed). New York: Oxford University Press.

Harris, S. (1997). Procedural vocabulary in law case reports. *English for Specific Purposes*, 16(4), 289-308.

Hayvarert, L. (2003). *A cognitive-functional approach to nominalization in English*. Berlin: Mouton de Gruyter.

Hsieh, S. K. (1998). *Characteristics of legal Mandarin: A corpus-linguistic approach to the criminal law*. Unpublished master's thesis, Fu-Jen Catholic University, Taipei, Taiwan.

Hurford, J. R., Heasley B., & Smith, M. (2007). *Semantics: A course book*. New York: Cambridge University Press.

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.

Hyland, K., & Tse, P. (2005). Hooking the reader: A corpus-based study of evaluative that in abstracts. *English for Specific Purposes*, 24, 123-139.

Hyland, K. (2008). Small bits of textual material: A discourse analysis of Swales's writing. *English for Specific Purposes*, 27, 143-160.

Lee, D., & Swales, J. (2006). A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes*, 25, 56-75.

Lehmann, C. (1988). Toward a typology of clause linkage. In J. Haiman and S. A. Thompson (Eds.), *Clausal combining in grammar and discourse* (pp. 181-225). Amsterdam: John Benjamins.

Levinson, C. (1983). *Pragmatics*. New York: Cambridge University Press.

Lewandowska-Tomaszczyk, B. (1996). Cross-linguistic and language-specific aspects of semantic prosody. *Language Science*, 18 (1-2), 153-178.

Li, J. (2010). Transitivity and lexical cohesion: Press representation of a political disaster and its actors. *Journal of Pragmatics*, 42, 3444-3458.

Liu, C. P. (1999). An analysis of collocational errors in EFL Writings. In *Proceedings of the Eighth International Symposium on English Teaching* (pp. 483-494). Taipei: Crane.

Lock, G. (1996). *Functional English grammar: An introduction for second language teachers*. New York: Cambridge University Press.

McEnery T., & Wilson A. (2001). *Corpus linguistics*. Edinburgh: Edinburgh University Press.

Mellinkoff, D. (1963). *The language of the law*. Boston: Little, Brown and Co.

Modiano, M. (2001). Linguistic imperialism, cultural integrity, and EIL. *ELT Journal*, 55 (4), 339-347.

Mudraya, O. (2006). Engineering English: A lexical frequency instructional model. *English for Specific Purposes*, 25, 235-256.

Nattinger, J. (1988). Some current trends in vocabulary teaching. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 62-82). New York: Longman.

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

Nelson, M. (2006). Semantic associations in business English: A corpus-based analysis. *English for Specific Purposes*, 25, 217-234.

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

Sarasúa, L. (2000). Cross lingual issues in patent retrieval. In *Online Proceedings of the ACM SIGIR 2000 Workshop on Patent Retrieval*. Athens, Greece.

Scott, M. (2008). WordSmith Tools version 5. Liverpool: Lexical Analysis Software.

Sheremetyeva, S. (2003). Natural language analysis of patent claims. In *Proceedings of the ACL 2003 Workshop on Patent Corpus Processing* (pp. 66-73). Sapporo, Japan.

Stubbs, M. (2009). The search for units of meaning: Sinclair on empirical semantics. *Applied Linguistics*, 30(1), 115-137.

Swales, J. M. (1983). Vocabulary work in LSP: A case of neglect. *Bulletin CILA*, 37, 21-34.

Swales, J. M. (2000). Language for specific purposes. *Annual Review of Applied Linguistics*, 20, 59-76.

Tsai, S. I. (2006). *Legal language used in laws of the Republic of China*. Unpublished doctoral dissertation. National Tsing Hua University, Hsinchu, Taiwan.

Tsai, Y. (2008). Supply and demand analysis of patent translation. *Translation Journal*, 12(3), http://accurapid.com/Journal/45patents.htm

World Intellectual Property Organization (2009). *WIPO intellectual property handbook: Policy, law and use*. Geneva: WIPO Publication.

Zhang, W. (2009). Semantic prosody and ESL/EFL vocabulary pedagogy. *TESL Canada Journal*, 26(2), 1-12.