THE USE OF PHRASAL VERBS IN ENGINEERING ENGLISH: A CORPUS-BASED STUDY OF EXTENDED UNITS OF MEANING

LEUNG SZE NING MAGGIE

PhD

The Hong Kong Polytechnic University

2019

The Hong Kong Polytechnic University

Department of English

The use of phrasal verbs in engineering English:
A corpus-based study of extended units of meaning

Leung Sze Ning Maggie

A thesis submitted in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

December 2017

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

Leung Sze Ning Maggie

Abstract

The present study explores the use of phrasal verbs in engineering English to illustrate the implications and significance of a fuller description of meanings of phrasal verbs in specialised contexts.

Phrasal verbs have attracted a considerable amount of attention over the past thirty years due to their extensive use in the English language and their syntactic and semantic characteristics. Previous studies have used different definitions of phrasal verbs. This study adopts an inclusive approach in defining phrasal verbs by including: Type I which is the combinations of a verb and an adverbial particle, Type II which comprises a verb and a prepositional particle, and Type III which is the combinations of a verb followed by an adverbial particle and a prepositional particle. All of the three types of combinations function as single units.

This study adopts a genre-based approach to examine the use of phrasal verbs. The corpus analysed was the 9.2-million-word Hong Kong Engineering Corpus (HKEC) composed of 31 genres of engineering texts collected from the engineering workplaces in Hong Kong. Phrasal verbs were extracted from the corpus using Wmatrix and WordSmith Tools ConcGram is used to perform concordancing as it is able to show all possible configurations in the concordances of a single search.

The most frequent phrasal verbs across the 31 genre-based sub-corpora were compared and examined. In particular, concordance analysis was carried out for the most frequent five phrasal verbs and their other inflectional forms. They were examined in terms of their frequencies, forms and co-selections in different sub-corpora using Sinclair's (1996, 2004) five categories of co-selections. The meanings and forms in which the most frequent

phrasal verbs exhibit show that some meanings and forms of phrasal verbs are specific to engineering English or particular engineering genres, and some meanings are specific to particular inflectional forms.

Through studying the co-selections of phrasal verbs in engineering English, the present study provides a much more thorough description of the meanings and use of phrasal verbs in specialised contexts. It has important implications for teaching and learning English for Specific Purposes (ESP) and professional communication in the workplace, and the design of grammar reference books and dictionaries. It may also make significant contribution to other studies on lexical items and phraseology in different contexts by replicating the methodology adopted in the current study.

Acknowledgements

I would like to express my most sincere gratitude to my supervisors Professor Martin Warren and Professor Winnie Cheng. I deeply appreciate their excellent guidance, insightful comments and continuous encouragement which are of tremendous contribution to my study.

I would also like to thank my colleagues and friends at the Department of English for their support.

Finally, I am deeply grateful to my family for their constant love and care. My mother is no longer with us, but my unceasing love and thanks go to her as well.

Table of Contents

List of Tables

List of Figures

Chapter 1   Introduction

This chapter serves to introduce the current study. It begins with a brief description of the background and relevant terms and concepts to provide a direction for this study. It then describes the research aims and questions. Finally, the chapter ends with an outline of all the chapters of the study.

1.1   Background

Phrasal verbs, such as *put off*, *rely on* and *come up with*, are a sub-type of English multi-word units. They are prominent and frequent in both spoken and written English language (Rundell & Fox, 2005; Gardner & Davies, 2007; Schmitt & Redwood, 2011; Kamarudin, 2013). Most previous research on phrasal verbs focused on analysing their syntactic configurations and semantic categories, or issues from a second language learning perspective. Unlike some sub-types of multi-word units such as idioms which are often fixed in form, phrasal verbs display complex syntactic variation (Schmitt & Redwood, 2011; White, 2012; Dehé, 2000; Ishikawa, 1999; Lindner, 1981; Celce-Murcia & Larsen-Freeman, 1999; Chen, 2013), for example, variation of particle positions and compositional use vs. idiomatic use. In terms of semantics, phrasal verbs are often opaque (Dehé, 2000; Schmitt & Redwood, 2011; Chen, 2013; Garnier & Schmitt, 2015). That means the meaning of a phrasal verb is often not equivalent to the sum of the meanings of the components (Gardner & Davies, 2007; Liu, 2011; Hampe, 2002; Zipp & Bernaisch, 2012; Chen, 2013), such as *give up* and *come up with*. What makes the semantic matter more complicated is the polysemous nature of phrasal verbs. Gardner and Davies (2007) identifid that the most frequent English phrasal verbs had an average of 5.6 meaning

senses. While it is considered crucial to master phrasal verbs in order for learners of English to sound natural and native-like (Siyanova & Schmitt, 2007), the syntactic and semantic complexities of phrasal verbs pose great challenges for learners of English (Schmitt & Redwood, 2011; Chen, 2013; Zarifi & Mukundan, 2014).

Unlike fixed multi-word units, phrasal verbs constitute one of the most productive patterns in the English language as they demonstrate productivity in terms of creating new phrasal verbs and meanings (Bolinger, 1971; Celce-Murcia & Larsen-Freeman, 1999; Campoy Cubillo, 2002; Downing & Locke, 2006; Jackendoff, 2002). This makes the collection of phrasal verbs grow and change continually. Yet, the emergence of new combinations is not random but "forms patterns which can to some extent be anticipated" (Sinclair et al., 1989, p. iv). This evolutionary process has made phrasal verbs of considerable interest to linguists.

The importance of phrasal verbs in English is also recognised as they are perceived to be "the multi-word middle ground" between lexis and grammar which were traditionally viewed and treated as two distinct components of linguistics studies (Zarifi & Mukundan, 2014, p.651; Owen, 1993; Alejo-González, 2010; Kamarudin, 2013). Following Firth, Sinclair and Halliday showed the interdependence of the two components and described them as the axes on the syntagmatic axis on which meaning is created (see Sinclair, 2004). Patterning of phrasal verbs with their distinct syntactic flexibility and semantic complexity makes them "a typical case of the difficulty of establishing the limits between grammar and lexis" (Alejo-González, 2010, p. 52).

Corpus evidence shows that phrasal verbs are particularly prevalent in spoken and informal contexts (Biber, Johansson, Leech, Conrad, & Finegan, 1999; Dempsey,

McCarthy, & McNamara, 2007). They tend to be associated with the description "informal, emotive and slangy" (McArthur, 1992, p. 774), and their single-word equivalents (e.g. *call off* vs. *cancel* and *carry on* vs. *continue*) are claimed to be more appropriate in formal contexts (Greenbaum, 1996; Kovács, 2007). However, it is argued that phrasal verbs may not always be appropriately substituted because phrasal verbs and their single-word synonyms do not necessarily share the same range of meaning and co-selection (Sinclair et al., 1989). In fact, phrasal verbs are used across different registers and genres and in some formal occasions, using phrasal verbs is more appropriate and natural in expressing particular meanings (Campoy Cubillo, 2002; Fletcher, 2005; Kamarudin, 2013). Despite the fact that phrasal verbs have a higher frequency of occurrence in spoken or informal contexts, studies should not be restricted to these contexts (Campoy Cubillo, 2002; Khir, 2012).

The patterning and multi-meaning potential of phrasal verbs suggest the importance of studying phrasal verbs, particularly in specialised contexts where specific forms and meanings of phrasal verbs may be revealed. Empirical studies of this language feature in different specialised contexts are, however, few and far between.

1.2 Variant forms of phrasal verbs: phraseological variation and inflectional variation

A phrasal verb constitutes two elements and has some syntactic flexibilities (a more detailed definition and description of phrasal verbs is given in Chapter 2). Therefore, there is possibility for a phrasal verb to display constituency variation and positional variation. Constituency variation refers to the form when the two words are non-contiguous and other words drop between them, for example, the phrasal verb *carry out*

displays as *carry…out*. Positional variation represents the form when the two words of the phrasal verbs occur in different positions relative to one another, such as *out carry* or *out…carry*.

As the same as single-word verbs, a phrasal verb can appear in several inflected forms (e.g. *carry out, carries out, carrying out* and *carried out*). The inflected forms are often lemmatised, i.e. grouped together, and identified by the base form so that they can be analysed as a single item (e.g. Schmitt & Redwood, 2011). This lemmatisation approach implicitly presumes that all the inflected forms share the same status in terms of frequency and meanings. This approach is particularly obvious and typical in dictionary entries and headwords, in discussions in grammar reference books, and in most computational and corpus research. While this long-held approach seems sensible in the sense that grouping together a set of seemingly closely-related items makes it easier for language users to locate them in works of reference and for language researchers to process and classify a large collection of texts, the assumption that all forms of a particular verb or phrasal verb necessarily share the same use should not be taken for granted (Sinclair, 1991; Tognini-Bonelli, 2001; Hunston, 2003; Knowles & Don, 2004; Newman, 2016). Some studies provide supporting evidence to this position, however, with little thorough analysis on phrasal verbs to show the specific patterning on the form-meaning association.

The current study draws on the views and issues to conduct a corpus-based study on the use of phrasal verbs and their variant forms, with special reference to the professional engineering genres in Hong Kong, by adopting the analytical framework of the extended unit of meaning (Sinclair, 1996, 1998) (see Chapter 2), which takes

particular account of word co-selections. The patterning of the form and meaning association in relation to genre specificity in specialised contexts is of particular interest in the current study. The following section explains the rationale of the study.

1.3 Rationale for the study

1.3.1 Why study phrasal verbs in engineering contexts?

Research has shown that linguistic variation occurs among different genres (Biber, 1992; Kennedy, 2002; Goźdź-Roszkowski, 2011; Bhatia, 2013). As far as phrasal verbs are concerned, they are generally perceived as mainly used in spoken and informal contexts rather than formal contexts. However, phrasal verbs are indeed rather frequently used in some formal specialised contexts (for example, academic writing and political speeches) and may be used to express specific meanings (Campoy Cubillo, 2002; Milizia, 2012).

The use of phrasal verbs in academic contexts has attracted more attention, particularly with focuses on comparison of the use of phrasal verbs between native speakers of English and non-native speakers (Chen, 2013), difficulty faced by learners of English (Akbari, 2009; Kamarudin, 2013, Akbari, 2017), and lists of phrasal verbs for pedagogical purposes (Gardner & Davies, 2007). While these studies provide useful resources in terms of teaching and learning phrasal verbs in general English or the contexts of English for second language learning, the use of phrasal verbs in more formal and specialised contexts is under-explored. Specific use of phrasal verbs is expected to occur when they are commonly used in particular specialised contexts (Campoy Cubillo, 2002).

The engineering industry has been one of the key industries in Hong Kong contributing to the growth of Hong Kong's economy and other sectors as well as creating employment opportunities (Census and Statistics Department, 2017). The engineering profession engages in a wide range of disciplinary and business activities, such as building engineering services, electrical and mechanical engineering services, industrial research laboratory services, computer hardware consultancy and other commercial research and development and testing services (Tsui, 2016). Most activities and people involved in engineering are professionally licensed and governed. Genres involved in the engineering profession, which engineers encounter in their daily workplaces, represent the formal and technical records and communication of their works and duties. Genres in the current study are texts collected from a professional setting, engineering, in this case. Each genre was identified by qualified and experienced engineers who are expert members of their professional community. Even in such formal and specialised genres, phrasal verbs are frequently used. Studying phrasal verbs in engineering genres does not only reveal the patterns in relation to the specific forms and meanings, but also enhance the understanding of the real-world language use in the engineering contexts in Hong Kong.

1.3.2 Why study extended units of meaning?

It has been widely observed and recognised that meanings are not made by individual words but by selecting words in combination (Sinclair, 1991, 2004, 2007, 2010; Biber, Conrad, & Cortes, 2004; Carter & McCarthy, 2006; Cheng, Greaves & Warren, 2006; Greaves and Warren, 2007). This is termed 'the idiom principle' (Sinclair, 1991, 2004) which regards the co-occurrences of words as the result of co-selection rather than open

choices bound only by grammar constraints. One of the major theoretical proposals coming out of corpus linguistics in relation to describing meaning creation is Sinclair's model of extended units of meaning (Stubbs, 2009). The model describes the process of word co-selection as a combination of the five categories of co-selection (Sinclair, 1996, 1998): the invariable core word(s), collocation, colligation, semantic preference and semantic prosody (more details in Chapter 2). The model has gained attentions, but has rarely been foregrounded (Greaves & Warren, 2007). Most studies giving space and attention to this model have focused on some categories of co-selection rather than all the five ones. Milizia (2012) examines some frequent phrasal verbs occurring in spoken political language, but analyses the co-selections only for particular ones.

This study aims to provide a comprehensive account of phrasal verbs by systematically analysing the most frequent phrasal verbs in engineering genres with the five categories of co-selection of the lexical item (Sinclair, 1996). Phrasal verbs themselves are the co-selection of two or three words to create particular meanings. The adoption of Sinclair's (1996, 2004) model to examine phrasal verbs is to go one step further by considering how phrasal verbs co-select with other words for a thorough description of their extended units of meaning and use.

1.3.3 Why study inflectional forms separately

The traditional way of treating and presenting phrasal verbs, whether in research studies or teaching and reference materials, is to use lemma as distinct item or entry. For example, a dictionary definition for the entry of BASE ON is for the whole set of inflectional forms, *base on, bases on, basing on* and *based on*. Another example is the explanation and guidance of usage of a phrasal verb provided in grammar reference books or teaching

materials. Very often, a single explanation is provided for a phrasal verb lemma. While it has some seemingly practical reasons for this approach, it fails to consider the potential variation each inflectional form may exhibit in terms of frequency of use and meanings. In other words, the inflectional forms do not only differ in grammatical form.

Some studies may have addressed this issue. However, they seem scanty and limited examples were selected for research. Few studies have fore-grounded the variation of inflection forms and treated them as part of a potential distinct unit of meaning. Considering the potential variation of different inflectional forms and the contribution to learners' understanding of language, this study aims to provide a full account of phrasal verbs by treating each inflectional form as a distinct item.

1.4 Objectives of the study

The principal objective of this study is to explore the use of phrasal verbs in specialised genres within the engineering profession. Phrasal verbs are generally less common in formal and professional contexts. The study focuses on their use in specialised genres to identify and examine genre-specific phrasal verbs.

It is a widely accepted way to simply list and present phrasal verbs with their lemmas in dictionaries, grammar books, teaching materials or even in different empirical studies. Although additional information about particular inflections is occasionally provided, presenting them in this way implies the assumption that the inflectional forms of the lemma have equivalent status in terms of their frequency of use, usage and meanings. It is argued that this approach should not be taken for granted since there is a

lack of evidence to suggest the equivalent status of inflectional forms of individual phrasal verbs.

This study examines the use of the most frequent phrasal verbs in the 9.2 million word Hong Kong Engineering Corpus (HKEC), a corpus of English based on a range of genres drawn from the engineering profession (Warren, 2010a). In particular, it attempts to achieve the following objectives:

1) To identify the phrasal verbs which are most frequently used within and across the engineering genres in order to find out genre-specific phrasal verbs;

2) To examine and describe the frequencies and co-selections of the most frequent phrasal verbs and their variant forms in engineering genres in order to make a thorough description of the extended units of meaning they express in the specialised contexts and uncover the extent of form-meaning association;

3) To compare and contrast the frequencies and co-selections of the inflectional forms of a phrasal verb lemma in order to determine whether they belong to the same lemma or are different unique phrasal verbs.

1.5 Research questions

To achieve the above research objectives, the study takes a corpus-based approach to identify the most common phrasal verbs across the 31 genres within the HKEC and to address the following research questions:

1) How do phrasal verbs within and across the engineering genres investigated compare in frequencies? What are engineering genre-specific phrasal verbs?

2) To what extent can a phrasal verb and its variant forms be specific to engineering genres in terms of its frequency and co-selection?

3) How are the inflectional forms of a phrasal verb lemma similar to and/or different from each other in terms of their frequencies and patterns of co-selection?

The selected phrasal verbs were analysed and compared across the genre-based sub-corpora in terms of their frequency distribution, usage and the five categories of co-selections, i.e. the extended units of meaning (Sinclair, 1996, 1998), to examine the extent to which the usage of a phrasal verb may be specific to a specialised genre. The other forms of the most frequent phrasal verbs, including the non-contiguous forms and inflectional forms, were also examined in terms of their frequency and extended units of meaning to determine if they are the phraseological variation of the same phrasal verb (in the case of non-contiguous forms), comprising the same lemma (in the case of inflectional forms) or distinct phrasal verbs. The findings are discussed in relation to the nature and communicative purposes of the genres.

1.6 Significance of the study

This study strives to contribute to the research of phrasal verbs with particular reference to specialised contexts. The overall findings of the present study will enhance the understanding of the specific language use in the engineering profession in Hong Kong. Researchers, teachers, students, reference materials providers, engineers and professionals-in-training, who are interested in this important feature in the English language, will become more aware of the specific use of phrasal verbs and their variant forms in the real-world professional specialised contexts. Language teachers, students and

engineering professionals in particular will be informed of the phrasal verbs which are most frequent in engineering genres in Hong Kong, and the meanings that are specific to the genres. The results of the study will also serve as a basis for designing effective language reference materials, such as dictionaries and grammar reference books. It will shed light on the association of phrasal verb lemma and inflected forms.

The research approach adopted in the study also shows a potential approach to data-driven learning. Starting with the most frequent phrasal verbs generated from the corpora and analysing their co-selections based on the concordances, language teachers and learners can develop an in-depth knowledge of the lexical patterning and meaning they discover, which they may not be able to predict, based on the actual text evidence (Wang, 2001).

As far as research into phrasal verbs is concerned, there is not much systematic and thorough analysis and description of phrasal verbs in professional specialised contexts. Using the model of extended units of meaning of the lexical item, this study provides a full account of the use and meanings of phrasal verbs with particular reference to the specific co-texts and contexts. The methodology and analytical framework adopted in the current study will provide insights to future research into phrasal verbs in other specialised contexts. It is hoped that future research will replicate the methodology adopted in the current study to investigate other lexical items and phraseologies in different contexts. There are implications for language studies, English for Academic Purposes and English for Specific Purposes.

1.7 Outline of the study

Chapter 2 presents a literature review of concepts and theories which underpin the study. It also provides a definition of phrasal verbs adopted in the current study and a description of some common syntactic properties and semantic categorisation of phrasal verbs. Chapter 3 presents and explains the data and methodology. It describes the size, source and nature of the corpus. A description of the 31 genre-based sub-corpora is also given. The tools and procedures of the extraction of phrasal verbs from the corpus, and the steps taken to analyse the phrasal verbs are also outlined. In Chapter 4, the findings and analysis are reported and discussed. It starts with the quantitative description: identifies the most frequent phrasal verbs in each genre sub-corpora and compare them across the genres (4.1). Chapter 4.2 to 4.6 analyses and discusses the five most frequent phrasal verbs, namely *based on*, *carried out*, *comply with*, *refer to* and *set out*. Their phraseological variant forms and inflectional counterparts are also analysed and discussed. Finally, Chapter 5 provides the conclusions of the study. It summarises the major findings of the study and discusses the implications in terms of pedagogy, methodology and research. Limitations of the study are pointed out in the final section.

Chapter 2   Literature review

As introduced in Chapter 1, the present study focuses on the use of phrasal verbs in engineering English. This chapter is devoted to a review of the background of the current study, concepts, theories and methodology in relation to studies on phrasal verbs which underpin my study.

2.1 Background

With the advance of computer technology and the first substantial collections of texts in machine-readable format available about fifty years ago, studies of various language features, meanings and structure of unprecedented scope and complexity have been made possible (Conrad, 2000; Sinclair, 2004). These studies are considered as part of corpus linguistics which is the empirical study of language based on computer-assisted methods to analyse large collections of naturally-occurring language texts (Conrad, 2000, Jackson, 2007). It is a 'bottom-up' approach that analyses the evidence from a corpus to find and investigate "probabilities, trends, patterns, co-occurrences of elements, features or groupings of features" (Teubert & Krishnamurthy, 2007, p. 6). For these analyses to be effectively and meaningfully conducted, a corpus must be treated not simply as a collection of language texts but rather as a collection of naturally occurring texts in electronic form that may be representative of a given language or language variety (Sinclair, 1991, 2005; Tognini-Bonelli, 2001; Hunston, 2002).

In corpus linguistics, the common categorisations of corpora are 'general' and 'specialised' (Cheng, 2012, Warren, 2013). General English corpora are for studies which aim to examine the use of a language as a whole, such as the British National Corpus

(BNC) which has 100 million words, Bank of English (BoE) which currently consists of 450 million words and the Corpus of Contemporary American English (COCA) which comprises about 520 million words. While large corpora of general English are useful and often used as reference materials for language learning and research, it has been recognised that "single, monolithic descriptions of English are insufficient" as suggested by the existence and development of the field of research on English for specific purposes (Conrad, 2000, p. 556). Corpus linguistics is growing and research shows that considerable differences can occur across varieties or registers of English (Conrad, 2000), such as British English vs. American English, and academic text vs. newspaper text. Specialised corpora are now being created to investigate a wide range of registers, genres and varieties (Flowerdew, 2004; Cheng, 2012). Specialised corpora are often smaller in size compared to general ones since they are designed to describe patterns of language use in more specific contexts.

Corpus linguistics has since been applied to different language studies, ranging from tracking language change and variation, to the production of dictionaries, teaching and reference materials, and to studying different aspects of linguistics. Lexical studies have flourished due to its perceived importance in language and the convenience corpus software provides. Using corpus tools to search a corpus efficiently reveals the most frequent words and their frequencies in the corpus, and also the information about word formation (Moon, 2010). A basic function is to search for a word to generate and study its concordance, which refers to "a collection of the occurrences of a word form, each in its own textual environment" (Sinclair, 1991, p. 32). In other words, the concordance provides the immediate co-text on both sides of the search item. The co-text, i.e. the surrounding textual environment, provides evidence and resources to describe meanings.

14

Traditionally, people equate the notion of the word to a unit of meaning. Corpus linguistics has shown that meaning is not created by single words but by the co-selection of words (Sinclair, 1987, 1991, 2004; Stubbs, 2001; Hunston, 2002; Teubert, 2004). Traditionally, people equate the notion of the word to a unit of meaning. However, single words are problematic as a result of "their polysemy, their ambiguity and their fuzziness" (Teubert, 2004, p. 176). This can be demonstrated by the fact that even a simple word may have more than one meaning defined in any large dictionary. Teubert (2004, p. 176) provides the example of strike. No fewer than eleven senses may be ascribed to it, and so co-text or a context is essential to the realisation of its meaning. It can be said with some justification that a word "has no meaning on its own" (ibid.). Indeed, words are used in texts with a tendency for some words to co-occur more frequently than others. In more extreme cases, there are words which have no individual existence and are used only in fixed phrases, such as dint in by dint of and sleight as in sleight of hand (Stubbs, 2002, p. 217). Therefore, a unit of meaning relates to the notion of phraseology and is the combination of a word with "all those words within its textual context that are needed to disambiguate this word, to make it monosemous" (Teubert, 2004, p. 174).

2.2 Phraseology

Phraseology is used as a broad notion to cover recurrent co-occurrence of words (Clear, 1993; Greaves & Warren, 2007). It includes multi-word sequences, ranging from idioms, phrasal verbs, lexical bundles, fixed sentences, to collocational frameworks. Some phraseologies comprise recurrent adjoining words, others are examples of non-adjacent word co-occurrence (Cheng, 2012).

The relation and phenomenon of how some words tend to come together more often than others are argued to be following what Sinclair (1991) calls the 'idiom principle' as opposed to the 'open-choice principle'. The open-choice principle suggests that language operate based on a vast number of open choices with the only restriction of the general grammatical rules of English (Sinclair, 1991). The idiom principle is based on a different view that words, whether they are in spoken or written texts, are not used randomly and are not simply bound by the grammar. Rather, language is phraseological in nature and meaning is created by the co-selection of words, i.e. the phrase as a whole rather than the individual parts of it (ibid., 1987, 1991). It may be easier to agree on and recognise fixed or semi-pre-constructed phrases, but Sinclair goes further by including also phrases that are not fixed, i.e. patterns of co-selection. In fact, many phraseologies are not fixed and display a degree of variation (Sinclair, 1991; Warren, 2011; Cheng, 2012; Cheng & Leung, 2012).

2.3 Extended units of meaning

One of the most influential theoretical models in corpus linguistics is Sinclair's notion of 'extended units of meaning', or 'lexical items' (Sinclair, 1996, 1998, 2004). It is a model for identifying and describing meanings and it accounts for the phraseological tendency in language. An extended unit of meaning is comprised of up to five categories of co-selection, two of which are obligatory, and three are optional. The obligatory categories are the core and semantic prosody, and the optional elements are collocation (Firth, 1935, 1957), colligation (Firth, 1935, 1957) and semantic preference.

The five categories of co-selection constitute a holistic model put forward by

Sinclair (1996, 1998, 2004) to identify and describe 'extended units of meanings' (Sinclair, 1996, 2004), or what he called later 'lexical items' (1998). Extended units of meanings represent "an element of meaning which is the function of the item in its co-text and context" (Sinclair, 2004, p. 121). This model of describing the extended units of meaning contributes to a fuller description of meaning as it is concerned with the particular meaning that is created by the combinations of the associated textual elements, i.e. the co-selection (Cheng, 2012). Sinclair introduced at a later stage the term 'meaning-shift unit' (MSU) which he sees more suitable to describe the fact that "the co-selection of the constituents causes a shift in the ambient meaning" (Sinclair, 2007, p. 3).

The notion of collocation has been in existence for a very long time. As pointed out by Hoey (2005), the word collocation was found in the 1928 edition of Webster's New International Dictionary. It was later used by Firth (1957, 1968) as a technical linguistic term: "collocations of a given word are statements of the habitual or customary places of that word in collocational order" (Firth, 1968, p. 181). Generally speaking, it represents the relationship between the node item and individual word-form which frequently co-occurs with it. Sinclair (2004, p. 141) described collocation as "the co-occurrence of words with no more than four intervening words". This span is a widely used measure for collocation as it covers most of the habitual collocates within the same lexical or grammatical constructions. Yet, it risks missing those which are part of the larger grammatical construction but fall out of the span. Hoey (1991) points out that collocation is created by semantic relations within texts over greater distances. In light of this, this study considers possible collocation patterns within a wider co-text.

In addition to lexical collocation, texts exhibit the co-occurrence of grammatical

classes, which is termed colligation (Firth, 1968). Stubbs (2001, p. 65) provided an example from his data where "the word-form *cases* frequently co-occurs with the grammatical category of quantifier in prepositional phrases such as *in some cases*, *in many cases*". In an analysis of the concordance of the word *persistent*, Huston (2007) demonstrated how the syntagmatic patterns have an influence on the meaning or connotation. When *persistent* is used as an adjective followed by a noun, the word is consistently used in a context with unfavourable connotation, whereas when it is used as a predicate, the connotation is no longer negative.

The other optional category is semantic preference which refers to the tendency towards "the restriction of regular co-occurrence to items which share a semantic feature" (Sinclair, 2004, p. 142). It is about the "similarity of meaning regardless of word class" (ibid.). An example of semantic preference cited by Stubbs (2001) is the frequent co-occurrence of the word *large* with other items sharing the semantic feature of "quantities and sizes". The roles of these semantic categories amount to "fine-tuning the meaning and giving semantic cohesion to the text as a whole" (ibid., p. 141).

For the two obligatory categories, the core is "invariable, and constitutes the evidence of the occurrence of the item as a whole" (Sinclair, 2004, p. 141). In the example of *naked eye*, Sinclair (2004) found that the article *the* always occurs immediately to the left of *naked eye*, therefore he pointed out that *the* is an inherent component and the core is thus *the naked eye* rather than simply *naked eye*. Semantic prosody deals with aspects of meaning. In Hunston and Thompson's (2000, p. 5) words, it refers to "the speaker's or writer's attitude or stance towards, viewpoint or feelings about the entities and propositions that he or she is talking about". Thus, semantic prosody is an

evaluation of the meaning of an item. Sinclair put it as trying to give an answer to the question "for what reason was this particular item chosen to express the meaning, rather than a close alternative?" (2010, p. 45). In short, semantic prosody shows "how the rest of the item is to be interpreted functionally" (Sinclair, 2004, p. 34) and thus determines the overall attitudinal, tactical and pragmatic meaning of the item (Sinclair, 2010). It is thus an essential element of meaning making and interpretation, and provides convincing support of how words come together as constituents of meaning (Morley & Partington, 2009). Notwithstanding its critical role, semantic prosody is not realised through the item itself, but, as its name suggests, it is "established through the proximity of a consistent series of collocates" (Louw, 2000, p. 57).

Despite the consistency in the co-occurrence of items, semantic prosody is not easy to identify or determine because it is "a subtle element of attitudinal, often pragmatic meaning" which is "not subject to any conventions of linguistic realization, and so is subject to enormous variation" (Sinclair, 2004, pp. 144-145). Partington (2004, p. 132) also recognises the difficulty and expresses the view that semantic prosody "is much less evident to the naked eye". In the Cobuild corpus, for example, Sinclair (1987) noticed that the subjects which the phrasal verb *set in* takes predominantly refer to unpleasant events or matters, such as "rot", "decay" and "ill-will", it can thus be said that *set in* has an unfavourable semantic prosody. This example has given rise to the general conception with regard to the notion of semantic prosody in that it seems to present a binary distinction between 'good' or 'bad' or a gradable scale of connotations depending on how often an item is used in favourable, unfavourable or neutral context (e.g. Partington, 2004). While this good-bad, or positive-negative distinction is considered to be the basic insight of the concept of semantic prosody by some linguists (for example, Morley &

Partington, 2009; Wei & Li, 2014), it is viewed by some linguists as an over-simplification (for example, Hunston, 2007) since not all items can be evaluated simply based on the concepts of 'good' or 'bad'. For example, the semantic prosody postulated from the concordance of the expression naked eye is 'difficulty', based on the co-occurring words such as "small", "faint", "weak", and "difficult", as in "too faint to be seen with the naked eye", or by the use of a negative as in "it is not really visible to the naked eye" (Sinclair, 2004, pp. 30-35). Such semantic prosody cannot simply be regarded as 'good' or 'bad' as it may be reflecting a fact or phenomenon of the real world. The notion of semantic prosody is subtle and its full meaning can only be realised through reference to describing the discourse function of the item itself and the textual environment all together. All these elements, i.e. the node, the co-text and context, or in other words, the five categories of co-selection just described, constitute to a unit of meaning, rather than the node alone.

Semantic preference and semantic prosody are two distinct concepts, yet, it is sometimes hard to determine the precise relationship between them due to their interdependency (Xiao & McEnery, 2006). This might even be one of the long-standing problems (McEnery & Hardie, 2012). Stubbs (2001, p. 66) recognised that "the distinction […] is not entirely clear-cut". Both concepts seem to be related to the semantic sense of words, but, according to Sinclair (1996, 1998) and Stubbs (2001), semantic prosody is at a further level of abstraction than semantic preference. This is reflected through the respective operations or realisations of the two. Semantic preference "generally remains relatively closely tied to the phenomenon of collocation" (Partington, 2004, p. 150) as it represents the semantic feature shared by the regular co-occurring words, whereas semantic prosody "can sometimes similarly be described as the

connection between two items, but often its effect extends wider, over a considerable stretch of text" (ibid.). Stubbs' (2001, pp. 89-95) concordance analysis of the word *undergo* can be used to illustrate the relationship of the two categories. Based on the investigation of its collocates, *undergo* is found to have a number of semantic preferences, namely "medicine", "change", "tests" and "involuntariness". These semantic preferences, all suggesting that "people generally are forced to undergo something they would rather not" (Partington, 2004, p. 150), combine to form a strong overriding semantic prosody which is unfavourable. Partington summarises the interaction of semantic preference and semantic prosody:

> …the former, preference, contributes powerfully to building the latter, prosody; conversely, the latter dictates the general environment which constrains the preferential choices of the node item. (ibid., p. 151)

Language users first make a functional choice, i.e. semantic prosody, which leads to the selection of the core. The semantic preference, determined by the semantic prosody, then controls the patterns of collocation and colligation (Sinclair, 2004, p. 34).

Another key distinction of the two categories, particularly in the model of the five categories of co-selection, is that semantic preference is an optional item while semantic prosody is an obligatory one. As Sinclair (1996, p. 88) remarked on semantic prosody, "without it the string of words just 'means' – it is not put to use in a viable communication". The main purpose of communication is to express and exchange attitudes, and it is semantic prosody which represents this pragmatic role (Stubbs, 2001).

Regardless of the distinction, semantic prosody and semantic preference are both conceivably the most important concepts in corpus linguistics studies, particularly for

semantic prosody (Wei & Li, 2014). Some linguistic examples have been discussed in the literature with identifiable semantic prosodies, including SET IN (Sinclair, 1987), BENT ON (Louw, 1993), CAUSE, PROVIDE (Stubbs, 1995), UNDERGO (Stubbs, 2001), HAPPEN, SET IN, OCCUR, TAKE PLACE (Partington, 2004), *bordering on* (Schmitt & Carter, 2004), *persistent* (Hunston, 2007), and CHALLENGE (Lin & Chung, 2016). This concept of semantic prosody is also researched into words in languages other than English, such as Italian (Tognini-Bonelli, 2001) and Chinese (Xiao & McEnery, 2006) and cross-linguistic and contrastive linguistic studies (Wei & Li, 2014).

In addition to addressing practical implication for corpus-based contrastive linguistic studies, Wei and Li (2014) investigate four translation pairs of English and Chinese verbs and show that an item may have more than one semantic prosody, thus associating with more than one unit of meaning. As Partington (2004) puts it, "unwavering prosody" such as the one associated with SET IN is highly uncommon. While this may not be surprising, there is a tendency for researchers to treat counter-instances of the so-called default semantic prosody as being exploited by language users for special communicative purposes, such as deploying irony (Louw, 2000; Partington, 2004, 2007; Morley & Partington, 2009). Wei and Li (2014) demonstrate with further evidence that counter-instances of the typical semantic prosody can in fact express different semantic prosodies and constitute different units of meaning. Therefore, they argue that counter-instances should not be "downplayed" but indeed deserve their own description and discussion. Therefore, when different semantic prosodies as well as patterns of co-selection are associated with an item, the most frequent pattern can be referred to as "the Major Patterning"; and other less frequent patterns of co-selection can be called "the Minor Patterning" (Wei & Li, 2014, p. 109). This is an important implication for studies

of semantic prosody and co-selections.

Traditionally there has been a clear distinction made between lexis and grammar in the description of languages owing to the demands of dealing with language systems simultaneously (Sinclair, 2000). Grammar seems to be traditionally prioritised in explaining many features of language (McEnery & Hardie, 2012). Later with the development of computer techniques, particularly computational and corpus linguistics, the approach of separating grammar and lexis in describing a language is abolished. With the emphasis on the the Idiom Principle and phraseological nature of language, a model of language where lexis has a much larger role than in traditional views of language has been proposed. The five categories of co-selection, describing how co-occurrences are selected and the process by which the extended units of meaning are constructed, are an account of such view and model of language. Two closely related approaches of language description do not only emphasise the privileging of the role of lexis, but also unify the systems of lexis and grammar "largely or entirely" (McEnery & Hardie, 2012, p. 143). They are briefly discussed below.

With reference to psycholinguistics, Hoey (2004, 2005) proposed an explanation for the choices of words, or more exactly, the recurrent co-occurrence of words:

> As a word is acquired through encounters with it in speech and writing, it becomes cumulatively loaded with the contexts and co-texts in which it is encountered, and our knowledge of it includes the fact that it co-occurs with certain other words in certain kinds of context. The same applies to word sequences built out of these words; these too become loaded with the contexts and co-texts in which they occur (ibid, 2005, p. 8)

For this, Hoey called it 'lexical priming' which is a psychological phenomenon of

language users' repeated acquaintances, i.e. reading and hearing of, with words in different textual environments and contexts leading to the mental association of words which frequently co-occur. This priming process is an interactive one as a speaker or writer's unintentional reproduction of language as influenced by priming would in turn prime the hearer or reader.

This notion of lexical priming is compatible with Sinclair's model of five categories of co-selection in terms of the different aspects in the priming process. A word or a phrase is said to be primed to co-occur more regularly with certain items than others (collocation), or regularly with particular grammatical categories (colligation). In addition, users of language develop from their previous encounters understanding that a word or phrase is primed to associate with a particular semantic context. For example, Hoey drew on Bastow's (2003) data from a study of US defence speeches in which the phrase *young men and women* has a semantic association with 'compliment' because the phrase is frequently used with expressions showing praise and admiration, as in "bright young men and women", "dedicated young men and women", "very capable young men and women", "talented young men and women", "as impressive as those young men and women are", etc. Hoey (2005) labelled such aspect of priming as 'semantic association' which resembles and grows out of the concept of Sinclair's semantic preference. A further aspect of priming is a pragmatic aspect which Hoey named it 'pragmatic association'. This aspect of priming occurs "when a word or word sequence is associated with a set of features that all serve the same or similar pragmatic functions (e.g. indicating vagueness, uncertainty)" (Hoey, 2005, p. 26). This definition of pragmatic association clearly corresponds to the notion of semantic prosody in Sinclair's model of five categories of co-selection. Using the word *sixty* as an example, Hoey (2005) showed that the word has the

semantic associations of 'units of time', 'units of distance' and 'age', and that it has a typical pragmatic association of 'vagueness'. Hoey (2005) commented on the fuzzy distinction between semantic association and pragmatic association and noted that it is similar to the relationship between semantic preference and semantic prosody.

Hunston and Francis (2000) also take theoretical stance from Sinclair's (1991) idiom principle and units of meaning and proposed 'pattern grammar' as an approach to the grammar of English for language description and pedagogy. In this approach, a pattern is defined as phraseology consisting of grammatical words, parts of speech and types of clause which frequently co-occur with particular lexical items. Examples are *V of n*, *V that*, *ADJ of n*, etc. Patterns in natural language stand out and become more easily identified with the use of corpus concordances. Pattern grammar, despite the term 'grammar', prioritizes the role of lexis and reinforces the idea that patterns and lexis do not co-occur randomly. The co-occurrence of lexis and pattern is associated with particular meanings. For example, Hunston and Francis (2000, pp. 100-101) noticed that the phrase *V way prep/adv* is frequently typically associated with the meaning of "someone uses clever, devious or forceful language to achieve a goal, usually extricating themselves from a difficult situation, or getting into a desirable situation", and that many of the verbs used most frequently in that pattern are semantically related to talking, such as "talk", "negotiate", "lie", "argue" and "wheedle" as in "he talked his way into the post of chief costume designer", "she does not try to lie her way out of trouble" and "argue his way out of tough situations". Therefore, identification of language patterns is compatible with Sinclair's model as the pattern can be considered as a manifestation of colligation. In terms of the semantic preference, those verbs frequently used as collocates share the semantic feature of 'talking' in the example above. The semantic prosody in this case

would be using strategic language to achieve a goal.

This study analyses the five categories of co-selections as it entails careful consideration of the co-text and context in which an item is used and the particular extended units of meaning that appear in that context. The analysis is done by examining the concordance to reveal quality evidence about the language (Sinclair, 1991). Key Word in Context (KWIC) is the most universally used format for concordances and displays a whole line of text with the item under examination appearing in the centre of each line.

Cheng, Greaves, Sinclair and Warren (2009) examine the patterns of co-selection of the two positional variants of the phraseology PLAY/ROLE and found that it is a "meaning shift unit" (Sinclair, 2007) as it has different units of meaning when it displays positional variation, i.e. PLAY/ROLE and ROLE/PLAY. In a study based on a corpus of spoken discourse about the SARS crisis[1], Cheng (2006) adopts this descriptive model of extended units of meaning and provides an account of seven most frequently occurring lexical words in the SARS corpus. Even for a place name which is supposed to denote a more restricted concept, like *Hong Kong* in her study, this model is able to describe the particular textual meaning within and across genres. Examining the concordance of *Hong Kong* in the SARS corpus, Cheng (2006) identifies the semantic preferences of 'business and economy' and 'medical field', and the semantic prosody of 'optimism', therefore constructing the function of *Hong Kong* in such context as 'optimism about recovery in business, economy and medical field' (ibid., p. 332).

---

[1] SARS crisis refers to the outbreak of severe acute respiratory syndrome (SARS) between November 2002 and July 2003 (World Health Organisation, 2003; Department of Health, HKSAR, 2003).

Patterns of word co-selection and phraseological profile were also examined in the contexts of defamation case law (Cheng, Cheng & Li, 2016) and ordinances (Warren & Leung, 2013, 2016) in Hong Kong. These more recent studies further show the significance of studying patterns of co-selections in relation to the understanding of meaning creation and expression, and also the characteristics of the genres.

In the case of phrasal verbs, Sinclair (1991) uses the model to examine them on concordances. A group containing the verb *set* were selected for description, namely *set in, set about, set off, set out and set up*. It was not a detailed examination and description of those phrasal verbs, but points were clearly made about how a phrasal verb has a particular use and characteristic environment for each of its senses and the different patterns of choice distinguish each sense from the other. Concordances of phrasal verbs provide a qualitative account of the use of phrasal verbs in particular contexts (see Milizia 2012, 2013). In Milizia's (2012, 2013) studies of phrasal verbs, three of the five categories of co-selection are analysed: collocation, semantic preference and semantic prosody. These three items are part of the five categories of co-selection. For example, *deal with* in political discourse is found to carry an unfavourable semantic prosody (Partington, 2003; Milizia, 2013) and typical collocates in a corpus of political speeches delivered by British and American politicians are related to 'crisis', 'a H1N1 virus', 'a broken health care system', 'a common threat', 'a budget's deficit', 'criminality', 'corruption' and so forth (Milizia, 2013, p. 149).

2.4 Defining phrasal verbs

In any study or discussion of phrasal verbs, their definition is the first issue to be dealt with and this has attracted considerable attention over the years. Phrasal verbs, which consist of more than the verb itself, are a part of the class of multi-word verbs in many English grammar books and dictionaries (see, for example, Quirk, Greenbaum, Leech, & Svartvik, 1985; Greenbaum, 1996; Biber et al., 1999; Nelson, 2001; Carter & McCarthy, 2006). A phrasal verb is generally defined as a combination of a verb and a particle which functions as a single unit (Carter & McCarthy, 2006; Thim, 2012). Based on how they are formed, phrasal verbs have been labelled differently by different researchers. The most widely known ones are 'phrasal verbs' (Fairclough, 1965; Bolinger, 1971; Dixon, 1982; Courtney, 1983; Dalle, 1983; Quirk et al., 1985; Sinclair, 1989; Darwin & Gray, 1999; Gardner & Davies, 2007; Hampe, 2012; Liu, 2011), 'verb-particle combinations' or 'verb-particle constructions' (Lipka, 1972; Fraser, 1965; Lindner, 1981; Hunter & Prideaux, 1983; Dehé, Jackendoff, McIntyre & Urban, 2002; Villavicencio, 2006), and 'particle verbs' (Zeller, 2001; Dehé, 2002; Elenbaas, 2003; Schneider, 2004; Zipp & Bernaisch, 2012), 'two-word verb' (Meyer, 1975), and 'discontinuous verb' (Live, 1965), among others.

Describing phrasal verbs as the combination of a verb and a particle seems to over-simplify the issue. The difficulty of adopting such a simple definition lies in what is regarded as a particle and how the verb and the particle function in combination. Typically, particles used in phrasal verbs are of two kinds, namely adverbial particles and prepositional particles. Some studies treat both combinations as phrasal verbs (Courtney, 1983; Dalle, 1983; Sinclair, 1989; Halliday, 2004), whereas others recognise only the combination of a verb and an adverbial particle and distinguish this from the other

28

combination type which they typically call a prepositional verb (Bolinger, 1971; Lipka, 1972; Fraser, 1976; Lindner, 1981; Quirk et al., 1985; Biber et al., 1999; Celce-Murcia & Larsen-Freeman, 1999; Hampe, 2002; Villavicencio, 2003; Carter & McCarthy, 2006; Downing & Locke, 2006).

In fact, there is a third type of combination which is a verb followed by both an adverbial particle and a prepositional particle, such as come up with. Similarly, some separate this type and call them 'phrasal-prepositional verbs' (Greenbaum, 1996; Biber et al., 1999; Carter & McCarthy, 2006), but some take an inclusive approach in including this type as a 'phrasal verb' as well (Sinclair et al., 1989; Sinclair et al., 1990; Halliday, 2004). There are those who do not use the term 'phrasal verb' or 'verb-particle combination' at all. Huddleston and Pullum (2002) do not consider the combinations of a verb and an adverbial particle as syntactic constituents, so they reject the term 'phrasal verb' in *The Cambridge Grammar of the English Language*. Instead, they use the term "'verb-particle-object' construction" (Huddleston & Pullum, 2002, p. 280) and "verb + intransitive preposition idioms" (ibid., p. 286). However, they do use the term 'prepositional verbs' to describe the combinations of a verb and a prepositional particle. In spite of the different terms used or the varying coverage of the terms, the similarity and close relations of the three types of combinations have been the subject of different studies and analysis. In addition, those who differentiate between the three types typically discuss them in close proximity in their grammars.

Consider examples (1) to (6) which are drawn from the British National Corpus (BNC, 2007) below:

(1) All you have to do is **pick up** a pen, or phone and send your first ten pounds.

(2) Smoking a cigarette would make me want to **throw up**.

(3) This booklet gives you guidance on how to **deal with** the problem in the workplace

(4) We will **refer to** such issues in Chapter 8.

(5) You can **look forward to** a substantial return on your investment.

(6) People will **put up with** hardship as the price of freedom.

It is easy to identify a 'phrasal-prepositional verb' as it consists of more than one particle but for examples (1) to (4), at first sight, it is not easy to distinguish different types of combinations, because many of the adverbial particles and prepositional particles overlap in terms of the form. Some scholars provide a list of such overlapping, or dual-functional, particles. The lists may not include identical or the same number of particles but typical ones serving both adverbial and prepositional functions include (see, for example, Lindner, 1981, p. 3; Quirk et al., 1985, p. 1151; O'Dowd, 1998, p. 12):

about, above, across, after, along, around, by, down,

in, off, on, out, over, through, under, up

In other words, all these words which can be used as adverbial particles can also be used as prepositional particles (Biber et al., 1999).

Although it is not easy to determine whether it is more widely accepted for the notion of 'phrasal verbs' to include only the combinations with adverbial particles, or include also the other two types, most experts, if not all, would recognise that the three types are very similar. First, they behave syntactically like a single unit (Quirk et al., 1985; Greenbaum, 1996; Biber et al., 1999; Carter & McCarthy, 2006). Irrespective of which

type of combination, what follows the verb is "morphologically invariable" and therefore given the status of a particle (Quirk et al., 1985, p. 1150), rather than simply as an adverb or a preposition. The particles "follow and are closely associated with verbs" (ibid.). This is demonstrated by the example given by Quirk et al. (ibid.):

…the sequence *disposed of* also functions as a single unit, such that for some purposes the sentence can be reasonably divided into:

    [We] [disposed of] [the problem].
rather than into:

    [We] [disposed] [of the problem].

Thus, it is shown that the verb and the particle together combine to form a syntactic unit. This is distinct from other grammatical units, for example, a verb plus a preposition. In some cases, the first part is not even used independently as a verb, such as *sum, tamper* and *zero* as in *sum up*, *tamper with*, and *zero in on* (Sinclair et al., 1990, p. 162). This further proves the status of the combinations of a verb and particle(s) as a syntactically single unit.

Second, there is a close semantic relationship between the verb and the particles, and together they form a single unit of meaning (Carter & McCarthy, 2006). Very often, the meanings of the combinations cannot be derived from the usual individual meaning of the verb and the particle(s) because their individual meanings change or extend when combining together and may create a new meaning (Sinclair et al., 1990; Biber et al., 1999). Thus, many combinations of a verb and particle(s) have a single-word synonym which is semantically unrelated to both the verb and the particle:

    carry on = continue (Greenbaum, 1996, p. 279)

31

look at = observe (Biber et al., 1999, p. 403)

put up with = tolerate (Carter & McCarthy, 2006, p. 430)

For some combinations, the verb may retain its usual meaning, but the particle adds some meaning to the whole combination, for example, *spread out* (Sinclair et al., 1989, p. v). Third, from a functional perspective, as pointed out by Halliday (2004) who treats all the three types as 'phrasal verbs', all the three types of the combinations represent a single process. In short, the three types of combinations are similar in different respects. Even sources which treat the three types as different concepts tend to put them under the same larger category of multi-word combinations based on similarity, and discuss them in close relation to each other (see for example, Quirk et al., 1985; Greenbaum, 1996; Biber et al., 1999; Carter & McCarthy, 2006).

This study adopts the term 'phrasal verbs' in an inclusive sense in line with Sinclair et al. (1989), Sinclair et al. (1990) and Halliday (2004), by covering the three types of combinations which are listed as follows:

Type I: a verb + an adverbial particle (e.g. *pick up* and *throw up* in examples (7) and (8))

   Examples: (7) All you have to do is **pick up** a pen, or phone and send your first ten pounds.

   (8) Smoking a cigarette would make me want to **throw up**.

Type II: a verb + a prepositional particle (e.g. *deal with* and *refer to* in examples (9) and (10))

   Examples: (9) This booklet gives you guidance on how to **deal with** the problem in the workplace

(10) We will **refer to** such issues in Chapter 8.

Type III: a verb + an adverbial particle + a prepositional particle (e.g. *look forward to* and *put up with* in examples (11) and (12))

Examples: (11) You can **look forward to** a substantial return on your investment.

(12) People will **put up with** hardship as the price of freedom.

In addition to the conceptual consideration, there is a practical advantage of adopting the inclusive sense of 'phrasal verbs'. Without setting too many parameters, it may be possible to uncover all possible phrasal verbs which may or may not have come to the knowledge of language users. This is because using an inclusive definition means simply extracting any co-occurrences of verbs and particles without pre-specifying limitations as to which particles should be included and which should be excluded.

That being said, it does not mean that a simple search for all combinations of a verb and a particle will immediately result in a comprehensive list of phrasal verbs because the resulting list will probably include phrasal verbs and free combinations. Free combinations may share the same structure as phrasal verbs, but they do not have the same grammatical and semantic status. As the term 'free combinations' suggests, each element, i.e. the verb, and the adverb or the preposition, carries its individual meaning and their combination is based on independent semantic choice (Quirk et al., 1985). Therefore, they can be syntactically viewed as a verb and an adverb, for example *go back*, or a verb followed by a prepositional phrase which functions as an adverbial such as *live in* (Biber et al., 1999).

In the subsequent sub-sections, the characteristics of each type of phrasal verbs are

briefly described in terms of their syntactic features. Unlike many other studies of phrasal verbs, the focus of this study is not to examine and characterise phrasal verbs based on their syntactic or semantic features. However, it is essential to provide a brief account of the phrasal verbs that are potentially included in the current study.

2.5 Syntactic characteristics of phrasal verbs

Based on the combination of the verb and different types of particles, there are three main types of phrasal verbs.

2.5.1 Type I: combinations of a verb and an adverbial particle

There are two sub-types of combinations with a verb and an adverbial particle. One is intransitive, the other is transitive (Quirk et al., 1985; Greenbaum, 1996; Biber et al., 1999; Carter & McCarthy, 2006).

Intransitive Type I phrasal verbs refer to those which consist of a verb and an adverbial particle. They do not take an object, for example, *get on* as in *how are you getting on* (Quirk et al., 1985, p. 1152) and *fit in* as in *he just doesn't fit in* (Biber et al., 1999, p. 406). The combination of the verb and the adverbial particle changes the distinct meaning of the individual elements and a new extended unit of meaning emerges.

As for free combinations such as *walk past* as in *he walked past* (Quirk et al., 1985, p. 1152) and *go in* as in *if this was new, I wouldn't let people go in* (Biber et al., 1999, p. 404), the meanings of the verbs and the adverbs or the prepositions are the same as those when they are used on their own or when used in other combinations. Although

idiomaticity, or the transparency or change of the meanings, is not the only consideration in recognising phrasal verbs, this does seem to be the main features of intransitive Type I phrasal verbs. Some exceptions might be determined based on the frequency of using the elements in combination. Corpus evidence shows that certain combinations occur more frequently than others, for example, in the combination *fight back*, the two elements have almost the same meaning when they are used separately, but the words *fight* and *back* have a higher frequency of co-occurrence in the corpus data than *fight* with other particles (Sinclair et al., 1989). Considering the high tendency for these two words to come together as a phrase and the notion of phraseology, we may include more fixed combinations, such as *fight back*, as phrasal verbs.

Transitive Type I phrasal verbs have the same structure but takes a noun phrase as a direct object (Biber et al., 1999, p. 408; Huddleston & Pullum, 2002, p. 280):

(13a) Margotte rarely **turned on** the television set.

(14a) She **brought down** the bed.

The main distinction from free combinations is the flexibility of the position of the particle in a transitive Type I phrasal verb, i.e. the particle can either precede the direct object or follow it (Bolinger, 1971; Fraser, 1976; Quirk et al., 1985; Sinclair et al., 1990; Greenbaum, 1996; Biber et al., 1999; Jackendoff, 2002; Thim, 2012). Thus, the two examples above can be rewritten as follows:

(13b) Margotte rarely **turned** the television set **on**.

(14b) She **brought** the bed **down**.

Free combinations in similar structures, such as she brought the bed downstairs (Huddleston & Pullum, 2002, p. 280), do not have the same syntactic characteristics. The word downstairs is an adverb but not a particle, so it has to follow the direct object.

The two forms of transitive Type I phrasal verbs with different positions of the direct object are referred to by different terms, for example, the continuous order (examples (13a) and (14a)) vs. discontinuous order (examples (13b) and (14b)) (Dehé, 2001, 2002), the left-hand position of the particle vs. the right-hand position (Jackendoff, 2002), and the joined order vs. split order (Lohse, Hawkins & Wasow, 2004). This characteristic of the adverbial particle to have a variable position in Type I phrasal verbs also distinguishes them from Type II with a prepositional particle (Fraser, 1976).

The position of the particle may appear to be totally optional. Nonetheless, the nature of the direct object has an impact on the position. When the direct object is a pronoun, the discontinuous order is required (Sroka, 1962, 1972; Live, 1965; Bolinger, 1971; Lindner, 1981; Celce-Murcia & Larsen-Freeman, 1999; Jackendoff, 2002; Downing & Locke, 2006; Thim, 2012), as in:

(15) For example, a television continues to draw a little power after the user **switches** it **off** with the remote control. (HKEC)

This can be explained in terms of the news value of the direct object (Bolinger, 1971; P. Chen, 1986; Dehé, 2000, 2001, 2002). Dehé (2002) explains that

> If the nominal object has been mentioned before in the discourse or can be inferred from the preceding context the discontinuous construction is preferred, whereas the continuous order is preferred in cases where the object introduces

new information into the context (p. 78).

A pronoun denotes an item or entity which is mentioned before or known already, it carries lower news value, and hence it typically takes the position in between the verb and the direct object in a discontinuous order (Lindner, 1981; Dehé, 2001, 2002). In example (15), the pronoun 'it' refers to 'a television' which occurs earlier in the sentence, therefore it is a known entity to the readers and carries less news value.

Type I phrasal verbs can produce action nominals. For example (Lindner, 1981, p. 17):

(16) His **looking up** of the information pleased the teacher.

(17) John's **thinking over** of the example assisted him in passing the exam.

In examples (16) and (17), look up and think over are phrasal verbs with an adverbial particle which can be nominalised in this way by the addition of the suffix –*ing* to the verb, i.e. changing the verbs *look* and *think* to gerunds. The position of the particle is restricted to being contiguous with the verb on the right side. It is observed that when the phrasal verbs are nominalised, a prepositional complement in the form of *of + NP*, i.e. *of the information* and *of the example* in examples (16) and (17) respectively, always follows the phrasal verbs (Fairclough, 1965; Fraser, 1965; Bolinger, 1971; Jackendoff, 2002; Thim, 2012). In some studies, this is not described as a prepositional complement, but rather just 'of-NP' (Jackendoff, 2002, p. 72) or in Lindner's (1981) words "the *of* in an action nominalization can appear between a particle and a following noun, but not between a preposition and a following noun" (p. 16). The noun phrase is originally the direct object of the phrasal verb, for example, the nominal in (16) can be re-written as:

He looks up the information…

Thus, the noun phrase together with *of* becomes the counterpart of the direct object in nominalizations (Jackendoff, 2002), nonetheless, "it behaves as a PP (prepositional phrase) with respect to particle position" (Jackendoff, 2002, p. 72).

Another feature of Type I phrasal verbs is the insertion of adverbial modifiers. They can be inserted in Type I phrasal verbs, but it is only possible in a discontinuous order, with the adverbial modifier being placed between the direct object and the particle. The following examples are from Jackendoff (2002, p. 71):

(18a) Bill **brought** the wagon right **back**.

(18b) Bill **brought** (*right) **back** the wagon.

(19a) Please **shut** the gas completely **off**.

(19b) Please **shut** (*completely) **off** the gas.

(20a) We **turned** the situation partway **around**.

(20b) We **turned** (*partway) **around** the situation.

Highlighting the modifiers and the particles in the examples, Jackendoff (2002) emphasises that the particles are bonded with the modifiers in these forms. The modifiers with an asterisk in examples (18b), (19b) and (20b) are inappropriate insertions.

Certain phrasal verbs of this type allow the fronting of the particle as in the following (from the BNC, in Thim, 2012, p. 24):

(21) A tap is turned, a pressure of 120 pounds per square inch applied to the piston, and

up goes the end of the coach as easily as though it were a doll's house.

This characteristic is rather unusual and is described as "one of the noticeable exceptions from subject-verb order in declarative sentences in present-day English" (Thim, 2012, p. 24). The particle immediately precedes the verb and the verb is followed by the subject which is a full noun phrase (Thim, 2012). However, if the subject is a pronoun, it takes the ordinary position of preceding the verb as in (Thim, 2012, p. 24):

(22) (and…) **up** it **goes**.

In fact, some sources regard this feature as a sign of a free combination (Quirk et al., 1985) since Type I phrasal verbs having this feature are often literal in meaning. In other words, both the verb and the particle retain their usual individual meanings, and thus, they are often treated as free combinations rather than phrasal verbs. Nevertheless, as pointed out earlier, some experts regard more frequent combinations as phrasal verbs. It is not always easy to draw an absolute distinction between phrasal verbs and free combinations. Therefore, for those who include frequent combinations as phrasal verbs (for example, Sinclair et al. (1989) include *fight back*, *go off* and *go up*), such phrasal verbs may display the feature of fronting the particle.

2.5.2 Type II: combinations of a verb and a prepositional particle

Type II phrasal verbs which consist of a verb and a prepositional particle can also be sub-classified into two sub-types. But unlike Type I phrasal verbs, all Type II phrasal verbs take at least a noun phrase as an object (Biber et al., 1999). So the sub-categories of this type of phrasal verbs do not differ in terms of the existence of a direct object, but rather

the number of objects. Quirk et al. (1985) call them 'Type I prepositional verbs' and 'Type II prepositional verbs'; Biber et al. (1999) use 'Pattern 1' and 'Pattern 2' to distinguish the two. This study adopts Greenbaum's (1996, p. 280) terms calling one sub-type 'monotransitive' and the other sub-type 'doubly transitive'.

Type II mono-transitive are those which take one noun phrase in the structure, for example *account for* as in *I can't possibly account for it* (Greenbaum, 1996, p. 283) and *come across* as in *you said you'd already come across that sort of stuff in the library* (Carter & McCarthy, 2006, p. 435). The structure of this sub-category resembles those of Type I-transitive, but the noun phrase of this sub-category is restricted to being placed after the particle. The noun phrase is not considered as the direct object of the phrasal verbs, rather, it needs to be introduced by the prepositional particle, thus it is sometimes called a prepositional object (Quirk et al., 1985; Sinclair et al., 1990; Greenbaum, 1996; Biber et al., 1999). The other sub-category doubly-transitive takes two noun phrases in the structure. To be exact, this sub-type of phrasal verbs consists of a transitive verb taking a direct object and a prepositional particle taking a prepositional object (Quirk et al., 1985; Sinclair et al., 1990; Greenbaum, 1996; Carter & McCarthy, 2006). Examples are *set on* as in *the farmer threatened to set his dogs on them* (Sinclair et al., 1990, p. 169) and *base on* as in *McGaughey bases his prediction on first-hand experience* (Biber et al., 1999, p. 414). The first noun phrase, such as *his prediction* in the previous example, is the direct object of the verb *bases*, and the second noun phrase which follows the particle, i.e. *first-hand experience*, is the object of the preposition or prepositional object. Most of these combinations allow passive forms with the direct object taking the subject position, for example, the previous example can be rewritten as *his prediction is based on first-hand experience*.

This sub-category may be seen as rather confusing as to where the boundary should be drawn to distinguish them from free combinations. The confusion or the difficulty lies in whether the preposition and its following noun phrase functions as the prepositional particle and its object, or whether they form a prepositional phrase functioning as an adverbial (Biber et al., 1999). Biber et al. (1999) suggest that the determining criterion is to see what the noun phrase following the preposition refers to. If it refers to a place or time, the structure is a free combination; but if the noun phrase identifies a person or thing, the structure is a phrasal verb (prepositional verb in Biber et al.'s term). However, this criterion seems to be inadequate in reflecting the special grammatical and semantic status of a phrasal verb, which emphasizes the bond between the verb and the particle as they form a single unit of meaning which has been discussed above. For instance, an example of a phrasal verb given in Biber et al.'s (1999) is *used in* as in *they are, however, widely used in the preparation of special cakes*, because the noun phrase following *in* identifies a thing rather than a place or time. Still, the extent to which the verb *used* and the preposition *in* extend their meanings to form a single unit, or to represent a single 'process', seems unconvincing. Nevertheless, this does not mean that the whole idea is rejected. It is agreed that when the noun phrase following the preposition refers to a place or time, the preposition and the noun phrase form a typical prepositional phrase denoting location or time. In addition to location and time, there are other types of prepositional phrases denoting other functions such as manner, purpose, accompaniment, role, matter, etc. All of these are regarded as circumstantial elements which are treated independently (Halliday, 2004). It is important to note that "a phrasal verb is a single Process, rather than Process plus circumstantial element" (Halliday, 2004, p. 352), and this is why a verb and a preposition is distinct from phrasal verbs. Thus, in

this study, while *account for* is included as a phrasal verb, *use for* is not because it is considered to be a verb and a preposition of purpose.

Type II phrasal verbs mainly differ from free combinations in terms of the choice of the preposition. A preposition is not chosen for its own usual meaning but rather is determined by the verb. Phrasal verbs are recognised as a case of "the verb GOVERNING the preposition" (Quirl et al., 1985, p. 1159). Consider the following example (adapted from Quirk et al., (1985, p. 1159)):

 (23a) I *accused* him *of* the crime.

(23b) I *accused* him **for* the crime

The choice of the prepositional particle *of* is made by reason of the verb *accused*. Although the meaning of example (23b) may also be understandable, it is regarded as ungrammatical because it is fossilized to combine *accuse* with *of*.

For this type of phrasal verb, the form of the object has no effect on the position of the particle. In other words, unlike Type I phrasal verbs, even if the object is a pronoun, it still follows the particle of the phrasal verb:

(24) If you would like to **dispose of** them, please note that spent HCFCs are classified as chemical waste…

Unlike Type I phrasal verbs, action nominalisations are not acceptable for this type of phrasal verbs. Thus, the following two examples are not grammatical (Lindner, 1981, p. 17):

(25) *Her **looking at** of the table surprised the storekeeper.

(26) *His **arguing with** of John disrupted the conversation.

Adverbial modifiers can be inserted in this type of phrasal verb. In such cases, the adverbial modifiers may precede the particle. For example:

(27) The degree of success in each project **depends** heavily **on** the nature of the project and the partnering management efficiency. (HKEC)

(28) However, if circumstances indicate that the potential interviewee may be reluctant to **talk** openly **about** the complaint, the investigator should arrange a private interview on another occasion. (HKEC)

Preposed constructions are also possible for this type of phrasal verbs. Instead of fronting only the particle, the whole phrase of the particle and the object noun phrase can occur in a sentence-initial position as in (Fraser, 1976, p. 2; Hampe, 2002, p. 16):

(29) **In** the street, the man **reeled** as if drunk.

(30) **At** the photo he **looked**.

In some cases when the noun phrase is absent, the particle is fronted with a relative or interrogative pronoun (Legum, 1968; Lindner, 1981), for example:

(31) …their main purpose is to provide standards for materials, **with** which suppliers must **comply**. (HKEC)

Also consider the following two examples with interrogative pronoun (Celce-Murcia & Larsen-Freeman, 1999, p. 430):

(32) **About** what does he **write**?

(33) **At** what did Peter **look**?

The referent of the relative or interrogative pronoun is a noun phrase, hence, *with which,*

about what, and *at what* in examples (31) to (33) have the same structure functionally as

examples (29) and (30). The form of questions in (32) and (33) are grammatical even

though they may seem somewhat unnatural (Celce-Murcia & Larsen-Freeman, 1999).

2.5.3 Type III: combinations with two particles

This type of phrasal verb consists of three words: a verb followed by an adverbial particle

and a prepositional particle (Quirk et al., 1985; Sinclair et al., 1990; Greenbaum, 1996;

Biber et al., 1999; Carter & McCarthy, 2006). They can also be classified into two sub-

types which are referred to in different terms, for example, 'Type I' and Type II' (Quirk et

al., 1985), 'intransitive' and 'transitive' (Sinclair et al., 1990) and 'pattern 1' and 'pattern

2' (Biber et al., 1999). This study adopts the terms from Greenbaum (1996):

'monotransitive' and 'doubly transitive'.

As with Type II phrasal verbs, the noun phrase following the monotransitive Type

III phrasal verbs is the object of the prepositional particle (Sinclair et al., 1990;

Greenbaum, 1996). Thus, *an innovative design* in example (34) is not the direct object of

the phrasal verb *come up with*, but the prepositional object of the particle *with*.

(34) Lambeth has **come up with** an innovative design for the project. (HKEC)

Doubly transitive refers to those verbs which take two noun phrases. The first noun

phrase which is the direct object comes between the verb and the first particle (Quirk et

al., 1985; Sinclair et al., 1990; Biber et al., 1999), for example (Carter & McCarthy, 2006, p. 436):

(35) He was just **taking** his frustration **out on** me, shouting at me and stuff.

In this example, *his frustration* is the direct object of the verb, and *me* is the prepositional object introduced by the particle *on*. While this is the prevalent structure of doubly transitive Type III phrasal verbs, Greenbaum (1996) states that in some of these phrasal verbs, the direct object can come between the two particles, for example (Greenbaum, 1996, p. 286):

(36) I would like to **take up** that issue **with** you.

Overall, there are only a small number of doubly transitive Type III phrasal verbs (Sinclair et al., 1990; Biber et al., 1999; Carter & McCarthy, 2006).

Fronting of the particle is possible for Type III phrasal verbs. The prepositional particle can be fronted in a relative clause (Greenbaum, 1996; Huddleston & Pullum, 2002), for example (Greenbaum, 1996, p. 286):

(37a) I've got the French written paper on Thursday which I'm not **looking forward to** at all.

(37b) I have the French written paper on Thursday **to** which I am not **looking forward**.

Greenbaum (1996) points out that the separation of verb and the two particles is not common in most of the cases, because there is a close link between them.

In practice, it is not always easy to make a clear-cut and absolute distinction between phrasal verbs and free combinations (Quirk et al., 1985; Biber et al., 1999;

Halliday 2004). Not only because there are often exceptions or fuzziness in the syntactic considerations, but also because of the productive nature of combining verbs and particles to form new units of meanings (Celce-Murcia & Larsen-Freeman, 1999; Downing & Locke, 2006). Phrasal verbs are constantly growing in number and evolving (Sinclair et al., 1989). Setting aside the fact that phrasal verbs are defined differently in different studies, even if the same definition or same set of criteria is used, it does not make the issue any simpler because some combinations may not have been recognised as phrasal verbs in the past or in certain contexts, but they may be in other contexts or with the passage of time.

2.6 Semantic categories of phrasal verbs

Many studies on phrasal verbs describe and classify them according to semantic categories. This section briefly describes the most common semantic categories of phrasal verbs.

Phrasal verbs, which include the three types, are described as behaving as a single unit of meaning (Biber et al., 1999; Carter & McCarthy, 2006). Quirk et al. (1985) state that a phrasal verb functions semantically as a single unit since it "acts as a single word lexically or syntactically" (p. 1150). For this general perception of the semantic and combinatorial characteristic, phrasal verbs are traditionally viewed to be associated with the term 'idiomatic'. Hampe (2002) points out that many scholars use idiomaticity to determine what they regard as 'phrasal verbs' which exclude 'free combinations' or what others call 'literal' phrasal verbs (see, for example, Palmer, 1965; Live, 1965; Fraser,

1976; Dixon, 1982; Quirk et al., 1985). However, this dichotomy has little influence on the current study because 'literal' phrasal verbs tend to occur frequently, carrying particular fixed meanings. Such verbs are unlikely to evolve new meanings. A closer investigation into the semantics of phrasal verbs finds that they express a wide range of meanings which can be expressed on a continuum ranging from non-idiomatic at one end to highly idiomatic on the other (Sinclair et al., 1989; Thim, 2012). In the literature, phrasal verbs are usually categorised based on their meanings along the continuum.

There is no absolute agreement on the semantic classification of phrasal verbs. Jackendoff (2002) has enumerated six types, namely, idiomatic combinations, combinations with directional particles, combinations with aspectual particles, the time-*away* constructions, *V/N-d out* constructions, and "*his heart out* family of constructions". However, it is assumed that some of these types may be grouped and classified as sub-types of the more broadly defined threefold classification model which is typically adopted by other scholars (see, for example, Quirk et al., 1985; Fraser, 1976; Bolinger, 1971; Celce-Murcia & Larsen-Freeman, 1999; Thim, 2012). Thim (2012) summarises and describes a three-way classification (shown in Table 1 below), namely compositional constructions, aspectual constructions and idiomatic constructions. The first two constructions can be identified by the types of particles they combine with, namely directional particles and aspectual particles respectively. The classification in fact subsumes compositional constructions and aspectual constructions under a larger category of compositional combinations as contrasted with non-compositional combinations, i.e. idiomatic constructions (Thim, 2012). This threefold classification seems long established; yet, there is an inherent limitation in adopting this model to phrasal verbs defined with a broad sense, because the studies using this classification examine only phrasal verbs with

an adverbial particle, i.e. Type I phrasal verbs in this study. Type II phrasal verbs tend to be treated as idiomatic.

Table 1: A three-way semantic classification of verb-adverb combinations

Category 1	Compositional constructions
Category 2	Aspectual constructions
Category 3	Idiomatic constructions

Drawing on the *COBUILD* database, Sinclair et al. (1989) list and provide meanings for over three thousand phrasal verbs which cover all three types of phrasal verbs, and categorises them into four semantic types. The categories are not named in his work, but based on the descriptions given, three of the four categories appear to be very close to those in Thim's (2012) three-way classification. Sinclair describes the fourth category in terms of the tendency of the verb to combine with a particular particle rather than the semantic senses contributed by the two items. It can be argued that the fourth category is not a semantic category at all and this is discussed below. For the purpose of identification, the fourth category is referred to here as 'habitual constructions'. For the other three categories, due to their resemblance to the widely recognised three-way classification, the same terms are used. Table 2 shows the four categories.

Table 2: Four semantic categories of phrasal verbs

Category 1	Compositional constructions
Category 2	Aspectual constructions
Category 3	Idiomatic constructions
Category 4	Habitual constructions

Despite the fact that Sinclair et al. (1989) use the same broad definition of phrasal verbs, the viability of this four-category classification needs further attestation because it has no exemplification on whether all the four categories apply to all three types of phrasal verbs, or whether a certain category only covers a particular type of phrasal verb. Nevertheless, the aim of the current study is not to analyse and classify phrasal verbs based on these semantic categories. The four-category model is discussed briefly below in an attempt to clarify certain characteristics of phrasal verbs.

2.6.1 Compositional constructions

The term 'compositional' suggests the nature of this type in which the meanings of the phrasal verbs can be interpreted as the sum of the respective meanings of each of the two components, i.e. the verb and the particle. In other words, the meaning of the phrasal verb is transparent and is easily predicted. Because of their compositional nature, such verbs are sometimes referred to as 'literal' phrasal verbs. These constructions are typically comprised of a verb and a directional particle, for example (the BNC):

 (38) They had to **carry** a table **out** with them, which was set up the required ten metres

 from the shed wall, and the pistols were laid out, and the targets pinned up.

(39) I just had to **put in** my ten pence.

Thim (2012) demonstrates the formation nature of the compositional category with the examples from Jackendoff (2002) and uses them in the following example (Thim, 2012, p. 14):

(40)

$$\text{George} \left\{ \begin{array}{c} \textbf{tossed} \\ \textbf{took} \\ \textbf{put} \\ \textbf{carried} \\ \textbf{threw} \end{array} \right\} \text{the food} \left\{ \begin{array}{c} \textbf{up}. \\ \textbf{in}. \\ \textbf{away}. \\ \textbf{back}. \\ \textbf{out}. \end{array} \right\}$$
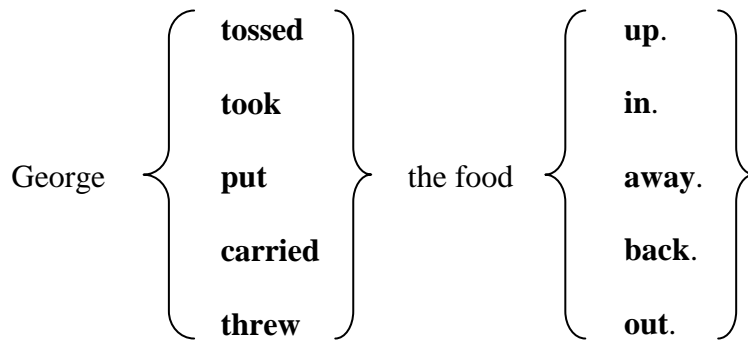
The example reflects the flexibility, or productivity, of this construction as "the paradigmatic insertion of any verb and any particle seems possible, as long as the combination of verb and particle allows an interpretation of motion through space, with the particle expressing the direction and the verb expressing the kind of the verbal action" (Thim, 2012, p. 14-15).

Despite the rather productive nature and apparent potential for free combinations, the corpus data examined in Sinclair's (1989) study shows that certain combinations occur more frequently than others. An example given is *fight back*. The verb *fight* retains the same lexical meaning as when it is used on its own, and the particle *back* carries a similar meaning in other combinations such as *strike back* and *phone back*. However, *fight* and *back* have a higher frequency of co-occurrence in the corpus data than *fight* with other particles. This suggests that this kind of construction is not exactly 'free' after all (cf. 'free combinations' in Quirk et al., 1985, p. 1152). This may argue against excluding compositional constructions as it is assumed that the contexts and nature of the data may influence the high frequencies of certain compositional constructions.

The examples of compositional constructions given above are all Type I phrasal

verbs, i.e. those with an adverbial particle, possibly because, as mentioned earlier, all of the studies categorising phrasal verbs based on their semantic characteristics focus only on those with an adverbial particle.

2.6.2 Aspectual constructions

The second semantic category comprises aspectual constructions which are combinations of a verb and an aspectual particle. Generally, aspect "relates to considerations such as the completion or lack of completion of events or states described by a verb" (Biber et al., 1999, p. 460). As for particles, they can also serve an aspectual function:

> They may add the concept of a goal or an endpoint to durative situations
> which otherwise have no necessary terminus. That is, the particles may affect
> the intrinsic temporal nature of a situation and hence alter its aktionsart from
> 'atelic' to 'telic'. (Brinton, 1985, p. 160)

Yet, the meanings of aspectual particles are not restricted to only atelic (i.e. incomplete) or telic (i.e. expressing an end-point). They can relate to marking states of beginning, completion, continuation and even high intensity (Downing & Locke, 2006). Based on a systematic classification of these aspectual meanings of particles (Celce-Murcia & Larsen-Freeman, 1999), four classes of aspectual particles are identified, namely inceptive, continuative, iterative and completive.

Completive aspectual particles

Completive aspectual particles mark a telic aspect in the sense that the verbal events are "directed towards a final stage" (Thim, 2012, p. 17). Typical particles of this sub-class include *up, out, off, down, over* and *through* (Celce-Murcia & Larsen-Freeman, 1999; Thim, 2012). Examples are (Thim, 2012, p. 18):

(41) With that beat I needed a really stomping guitar line to go with it so I **worked** it **through** in my head and then worked out the chords on the piano.

(42) Bill Murray spent £50,000 on setting up his restaurant at Telegraph Hill, near Exeter, Devon, two years ago but said the business started to go downhill when he **handed** it **over** to a manager to run.

The combination of a verb and an aspectual particle expresses a meaning of *V + NP completely*.

Continuative aspectual particles

Continuative aspectual particles denote an atelic aspect to show that the verbal event continues (Jackendoff, 2002; Thim, 2012). Typical examples are *away*, *on*, *along, through* as in the following examples (Celce-Murcia & Larsen-Freeman 1999, p. 432; Jackendoff, 2002, p. 78; Thim, 2012, p. 18):

(43) Miriam **worked away** on her manuscript.

(44) Abraham **talked on**, not noticing her lack of attention.

(45) In the end, Mungo reasoned that the old man had probably been **driving along**, had

somehow caught a glimpse of him, and had taken a short cut from the road.

(46) She **read through** her lines in the play for the audition.

The construction of a verb and these particles roughly expresses the meaning of 'continue V-ing' (Jackendoff, 2002).

Iterative aspectual particles

This sub-class of aspectual particles denotes repetition (Celce-Murcia & Larsen-Freeman, 1999). Consider the following examples (Celce-Murcia & Larsen-Freeman, 1999, p. 433; Jackendoff, 2002, pp. 78-79):

(47) They **play** the aria **over**.

(48) They **cook** the food **over**.

(49) He **did** it **over** and **over** again until he got it right.

Phrasal verbs with *over* are used to show the meaning of 'V + NP again' (Jackendoff, 2002).

Inceptive aspectual particles

Inceptive aspectual particles have the meaning of marking a state of beginning (Celce-Murcia & Larsen-Freeman, 1999). *Off* and *out* may be used as in the following examples (Downing & Locke, 2006, p. 341):

(50) He sat in an armchair in front of the television and soon **dozed off**.

(51) Yasmin says that she didn't **start out** to be a model. (BNC)

The use of these particles in the phrasal verbs seems to emphasize the beginning of an event (Downing & Locke, 2006).

In short, the meanings of aspectual constructions are reasonably predictable as a result of the consistent meanings expressed by the aspectual particles (Celce-Murcia & Larsen-Freeman, 1999; Jackendoff, 2002; Thim, 2012). Yet, they are not as transparent as compositional constructions with directional particles, so they are sometimes known as 'semi-compositional' or 'semi-idiomatic' constructions. Also, the examples above show that the range of verbs that can combine with the aspectual particles to form readily understandable constructions is rather wide (Jackendoff, 2002), though not as wide as that for directional particles. This takes us further along the semantic cline of phrasal verbs towards the non-compositional end at which idiomatic constructions are found.

2.6.3 Idiomatic constructions

The third semantic type of phrasal verbs is idiomatic constructions. As the term 'idiomatic' suggests, the meaning of the combination cannot be deduced by the respective meanings of the individual verb and particle (Sinclair et al., 1989; Hampe, 2002; Thim, 2012), as in the following examples (Thim, 2012, pp. 19-20; Biber et al., 1999, p. 414; Carter & McCarthy, 2006, p. 435):

(52) My husband actually said to me that **giving up** smoking was easy because he's done

　　it plenty of times.

(53) After she **hangs up** on Mark, Martha takes a deep breath and dials a London number.

(54) I won't **stand for** it.

(55) You said you'd already **come across** that sort of stuff in the library.

The items in (52) and (53) are examples of Type I phrasal verbs (with an adverbial particle), (54) and (55) are Type II phrasal verbs (with a prepositional particle).

Although there are no examples of Type III phrasal verbs provided in the studies which describe this semantic category of phrasal verbs, the existence of a single-word synonym of many Type III phrasal verbs is "a common sign of idiomatic status" (Quirk et al., 1985, p. 1160). Using different terminology, Huddleston and Pullum (2002, pp. 286-287) put this type of phrasal verb under the larger category of 'constructions containing verb + intransitive preposition idioms' and describe them as "verbal idioms". The examples of Type III phrasal verbs given in the grammar reference books all seem to be idiomatic, for example:

face up to	*fob off with*	(Quirk et al., 1985, p. 1160)
come up with	*put up with*	(Biber et al., 1999, p. 423)
run up against	*take out on*	(Huddleston & Pullum, 2002, p. 287)
watch out for	*put down to*	(Carter & McCarthy, 2006, p. 436)

Thus, it is reasonable to conclude that most Type III phrasal verbs constitute idiomatic constructions, and therefore idiomatic constructions consist of all three types of phrasal

verbs.

Characteristically, idiomatic phrasal verbs are very restricted in their syntactic behaviour (Hampe, 2002; Thim, 2012). They tend to appear in continuous order, i.e. the particle follows the verb immediately, in the majority of the cases.

2.6.4 Habitual constructions

This category is identified as 'habitual constructions' following Sinclair's (1989) description:

> … combinations where the verb is always used with a particular preposition or adverb, and is not normally found without it. Examples are *refer to* and *rely on*. We also include similar cases where a verb is always used with a particular preposition or adverb in a particular meaning (Sinclair et al., 1989, p. v).

It seems that most Type II, Type III and idiomatic Type I phrasal verbs fall into this category. While compositional and aspectual constructions contain mainly Type I phrasal verbs, Type II and Type III phrasal verbs carry meanings that cannot be inferred by the individual meanings of the verbs or the particles, which are neither directional nor aspectual. The fact that the verb used in a Type II phrasal verb is always used with particular particles makes Type II phrasal verbs fall squarely within the class 'habitual constructions'. Indeed, this seems to be the case with the two examples provided by Sinclair et al. (1989, p. v), *refer to* and *rely on*, both of which are Type II phrasal verbs. In addition to Type II phrasal verbs, Type III and idiomatic Type I phrasal verb also have

this characteristic of co-selecting a particular verb with particular particles.

Such a description seems rather ambiguous as the above illustrations show that many phrasal verbs, irrespective of their semantic meanings, are habitual combinations. In other words, the selection of a verb and a particle to be combined as a phrasal verb is not random. Therefore this habitual sense of "the verb is always used with a particular preposition or adverb" (Sinclair et al., 1989) does not seem enough to distinguish them as a separate category of phrasal verbs. Moreover, such a description seems morphological rather than semantic. It might be doubted whether this 'habitual construction' stands alone as a separate semantic category at all. Alternatively, it may also be possible to treat these so called 'habitual constructions' as part of the idiomatic constructions as their meanings are restricted to the fixed combinations of a specific verb and a specific particle.

Sharp dividing lines are difficult to draw between the semantic categories here (Downing & Locke, 2006; Thim, 2012). However, analysis provides insights and since two of the semantic categories are regarded as compositional, or semi-compositional, the interpretation of phrasal verbs may not be as complex as is traditionally presented. Although idiomaticity seems to be closely associated with the definition of phrasal verbs in much of the literature, it might be interesting and useful to examine the proportion of the semantic categories of phrasal verbs in a large dataset. For example, Pelli (1976) examined 14,021 phrasal verbs and regards only six percent of them as idiomatic, with all others as compositional and able to be interpreted semantically.

2.7 Other studies on phrasal verbs

In addition to the considerable body of literature dealing with the syntactic structure and semantic characteristics of phrasal verbs, there have been some attempts to study phrasal verbs from other perspectives and in different contexts. These previous studies tend to concern phrasal verbs in relation to second language acquisition and applied linguistics, such as learner knowledge of phrasal verbs and the challenge phrasal verbs pose to English language learners (Liao & Fukuya, 2004; Schmitt & Redwood, 2011; Barekat & Baniasady, 2014; Garnier & Schmitt, 2015; Chen, 2017), the need to compile a list of essential or most frequent phrasal verbs for teaching and learning purposes (Gardner & Davies, 2007; Ishii, 2009; Liu, 2011), the use of phrasal verbs in different varieties of English (Schneider, 2004; Mendis, 2010; Zipp & Bernaisch, 2012; Nelson & Ren, 2012; Gilquin, 2015; Chen, 2013, 2016). Rather few studies examine the use of phrasal verbs in specialised contexts (Campoy Cubillo, 2002; Trebits, 2009; Milizia, 2012, 2013; Oprit-Maftei, 2014; Rosca & Baicchi, 2016).

Some studies on English phrasal verbs focus on the issues of the acquisition of phrasal verbs by second or foreign language learners of English. Many studies recognise that phrasal verbs pose difficulties for second language learners and such difficulties potentially result in learners' consistent avoidance of using them (Liao & Fukuya, 2004; Gardnes & Davies, 2007; Siyanova & Schmitt, 2007; Gilquin, 2009; Schmitt & Redwood, 2011). One of the major potential reasons is the structural differences between English and the first language of learners, particularly to learners whose first language is not a Germanic language since phrasal verbs are regarded as a distinctive structure of Germanic languages (Dagut & Laufer, 1985; Celce-Murcia & Larsen-Freeman, 1999; Darwin & Gray, 1999). Although Chinese also has verb + particle structures, these have different

syntactic and morphological characteristics from those in English (Liao & Fukuya, 2004).

Another major reason may be the general perception of the idiomaticity of phrasal verbs.

Phrasal verbs are always claimed to have idiomatic meanings and it might be overstated

that the meaning of a phrasal verb cannot be deduced from the individual meanings of the

constituent elements. Some learners might even perceive that phrasal verbs are "largely

unpredictable" (Xiao, 2009, p. 445). This is possibly true in the cases of phrasal verbs

categorised as idiomatic. Nonetheless, the systematicity of the underlying semantic

patterns of the majority of phrasal verbs have been increasingly recognised (Side, 1990;

Morgan, 1997; Sansome, 2000; Villavicencio & Copestake, 2002).

These underlying patterns may not be easy for learners to perceive. Sansome

(2000) conducted a test to six students of English as a second language to examine the

extent to which they could comprehend the meanings of sentences with phrasal verbs

which have varying degree of idiomaticity, but which follow the same underlying pattern.

There were occasions where some students were able to learn from doing the test; for

example, two students were able to understand a certain sentence after another one with

the same particle was explained. In general, however, they tended to lack awareness of

underlying patterns. Thus, in spite of the recognised underlying patterns of the meanings

of phrasal verbs, and particles in particular, students are largely unaware of the patterns

and find phrasal verbs confusing. This explains: "the natural tendency to avoid using what

they do not properly understand and to prefer the more familiar one-word verb" (Dagut &

Laufer, 1985, p. 78).

Another reason non-English native learners find phrasal verbs difficult to master is

the overwhelming use of them in English (Liu, 2011). Their appearance is not limited to

informal contexts, and they are found across various types of texts (Cornell, 1985).

Gardner and Davies (2007, p. 347) estimates that "a learner will encounter 1 PV, on average, in every 192 words of English". The actual encounter of phrasal verbs by learners may be even more overwhelming due to the fact that the type of phrasal verbs Gardner and Davies (2007) concerned with is only one sub-type of that of the current study, i.e. a verb plus an adverbial particle. Thus, it is not easy to decide which ones or how many to begin with during the learning stage. It is pointed out that many textbooks base the coverage of phrasal verbs on arbitrary selection or the intuition of the authors (Darwin & Gray, 1999; Sansome, 2000; Schmitt & Redwood, 2011). Such selection of phrasal verbs for teaching and learning is of little help in mastering their use in real world contexts. There may also be discrepancies between the explanations given in textbooks and the real meanings of the verbs as encountered in natural language use (Zarifi & Mukundan, 2014).

Some attempts have been made to identify and classify phrasal verbs based on corpus findings. Larger-scale work includes the compilation of dictionaries of phrasal verbs, for example, the *Longman Dictionary of Phrasal Verbs* (Courtney, 1983), the *Collins Cobuild Dictionary of Phrasal Verbs* (Sinclair et al., 1989, 2002), *Cambridge International Dictionary of Phrasal Verbs* (Walter & Pye, 1997), *Macmillan Phrasal Verbs Plus* (Rundell & Fox, 2005), and *English Phrasal Verbs in Use: Advanced* (McCarthy & O'Dell, 2007). These works list and provide definitions of a large number of entries of phrasal verbs based on corpus data. For example, the *Collins Cobuild Dictionary of Phrasal Verbs* lists over three thousand combinations of verbs and particles based on a corpus study of the Birmingham Collection of English Texts (Sinclair et al., 1989). Major grammar reference books also discuss phrasal verbs in separate sections, for example, *A Comprehensive Grammar of the English Language* (Quirk et al., 1985),

Collins Cobuild English Grammar (Sinclair et al., 1990), *Longman Grammar of Spoken and Written English* (Biber et al., 1999), *An Introduction to Functional Grammar* (Halliday, 2004), *Grammar for English Language Teachers* (Parrott, 2010), and *Oxford Modern English Grammar* (Aarts, 2011).

These extensive works are based on corpus-based research findings and their pedagogical value is recognised and appreciated due to the wide coverage of phrasal verb items or the in-depth description and explanation of the use and features. However, researchers and learners can be overwhelmed by the considerably varied coverage resulting from the different criteria and methods used to select phrasal verbs. Little information regarding the frequency of use of particular verbs is provided in these works. Biber et al. (1999), who separate the discussion of the three types of phrasal verbs, provide a cross-register analysis and some information on the frequency of particular phrasal verbs, but the discussion is limited to a small group of items (Liu, 2011). Therefore, other studies value the need to identify essential or most frequent phrasal verbs which could serve as a more specific starting point for learners and researchers. Gardner and Davies (2007) identify 100 most frequent phrasal verbs from the British National Corpus and found that these phrasal verbs account for more than half (51.4%) of all the phrasal verb occurrences in the corpus. While the list provides important insights about phrasal verbs, it contains only the most frequent phrasal verbs formed by the top 20 lexical verbs. In other words, highly frequent phrasal verbs which are not formed by those 20 verbs may risk missing. Liu (2011) aims to provide a comprehensive list of the most common phrasal verbs in American and British English by validating the phrasal verbs identified in the work of Biber et al. (1999) and Gardner and Davies (2007), and conduct a comparative examination in the Corpus of Contemporary American English (COCA)

and the British National Corpus (BNC). While the results of these studies have research and pedagogical implications, the definition of phrasal verbs used in these two studies includes only the combinations of a verb and an adverbial particle. Adopting a broad definition of phrasal verbs which resembles the one used in the current study, Ishii (2009) considers both the importance of frequency and the coverage in phrasal verb dictionaries and identifies a list of basic phrasal verbs.

In addition to coverage and frequency, perhaps another major shortcoming of many phrasal verb lists and dictionaries is the lack of useful information on the important meaning senses of each phrasal verb, particularly when most frequent phrasal verbs have multiple meanings. Garnier and Schmitt (2015) extend the work based on Liu (2011)'s 150 most frequent phrasal verbs and develop the Phrasal Verb Pedagogical List (the PHaVE List) by including more semantic information of these most frequent phrasal verbs. Instead of simply listing the phrasal verbs based on their frequencies, Garnier and Schmitt (2015) identified the most frequent meaning senses of each of the top 150 phrasal verbs based on their occurrences in the COCA. Unlike dictionaries which aim to provide all possible meaning senses of each entry, Garnier and Schmitt (2015) consider that to be too overwhelming for language teachers and learners. They found that an average of nearly two meaning senses usually cover 75% of the occurrences of the most frequent phrasal verbs, thus they provide the definition, percentage of occurrence, and an example sentence for each meaning sense. This pedagogical list of phrasal verbs provides useful insights by considering the most common meaning senses and providing more semantic information, however, information such as collocation, semantic preferences and connotations is only given in some phrasal verbs.

Other studies examining the use of phrasal verbs put the emphasis on cross-variety comparisons, for example, new varieties of English (Schneider, 2004), native and second language varieties of English (Zipp & Bernaisch, 2012), Sri Lankan English and British English (Mendis, 2010), African Englishes (Nelson & Ren, 2012), new Englishes (Schneider & Zipp, 2013), and dialects within the UK and the US (Haddican & Johnson, 2012). Variations across varieties and variety-specific features were reported. For instance, in Singaporean English, phrasal verbs "are used frequently, widely and creatively" as compared to Englishes in India, the Philippines, East Africa, and even when compared to British English (Schneider, 2004, p. 246). In Xiao's (2009) multidimensional analysis of variation across five varieties of English, Hong Kong English was reported to use significantly fewer phrasal verbs, particularly in private conversations. Apart from the variation in the frequency of use, features of use were also reported. For example, Haddican and Johnson (2012) investigated the regional and national effects on the placement of the particle in phrasal verbs, i.e. continuous order vs. discontinuous order. While there is no obvious variation between the North and South dialects in the UK, a trans-Atlantic difference is observed in that the discontinuous order is preferred in the UK whereas US speakers tend toward the continuous order.

Many of these cross-variety studies use the nation-based sub-corpora of the International Corpus of English (ICE) as data for analysis. This corpus provides very useful resources for the comparison of the general language use between regions or nations, not only because the corpus includes text types from a range of discourse, but the design of corpus is also made consistent across all its component sub-corpora for compatibility and balance.

These studies tend to use general English as data because they aim to explore the general patterns of use of phrasal verbs, rather than register-specific features. There are, however, studies focusing on the use of phrasal verbs in more specialised contexts.

2.7.1 Phrasal verbs in specialised contexts

Much of the research on phrasal verbs tends to be based on general English and rather less attention has been given to studying the use of English phrasal verbs in more specialised contexts (Campoy Cubillo, 2002; Trebits, 2009). Biber et al (1999, p. 24) consider that the research on phrasal verb use in general English may be "incomplete and…even misleading or inaccurate" since many studies have failed to consider the issue of register. There can be substantial differences in use between and across genres or registers which may risk being obscured if the data is analysed without considering registers separately (Biber et al., 1999). In the case of phrasal verbs, since it is established empirically that they are more frequent in spoken or informal contexts (Biber et al., 1999; Dempsey, McCarthy & McNamara, 2007); it may be argued that there is little purpose in studying them in other types of texts. Nevertheless, the fact that they are more frequent in spoken or informal contexts by no means restricts them to that level. Indeed, they are found across various kinds of texts irrespective of formality (Cornell, 1985; Fletcher, 2005). It is reasonably expected that the phrasal verbs commonly used in a particular register probably differ from those in other registers or general language. The polysemous nature of phrasal verbs ensures that the same verbs, when used in different contexts, express different meanings (Campoy Cubillo, 2002; Trebits, 2009; Khir, 2012; Milizia, 2013). More importantly, phrasal verbs can be specific to specialised or formal registers:

It is not really true anymore to say that a phrasal verb always has a formal equivalent. The form you use or choose often depends on the context. There are now many phrasal verbs such as *check in*, *plug in* or *log on* that have come into English over the last years from science, technology and computing and they are known to have no alternative forms expressed in simple verbs. So when you use any of these phrasal verbs above you are not using a slang word that should be replaced by a formal verb since the phrasal verbs is the only way of describing these actions (Khir, 2012, p. 99)

Trebits (2009, p. 477) points out that studying phrasal verbs in different specialised contexts is helpful not only in understanding the range of meanings and use they have, but also in discovering that "the meaning or meanings of a word in a given register can be very important to characterize the register itself".

Large and well-designed reference corpora of general English provide invaluable data for cross-register investigation of the use of different linguistic features. Similar amounts of texts were collected from across a range of registers for comparison. For example, the Longman Spoken and Written English Corpus (the LSWE Corpus) consists of four main registers, namely conversation, fiction, news and academic prose. Biber et al. s' (1999) analysed the three types of phrasal verbs separately across the four registers and found that all three types are more frequent in conversation and fiction but least frequent in academic prose. Type I and Type III are particularly rare in academic prose. The more common use of Type II in academic prose indicates that the tone of Type II phrasal verbs is not considered as informal as compared to Type I. Liu's (2011) cross-register investigation of phrasal verbs in the Corpus of Contemporary American English supports Biber et al. s' findings that phrasal verbs are used more commonly in fiction and spoken English.

Studies have been carried out into the roles played by phrasal verbs in specialised contexts: botany research articles (Campoy Cubillo, 2002), European Union documents (Trebits, 2009), marine engineering English (Qu, 2010), business and finance English (Breeze, 2012), cartoons and puns (Khir, 2012), political discourse (Milizia, 2012; 2013), and police investigation (Rosca & Baicchi, 2016). Khir (2012) carried out a qualitative analysis of phrasal verbs used in cartoons to show how the use of the phrasal verbs in the texts and visual elements interact to convey meanings and create humour at the same time. The multi-meaning potential of phrasal verbs is exploited in cartoons which deliberately make use of both the literal and idiomatic meanings of the same phrasal verb.

Another, more qualitative, approach identifies those phrasal verbs with the most productive particles and analyses them using a cognitive approach (Requejo & Díaz, 2008; Breeze, 2012; Mahpeykar & Tyler, 2015). It had been viewed that particles do not contribute to the meanings of phrasal verbs (Fraser, 1976). Challenging the traditional view, Lindner (1981) provides a detailed analysis of two particles *up* and *out* in phrasal verbs using a cognitive linguistics approach and suggests an organised semantic network from which different meaning senses of the particles are extended. Phrasal verbs with the particle *out* are found to be related to the container metaphor, whether it is in general English as in *stuck out* and *thought out* (Rudza-Ostyn, 2003), medical English as in *take out* and *breathe out* (Requejo & Díaz, 2008), or business and finance English as in *break out* and *squeeze out* (Breeze, 2012). *Out* is a complex particle in terms of semantics. Among its various senses, one is perfective, i.e. the signalling of the completion of an action (Breeze, 2012). Campoy Cubillo (2002) also finds similar meanings of the particle *out* in the investigation of phrasal verbs with the most frequent particles in a corpus of research articles on plant biology. The findings show that *out* is commonly used to

express "a (long) process leading to a final state", as in *dry out* and *fade out* (Campoy Cubillo, 2002, p. 107). Through analysing the conceptual metaphors associated with the particles and their semantic contribution to the phrasal verbs, these studies shed light on the underlying meanings of phrasal verbs. The meanings of phrasal verbs may not be as arbitrary as traditionally described and categorisation based on particles may foster learners' understanding (Sinclair et al., 1989; Breeze, 2012; Milizia, 2013).

Mahpeykar and Tyler (2015) argue that only partial account is provided if explore further by considering also the polysemy of the verb and investigating the interaction of meanings of both the verb and the particle. The four phrasal verbs analysed are formed by two most frequent particles, *up* and *out*, and two most frequent and polysemous verbs, *get* and *take*. Studying 200 instances of each of the phrasal verbs from COCA, Mahpeykar and Tyler (2015) found that phrasal verbs such as the four analysed in their studies are in many cases compositional and the meanings are often the result of the combination of one of the particular meaning senses of the verb and one of the senses of the particle. For example, 'having a date' is one of the senses of *take out*. This meaning is extended from the central sense of *take*, i.e. 'Get Hold of and Remove Sense' and the 'Not In Situ Sense' of the particle *out*. The conceptualisation is one person initiating to bring another person from a base location to a second location. The meaning is further contextualised and developed as having a date, i.e. taking a person away from home, due to its frequent use in daily life (Mahpeykar & Tyler, 2015, p.31). They argue that this approach of placing consideration to both the verb and the particle provide a fuller account for a wide range of phrasal verb meanings.

In other studies, a corpus-based approach is used to identify phrasal verbs and analyse them both quantitatively and qualitatively. This approach normally begins with the generation of a list of the most frequent phrasal verbs in a context-specific corpus. Frequencies may then be compared with those found in general English corpora. There are clearly implications here for the design of English courses and their associated instructional materials in that phrasal verbs that are frequent in, or even specific to, a register or genre may not be commonly encountered in general English (Trebits, 2009). Very often, the frequency list is a good starting point. Phrasal verbs may then be selected for further analysis using concordances. Milizia (2012, 2013) examined the phrasal verbs commonly used in British and American political speeches and selected two of the most frequent phrasal verbs in the corpus for concordance analysis, namely DEAL WITH and FIGURE OUT. It was found that a negative or difficult semantic prosody is always associated with DEAL WITH in political speeches as words describing different kinds of crisis or problems are used in its vicinity. As for FIGURE OUT, in addition to discussing the slight negative semantic and a semantic association of 'difficulty', Milizia (2012, 2013) shows that the phrasal verb is always followed by *how to* and *ways to* which together tend to form longer phraseologies.

Findings in these studies also suggest the specific usage of phrasal verbs in specialised contexts. Typically, *carry out* has scientists or researchers as its subjects in journal articles, but in plant biology insects may also be subjects (Campoy Cubillo, 2002, p. 107). Trebits (2009) compared the number of word-senses of the most frequent 25 phrasal verbs in English documents of the European Union and found a more restricted range of meanings than in general English. For example, *set up* and *take up* have fifteen and thirteen word-senses respectively in general English, but they are found to have only

two possible meanings in EU English. While this analysis is useful in reflecting the specific meanings of phrasal verbs in a particular context and thus raising the awareness of language users, no detailed discussion is provided to show how the phrasal verbs are used differently in English documents of the European Union and in general English.

In Milizia's (2012) study of phrasal verbs in political discourse, details are provided to show how their use differs from that in general English. For example, *give up* is used rather differently in the spoken language of American and British politicians in that "the words in the vicinity of *give up* in spoken politics are the following: *nuclear, weapons, holiday, work, violence, time and energy, alliances, dependence, boarder protection, guns, opt-out, the pound, the rebate, the right to call a general election, nationhood, opportunity*" (ibid., p. 132).

Rosca and Baicchi (2016) focus on phrasal verbs with the particle *up* in the spoken English used in crime and police investigation and propose some possible teaching activities for English for Police. Five meaning senses are identified for the particle *up* in the data, and some context-specific meanings and functions are discussed. For example, the most frequent phrasal verb with *up*, i.e. *pick up*, is the most polysemous in the corpus, with five different meanings, namely taking or arresting someone to a police station for questioning, collecting something in an illegal transaction, cleaning and tidying a place or scene, capturing a scene on film, detecting a sound or smell. It is also found that some phrasal verbs function as copular verbs leading various adjectives, for example, *come up*, as in "But we've run the prints and we still came up empty", and *end up*, as in "How do a blackmailer and his victim both end up dead?". While the findings are interesting and show context-specific meanings and usage, and useful pedagogical activities are drawn,

the authenticity of the data may be questioned because the texts are extracted from the script of an American television series instead of naturally occurring language.

Recognising the influence of contextual factors on meanings of phrasal verbs, these studies demonstrate the value of studying the phrasal verbs that are specific to the particular genres or registers.

Another issue about phrasal verbs in previous work is the presentation of them. Irrespective of the aims of earlier studies, all the inflectional forms of the same phrasal verb are lemmatised (see, for example, Gardner & Davies, 2007; Trebits, 2009; Liu, 2011; Alejo-González, 2012; Zarifi & Mukundan, 2012, 2014; Milizia, 2012, 2013). This is also how they are presented in phrasal verb dictionaries and grammar references (see, for example, Courtney, 1983; Walter & Pye, 1997; Biber et al., 1999; Rundell & Fox, 2005; Spears, 2005; McIntosh, 2006). This kind of presentation inherently assumes that the frequency information and usage description provided about a phrasal verb is valid across all its inflectional forms. However, it is actually evident that there can be a substantial difference in the use of different inflections (Liu, 2011; Milizia, 2012). For example, the phrasal verb lemma FIRE UP ranks 20th in the spoken political corpus (Milizia, 2012). If no further illustration on its inflections is provided, it may be assumed that FIRE UP is used quite commonly in spoken political discourse, whether it is in the form of *fire up*, *fires up*, *fired up* or *firing up*. In fact, there are only 8 occurrences of the base form *fire up* in the corpus, but it is the passive use *fired up*, as in "Are you fired up? Ready to go?", which makes the lemma so frequent (ibid.). Thus while lists of phrasal verb lemmas provided in previous studies are useful in some ways, "the concept of word lemma is critical to taking full advantage of the lists" (Gardner & Davies, 2007, p. 353). It is

difficult for language users to interpret frequency information or meaning descriptions if no further details about various inflections are provided.

Some phrasal verb dictionaries mark certain phrasal verbs as important or more commonly used. In the *Cambridge International Dictionary of Phrasal Verbs*, phrasal verbs which are considered very common and useful for learners of English are highlighted in a grey box with the definitions (Walter & Pye, 1997). In *Oxford Phrasal Verbs Dictionary for Learners of English*, a key symbol is marked next to a phrasal verb if it is regarded as important (McIntosh, 2006). *Macmillan Phrasal Verbs Plus* marks more frequent phrasal verbs in red and express the degree of frequency with a star rating (Rundell & Fox, 2005). While providing some reference to dictionary users as to what phrasal verbs may be more essential in terms of their frequency of occurrence, such an indication of frequency or importance is based on lemma rather than specifically on the inflections. Some dictionaries give a short note of typical usage for the lemma entries. For example, the entry BASE ON is qualified in some dictionaries as "usually passive" (Sinclair et al., 1989, p. 11) or "often passive" (Rundell & Fox, 2005, p. 21). Some dictionaries provide a typical syntactic structure of 'base sth on sth' while also mentioning that it is often or usually used in the passive form (Walter & Pye, 1997, p. 8; McIntosh, 2006, p. 12). It can be seen that although the information in dictionaries can be useful in providing the general concept of the use of BASE ON, it is inadequate to determine how dominant some inflectional forms may be and how inflectional forms are used in a given context. It is worth investigating inflected phrasal verbs separately so as to better understand whether certain forms of a phrasal verb are more associated with particular registers (Liu, 2011). Therefore, to provide a more accurate account of the use

of phrasal verbs, this study treats each inflectional form as an individual phrasal verb initially.

Given the highly polysemous nature of phrasal verbs, this study has adopted Sinclair's model of five categories of co-selection as the framework for analysing the extended units of meaning of phrasal verbs within and across the genres used in the engineering sector. As semantic prosody "is both dynamic and text or genre specific" (Cheng, 2006, p. 341), it is argued that if phrasal verbs have different extended units of meaning across genres, they may be considered as specific to the genres in constructing a particular meaning or fulfilling a specific function.

Chapter 3   Methodology

This chapter depicts the nature and content of the data used, followed by an outline of the methods of extracting phrasal verbs from the data as well as the model used for data analysis.

3.1 Data

The data for the present study is a corpus of specialised English called the Hong Kong Engineering Corpus (HKEC) (Warren, 2010a). It consists of 9.2 million words of texts collected from the engineering sector in Hong Kong. This profession-specific corpus was compiled by researchers in the Research Centre for Professional Communication in English (RCPCE) of the Department of English of The Hong Kong Polytechnic University in 2007 and 2008. It is publicly available on the website of the RCPCE at http://rcpce.engl.polyu.edu.hk/HKEC/. The HKEC was compiled for use by researchers, those who learn or teach the language of the engineering profession, and the engineering professional themselves to study the authentic patterns of language use in this register. Expert advice from government departments, professional associations and companies was sought for the design of the corpus and data collection, particularly in terms of the types of texts collected and the proportions of text types to ensure that the corpus contents were representative of the text types that engineering professionals produce or encounter in English in their daily workplace (Warren, 2010b).

The texts in the corpus were mainly collected by downloading them from websites of companies and professional bodies. The sources of the texts were primarily from the Hong Kong Institution of Engineers (HKIE), and the websites of engineering-related

Hong Kong government departments, professional associations and private companies (Cheng, 2010). Despite the fact that many of the texts were publicly accessible on the internet, they are protected by copyright. The RCPCE research team obtained consent from the copyright holders of the texts as far as possible before they were included in the corpus. The corpus may, therefore, not be perfectly representative and balanced. Concerns of accessibility were raised by some consenting copyright holders, as a result of which it was agreed to limit access to the corpus data to search only with limited co-text available (Warren, 2010b).

The corpus is made up of 31 genres. Table 3 shows the word count and percentage of individual genres within the corpus.

Table 3: Contents of the HKEC

Genre	Word count	%	Genre	Word count	%
Abstract	94,671	1.020%	Position Documents	75,660	0.815%
Agreements	127,895	1.378%	Plans	4,173	0.045%
About Us	647,013	6.970%	Publicity Material	599,407	6.457%
Code of Practice	997,228	10.742%	Product Descriptions	611,549	6.588%
Circular Letters	143,313	1.544%	Project Summaries	115,829	1.248%
Conference Proceedings	196,498	2.117%	Q & A	27,703	0.298%
Consultation Papers	111,494	1.201%	Reports	979,170	10.548%
Frequently Asked Questions	55,726	0.600%	Review Papers	106,506	1.147%
Fact Sheets	26,059	0.281%	Standards	136,024	1.465%
Guides	783,805	8.443%	Speeches	2,822	0.030%
Handbooks	67,284	0.725%	Tender Notices	4,242	0.046%
Letters to Editor	3,492	0.038%	Technical Papers	65,731	0.708%
Manuals	296,299	3.192%	Transaction Discussions (HKIE)	7,149	0.077%
Media Releases	1,566,742	16.877%	Transaction Notes (HKIE)	79,058	0.852%
Notes	156,255	1.683%	Transaction Proceedings (HKIE)	1,055,248	11.367%
Ordinances	139,176	1.499%	Total	9,283,211	100.0%

While most of the genres are self-explanatory, some may be more specific to the engineering register and it is not always easy for non-members of this professional community to understand them. A more detailed description of these genres is found in Warren's (2010a) study. 'About Us' is often an introductory section found on companies' websites which describes the background of the companies, nature of their work and business, and the products and services the companies provide. 'Circular letters' are printed notices or announcements intended for mass distribution. 'Letters to editor' are primarily the responses from government departments to letters written to the Editor of South China Morning Post in response to engineering-related issues. 'Position documents' are documents presenting comments or views on a specific topic or issue on behalf of government bodies or organisations. 'Review papers' are reports submitted to the Legislative Council related to the review, consultation or background brief on issues, policies or proposals handled by government departments. 'Tender notices' describe tendering opportunities for engineering projects or contracts. 'Technical papers' consists of research and development papers published in journals or presented at conferences and workshops. 'Transaction discussions', 'Transaction notes' and 'Transaction proceedings' are three kinds of texts from the quarterly periodical *HKIE Transactions*. 'Transaction discussions' consists of very short pieces of discussion on a current topic of interest; 'Transaction notes' describe work in progress; and 'Transaction proceedings' are full length research papers describing research projects and/or development in the industry.

3.2 Methods of data extraction

As explained in Chapter 2.1, phrasal verbs examined in this study are combinations of a verb and an adverbial particle or a prepositional particle, or both. So they can be extracted automatically from the corpus by identifying the combinations of these parts-of-speech. The 31 genre-based sub-corpora were first tagged using Wmatrix (Rayson, 2009). The Part-of-speech tagging system adopted was the latest version of Constituent Likelihood Automatic Word-tagging System (CLAWS4) (Garside & Smith, 1997) which can achieve 96-97% accuracy (Van Rooy & Schäfer, 2003). The output of this step was a set of tagged sub-corpora in which each of the words was attached with a tag based on the part-of-speech of the word. Table 4 shows the tags of potential lexical verbs and particles, adverbial and prepositional, which may combine to form phrasal verbs.

Table 4: POS tags potentially forming phrasal verbs

Potential lexical verb	
Wordtag	**Part-of-speech**
VV0	base form of lexical verb
VVD	past tense of lexical verb
VVG	-ing participle of lexical verb
VVI	infinitive
VVN	past participle of lexical verb
VVNK	past participle catenative
VVZ	-s form of lexical verb
Potential particle	
IF	*for* (as preposition)
II	general preposition
IO	*of* (as preposition)
IW	*with*, *without* (as prepositions)
RL	locative adverb
RP	prep. adverb, particle

(adopted from CLAWS Tagset)

The tagged sub-corpora were then searched using WordSmith Tools (Scott, 2012)

for all possible constructions of phrasal verbs by performing search queries for a verb and an adverb or a preposition within a five-word span, i.e. the combination of any of the tags under 'Potential lexical verb' with any under 'Potential particle' in Table 4. This step generated possible constructions. While some of them are real phrasal verbs, others may be verb plus prepositional phrase, or simply chance occurrences where the verb and the adverb or prepositional phrase are not meaningfully associated. Therefore the generated results needed to be studied manually to determine which combinations are potential phrasal verbs and which ones are not. For this purpose, ConcGram 1.0 (Greaves, 2009) was used for its concordancing function. ConcGram can show all the possible configurations, i.e. contiguous instances, non-contiguous instances and positional variations, in a single search and displays them all in a single concordance. For example, a search for *comply* and *with* may result in concordance lines which include the following instances:

1. *…the company shall* ***comply with****…*
2. *…the consultants* ***comply*** *fully* ***with*** *the terms and conditions …*
3. *…a minimum standard* ***with*** *which buildings must* ***comply****…*

The first line above is a contiguous instance in which the two search words occur adjacent to each other. The second line is a non-contiguous instance where other words are found between the two search words. In this case, the intervening word 'fully' acts as an adverbial and it does not affect the association of *comply* and *with* as a phrasal verb. The third line is an example of positional variation where the two search words occur in different positions relative to one another. Although the particle of a phrasal verb typically follows the verb, there are some circumstances, as described earlier, in which the particle can be fronted. In this instance, the relative determiner *which* refers to 'a

minimum standard', which is the object of *comply with*. These forms of variation are all phrasal verbs. Therefore, using the concordancing facility of ConcGram allows all these possible configurations of a phrasal verb to be captured in a single search.

The concordances were examined manually to determine which instances are phrasal verbs, and thus obtain the actual frequencies of the phrasal verbs. For example, *used + in* was one of the results generated as two words co-occurring within a span of five words as in the instances below:

… the cables are *used in* fixed installations …

… have been *used* successfully *in* similar situations …

However, the combination of *used* and *in* is not regarded as a phrasal verb in this study. The preposition *in* is treated as forming a prepositional phrase with the noun phrase following it, i.e. 'in similar situations', rather than forming a constituent with *used*, and thus the combination is regarded as a verb plus a prepositional phrase rather than a phrasal verb.

In addition to checking whether the combinations are actual phrasal verbs or other grammatical structures, the results were also examined to see if there are any Type III phrasal verbs. Since any Type III phrasal verbs are combinations of a verb followed by an adverbial particle and a prepositional particle, a resulting two-word combination from the query search is potentially part of a Type III phrasal verb. For example, a resulting combination of *look* and *forward* may indeed lead to the finding of a Type III phrasal verb *look forward to*.

Combinations that were determined to be potential phrasal verbs were examined manually to determine whether all of the instances were actual phrasal verbs or chance

occurrences, by reading each instance in the co-text. For example, the combination of *set* and *up* is a phrasal verb as in instances given below:

> … we *set up* a pilot C&D materials recycling facility …
>
> … encourage enterprises to *set up* their own RE systems …

but it is not a phrasal verb in instances such as:

> … a framework was *set* out to follow *up* detailed issues …
>
> … adjusting the *set* point *up* to 29℃', and 'A data *set* is built *up* based on …

As one of the research objectives of the study is to investigate whether inflectional forms of a phrasal verb necessarily share the same patterns of use, each of the inflectional forms of the phrasal verbs was initially treated and presented as a distinctive phrasal verb item. Thus, this study considers the frequency and meanings of each inflected form separately and discusses the extent to which they may share the same or have different meanings as derived by their co-selections. Following Stubbs (2001), this study presents lemmas in upper-case, and inflectional forms, as individual phrasal verbs in this study, in lower-case italics. For example, the inflectional forms *carry out, carries out, carrying out* and *carried out,* which comprise the lemma CARRY OUT, are regarded as four distinct phrasal verbs in this study. If it is found that *carry out* has the highest frequency in a list or a sub-corpus, it refers only to that specific form, whereas *carries out, carrying out* and *carried out* have their own individual frequencies which may not be highly ranked.

A second researcher was invited to check the list of phrasal verbs identified after being given the definition of phrasal verbs used in the study. This process helped to support validity of the data and findings.

3.3 Data analysis

The phrasal verbs identified and their respective frequencies were tabulated to provide a list of the most frequent phrasal verbs for each of the 31 genre-based sub-corpora. The ten most frequent phrasal verbs in each list were compared across the lists to determine shared and genre-specific phrasal verbs. Since only the top ten phrasal verbs in each sub-corpus were examined, a phrasal verb found only in one genre does not necessarily mean that it is unique because it may also occur in other sub-corpora but is not in the top ten. The most frequent phrasal verbs were tracked to see which of the following three categories they fall into: (1) genres in which the phrasal verb occurs among the top ten most frequent ones, (2) genres in which the phrasal verb occurs but is not in the top ten, and (3) genres in which it does not occur at all.

The sizes of the sub-corpora vary, and so the percentages of the frequency of occurrence in the sub-corpora are calculated and compared. The reason for using percentage of frequency rather than normalizing the frequency to a widely used basis of per million words is essentially due to the wide range of sizes of the sub-corpora. Particularly for those very small sub-corpora, which consist of less than 10,000 words, normalizing the frequencies by multiplying them too many times may misrepresent their actual patterns of use.

The most commonly used phrasal verbs across the genres were selected for qualitative analysis. Sinclair's (2004) model of lexical description was adopted for analysing the phrasal verbs. All the concordance lines were manually and carefully analysed in terms of five categories of co-selection, namely the core, collocation, colligation, semantic preference and semantic prosody, to determine whether the phrasal

verbs are part of the same or different extended units of meanings within and across the genres, and if different, how many possible units of meaning a phrasal verb and its co-selections may have within and across the genres.

The analytical procedure follows the sequence of colligation, collocation, semantic preference and semantic prosody, considering their level of realisation, from more tangible linguistic realisation to relatively more abstract and subtle realisation. The colligation pattern was analysed by first examining the left side of the core, i.e. the phrasal verb investigated, to identify any re-current grammatical structures. The same process was then repeated on the right side of the core. The most frequent syntagmatic configurations were then shown. For collocation, the actual co-occurrence of individual words or phrases on both sides of the core was examined and recorded. The collocate function of WordSmith Tools was also used to show the number of collocates because some of the concordances contain more than 1,000 instances. Collocation is closely tied to semantic preference (Partington, 2004). Therefore, to determine the semantic preference, the surrounding words or phrases on the left or right of the core were examined to see if there is any similarity of meaning. In other words, the collocates identified were classified based on features they share to arrive at the semantic preference. The final but obligatory category, semantic prosody, is not restricted to any conventional realisation (Sinclair, 2004), but any form of consistent pattern in the proximity of the core, irrespective of whether the pattern is established through the collocation, colligation or the semantic preference. Careful examination and consideration on collocates, semantic preference and the syntagmatic patterns help to identify the overriding semantic prosody.

It is noted that the analytical procedure often involved and extended to a wider co-

text which was not shown within the span of the concordance. This is partly due to the nature of analysing co-selection, and partly due to the fact that engineering texts are formal and technical. Terminology and sentences are often long, technical and complicated. Therefore, expanding the span of the concordance, or even referring back to the original text were often needed to investigate the wider co-text and contexts.

A second researcher was invited to study and discuss the concordances of the selected phrasal verbs in the sub-corpora in which they are most frequent to provide quality and validity assurance.

Here, the phrasal verb *set out* is used as an example. Figure 1 shows a sample of concordance lines of *set out* extracted from the HKEC.

```
1      as a Director.  The relevant procedures are set out in the Notice of AGM which  accompanies the
2   of the resolutions relating to those items are set out below: ? Election of Messrs. Jason Whittle,
3    year, the Company met the Code Provisions as set out in the Stock Exchange Code  contained in
4   and Board Committee Meetings held in 2007 are set out in the following table. The overall
5   public companies in the past three  years are set out in the Notice of AGM. Other details of
6    Other details of Directors' appointments are set out in  "Directors and Senior Management"
7    no less exacting than the  required standard set out in the Model Code.  All Directors have
8   2007 they complied with the  required standard set out in the Model Code and our own Code for
9  paid to each Non-executive Director in 2007 are set out in the Remuneration Report at page 120 of
10         of the Chairman and CEO are more fully set out in the CLP  Code. The interests in CLP's
11   by Directors, largely based on the Model Code set out in Appendix 10 of the Listing  Rules. Our
12    and work done during the Relevant Period are set out in the Audit Committee Report at page 118 of
13     is  conducted. This Code of Conduct is also set out in full on our website. Management and staff
14         the 10 managers, whose biographies are set out on page 102 of  this Annual Report) and
15   nature. 4. Project proponents are required to set out targets deliverables in their funding
```

Figure 1: Sample concordance lines of *set out* in the HKEC

It is found that there is no strong collocate of *set out*. Thus *set out* is the core itself. Regarding colligation, i.e. the co-occurrence of grammatical classes, 10 lines are found to have a verb *to be* at N-1 (8 times), N-2 (1 time) and N-3 (1 time) position, mainly in the form of 'are' (9 times). This pattern suggests that *set out* is used in the passive form in these lines. In lines 3, 7, 8 and 11, *set out* is a passive participle modifying the noun phrase it follows, such as 'the required standard' in lines 7 and 8. Hence, only one line is used in the active voice, i.e. line 15 in which *set out* follows an infinitive 'to'. The

overwhelming use of the passive form indicates that the focus is on the subject of *set out*, such as 'the relevant procedures' in line 1, and 'other details of Directors' appointments' in line 6. And it is not important to mention the agent of the verbal action. On the right side, s*et out* is colligationally followed by a prepositional phrase (13 times, 86.7%). The paradigmatic choice of the preposition is 'in' (12 times).

A semantic preference of 'written record' is observed to be associated with *set out* in 14 lines (93.3%). Examples are 'the Notice of AGM', 'the Model Code', 'the Remuneration Report', 'Appendix 10 of the Listing Rules', 'our Website', 'page 102 of this Annual Report'. These phrases share the same semantic feature of being a document or a piece of writing regardless of its format. A semantic prosody of 'official' is observed because the phrases found surrounding *set out* are 'the relevant procedures', 'the Code Provisions', 'details of Directors' appointments', 'the required standard', 'the Model Code', 'Code of Conduct', and 'biographies'. These seem to be more significant or regulatory information. Also, most of the noun phrases to the right contributing to the semantic preference of 'written record' tend to be official in nature and made public by authorities or corporations and cannot be altered easily. Thus, the analysis shows that the extended unit of meaning with *set out* as the core is 'referencing significant information in official written record'.

The above example demonstrates the analysis of the contiguous form of a phrasal verb. In the actual analysis of the most frequent phrasal verbs (Chapter 4), the contiguous form, non-contiguous form(s) and other inflectional forms are analysed one by one separately. The following chapter presents detailed analysis and discussions of findings.

Chapter 4   Findings and discussion

This chapter is devoted to presenting and discussing the findings related to the frequency and textual use of the most frequent phrasal verbs in the engineering profession in Hong Kong. In section 4.1, an overview of the most frequent phrasal verbs across the 31 genres is given. In addition to most commonly used phrasal verbs, unique phrasal verbs are identified. Sections 4.2 to 4.6 focus on the analysis and discussion of the five most frequent phrasal verbs in the corpus. The analysis is based on Sinclair's (1996, 1998, 2004) model of five categories of co-selection which comprise the extended unit of meaning. The inflectional forms of these five most frequent phrasal verbs are also examined to reveal whether they share the same extended unit of meaning or not. The final section, 4.7, summarises and discusses the overall findings.

4.1 The most frequent phrasal verbs in engineering sub-corpora

This section serves to provide a general picture of the most frequent phrasal verbs across the sub-corpora. The lists of the top ten most frequent phrasal verbs from the 31 sub-corpora (Appendix I) were generated and compared against each other. As shown in Appendix I, some of the lists have more than ten phrasal verbs. This happened when some phrasal verbs have the same frequencies in a particular sub-corpus. In such cases, more than one phrasal verb shares the same ranking. For example, in the list derived from Fact Sheets, five phrasal verbs recorded the same frequency, and they all rank seventh in the list. On the other hand, some lists have fewer than ten phrasal verbs. Due to the small size of some sub-corpora, frequencies of some phrasal verbs can be very low. The cut-off used

for inclusion of items in the lists is two. For example, in the list of Tender Notices, only five phrasal verbs have frequencies higher than two.

The 31 lists were compared to investigate which phrasal verbs are shared and which ones are unique. This comparison was only made across the lists of the top ten most frequent phrasal verbs. Therefore, a phrasal verb found to be shared by, for instance, ten sub-corpora may actually occur in sub-corpora other than those ten sub-corpora with, however, a lower ranking. For the same reason, a unique phrasal verb, meaning that it only occurs in a single list of the top ten most frequent phrasal verbs, does not necessarily mean that it has zero occurrence in other sub-corpora. It may also occur in other sub-corpora but is outside of the top ten.

A total of 98 distinct phrasal verbs are identified in the 31 lists. The number of sub-corpora in which each of these phrasal verbs occurs ranges from one to 30. There are 47 shared phrasal verbs, of which, eight are shared in more than ten sub-corpora, namely *based on* (26 sub-corpora), *carried out* (24 sub-corpora), *comply with* (16 sub-corpora), *refer to* (12 sub-corpora), *set out* (12 sub-corpora), *associated with* (12 sub-corpora), *carry out* (11 sub-corpora) *and set up* (11 sub-corpora). For the remaining 39 shared phrasal verbs, the number of top ten lists in which they are found ranges from two to eight. The most frequent phrasal verbs will be discussed in detail with the concordance findings and patterns of co-selection in sections 4.2 to 4.6. The rankings of the top five phrasal verbs in each sub-corpus are shown in Table 5 below.

Table 5: Ranking of the most frequent five phrasal verbs in the sub-corpora

No.	Genre-based sub-corpora	Ranking of the phrasal verbs in the sub-corpora				
		based on	*carried out*	*comply with*	*refer to*	*set out*
1	Abstract	1	2			5
2	Agreements	1				
3	About Us	1				2
4	Code of Practice	3	1	2	7	
5	Circular Letters	2				1
6	Conference Proceedings	1	2	10		
7	Consultation Papers	1	6	4	3	2
8	Frequently Asked Questions	2	6	3	1	
9	Fact Sheets		3			
10	Guides	3	1	2	6	5
11	Handbooks	5	1	7	10	8
12	Letters to Editor				3	
13	Manuals	2	1	3	5	
14	Media Releases	5	4		1	7
15	Notes	2	1	6	8	8
16	Ordinances		1	3		
17	Position Documents	2	9		7	
18	Plans					2
19	Publicity Material	1	3	2		
20	Product Descriptions	1	2	4		
21	Project Summaries	10	5			
22	Q & A	1	4	3		
23	Reports	2	1	8	6	7
24	Review Papers	1	7	4		3
25	Standards	6	4	1		2
26	Speeches		1			
27	Tender Notices	1				
28	Technical Papers	2				
29	Transaction Discussions (HKIE)	1	2		5	
30	Transaction Notes (HKIE)	1	2	7		
31	Transaction Proceedings (HKIE)	1	2			

The first three phrasal verbs, particularly *based on* and *carried out*, are the most

frequent phrasal verbs in the corpus, not only in terms of their overall frequencies but also

their distribution across the sub-corpora. It is found that they are among the top three most

frequent items in most of the sub-corpora. *Based on,* which is the most commonly used phrasal verbs across the engineering genres, is found in 26 lists of the top ten phrasal verbs. The data in Table 5 reveals that it is among the top three in 22 sub-corpora, and is the most frequently used phrasal verb in 13 sub-corpora, including Agreements, Consultation Papers, Publicity Materials and Transaction Proceedings. It ranks second in seven sub-corpora, including Circular Letters, Manuals, Reports and Technical Papers, and the third in two sub-corpora, namely Codes of Practice and Guides. In total, there are 3,671 occurrences of *based on* in the engineering corpus. In addition to these 26 sub-corpora in which *based on* is among the top ten, the phrasal verb is only absent from the sub-corpus of Letters to Editor.

Generally, *based on* is used predominantly in the passive voice which matches the description in *Longman Grammar of Spoken and Written English*. Biber et al. (1999) present this phrasal verb as *be based on* rather than simply *based on* indicating its habitual use as a passive verb. It is very commonly used in expressing the relationship between two things, i.e. how one thing is decided, determined or formed with the other thing as reference or foundation. Engineering always requires rigorous measurements and judgements in terms of details and processes that need to be carried out accurately and safely. Thus, the phrasal verb *based on* is used very frequently in engineering texts to describe the foundation or basis used for the decisions.

The second most commonly used phrasal verb across the engineering genres is *carried out*. As shown in Table 5, this phrasal verb is found in 24 lists of top ten most frequent phrasal verbs. It ranks first in eight sub-corpora, including Codes of Practice, Guides, Handbooks and Ordinances, second in six sub-corpora, including Abstracts,

Product Descriptions, Transaction Discussions and Transaction Notes, and the third in two sub-corpora, namely Fact Sheets and Publicity Materials. Thus, *carried out* is among the top three most frequent phrasal verbs in 16 sub-corpora. In fact, it has an overall occurrences of 3,785, which is even higher than that of *based on*. Its frequent occurrence is due to the nature of engineering which involves performing a wide variety of processes, and thus *carried out* is frequently used to describe the processes and works done. It has been reported that *based on* and *carried out* (and other forms of CARRY OUT) are among the most frequently used multi-word constructions in academic writing in both American and British English (Liu, 2011). Engineering texts tend to resemble academic writing in the sense that they describe the technical aspects of how certain routines and tasks are performed. The only two sub-corpora in which there is no trace of *carried out* are Letters to Editor and Plans.

The third most commonly used phrasal verb is *comply with* which occurs in 16 lists of the top ten most frequent phrasal verbs and is among the top three in eight sub-corpora. It ranks first in Standards. In Codes of Practice, Guides and Publicity Materials, *comply with* ranks second most frequent and in Frequently Asked Questions, Manuals, Ordinances and Q & A, it ranks third. Overall, it is found in 26 sub-corpora and has a total frequency of 1,831. In addition to stating the particular judgements or evidence which shows how the decisions are made rigorously regarding the design and processes of the engineering tasks, it is also very common for engineering texts to ensure that designs, procedures, products, or human behaviour comply with stipulated requirements or standards.

The fact that these three phrasal verbs are highly frequent in engineering genres by

no means guarantees the same patterns for their inflectional counterparts. *Based on*, as the most frequently and commonly used phrasal verb in engineering sub-corpora, is the only inflectional form of BASE ON found across the lists of top ten phrasal verbs. Since the lists only show the top ten phrasal verbs in each sub-corpus, all the sub-corpora were searched to discover whether *base on*, *bases on* and *basing on* occur. It was found that they have very limited frequencies across the sub-corpora. For example, *base on* occurs 49 times in the whole corpus, *basing on* occurs 16 times and *bases on* only occurs five times.

The inflectional forms of CARRY OUT are generally more commonly used in engineering genres. In addition to *carried out* which is found in 24 lists of the top ten phrasal verbs, *carry out* is found in 11 lists and *carrying out* in six lists. Overall, *carry out* occurs 1,474 times in the whole corpus and *carrying out* has a total frequency of 879. The other form *carries out* is much less frequently used in engineering genres. It has a total frequency of 88 and is found across 17 sub-corpora.

As for COMPLY WITH, *comply with* is the only phrasal verb among other forms found across the lists of top ten phrasal verbs. *Complied with* and *complying with* are found to be unique in the lists of Tender Notices and Review Papers respectively. All sub-corpora were searched to investigate whether there is an occurrence of these "unique" phrasal verbs. In the case of *complied with* and *complying with*, they are both found in more than one sub-corpus. *Complied with* is used in 19 sub-corpora with a total frequency of 347, and *complying with* has a total frequency of 381 distributed across 23 sub-corpora.

There are two top ten lists in which these three most frequent phrasal verbs are not found: Letters to Editor and Plans. In addition, these two sub-corpora have the highest number of unique phrasal verbs. While most sub-corpora have one or two unique phrasal

verbs, Letters to Editor and Plans have five and four unique items respectively. This indicates that the use of phrasal verbs in these two sub-corpora is considerably different as compared to the other sub-corpora.

The top three phrasal verbs in Letters to Editor are *thank for*, *signed up* and *refer to*. This sub-corpus is a small collection of letters written by officials of the Environmental Protection Department in response to earlier published letters addressing environmental issues. The writers often allude to the particular previous letter they would like to discuss or follow up on and express appreciation for people's interest in environmental issues. Some letters also mention participants signing up for certain programmes launched by the Environmental Protection Department. These are the contexts in which the top three phrasal verbs are used in Letters to Editor. The unique phrasal verbs in the top ten list are *signed up, point out, drop off, extended to* and *look forward to*.

In Plans, the top three most frequent phrasal verbs are *provide for*, *set out* and *allow for*. In fact, this sub-corpus only contains two pieces of texts: proposals to amend current ordinances and to introduce new schemes respectively. The top three phrasal verbs are mainly used in the contexts of how the Environmental Protection Department made adequate consideration of and preparation for proposed changes, and what and where they would detail the regulations and ordinances or purpose of the legislation. The unique phrasal verbs among the top ten found in Plans are *allow for, consults on, give out* and *consulted on*. Letters to Editor and Plans are very small in size with less than five thousand words each. The rather different sets of the most frequent phrasal verbs found in these two sub-corpora can be attributed to both the nature and size of the two.

The top five most commonly used phrasal verbs and their inflected counterparts are analysed in greater detail for their co-selection in sections 4.2 to 4.6 below.

4.2 Use of phrasal verbs: BASE ON

This section is devoted to the analysis and discussion of the co-selection across the sub-corpora of the most frequent phrasal verb *based on*. The section then examines the co-selections of other inflectional forms: *base on*, *basing on* and *bases on*.

4.2.1 *Based on*

The phrasal verb *based on* is the most frequently found phrasal verb in the Hong Kong Engineering Corpus. It occurs 3,671 times (0.0398%) in the corpus and is found in all of the sub-corpora except Letters to Editor. Table 6 shows the frequencies and percentages of occurrence of *based on* in each sub-corpus. In the table, the sub-corpora are ranked according to the percentage of occurrence of *based on* in each of the sub-corpora.

Table 6: Frequencies and percentage of occurrence of *based on* in the sub-corpora

	Sub-corpora	Size of sub-corpora	Frequency of *based on*	%
1	Tender Notices	4,242	6	0.1414%
2	Transaction Discussions (HKIE)	7,149	8	0.1119%
3	Agreements	127,895	138	0.1079%
4	Abstract	94,671	84	0.0887%
5	Transaction Proceedings (HKIE)	1,055,248	894	0.0847%
6	Transaction Notes (HKIE)	79,058	63	0.0797%
7	Q & A	27,703	15	0.0541%
8	Publicity Material	599,407	313	0.0522%
9	Technical Papers	65,731	32	0.0487%

10	Consultation Papers	111,494	53	0.0475%
11	Conference Proceedings	196,498	92	0.0468%
12	Notes	156,255	71	0.0454%
13	Frequently Asked Questions	55,726	25	0.0449%
14	Circular Letters	143,313	64	0.0447%
15	Position Documents	75,660	33	0.0436%
16	Manuals	296,299	124	0.0418%
17	Review Papers	106,506	44	0.0413%
18	Handbooks	67,284	27	0.0401%
19	Speeches	2,822	1	0.0354%
20	Guides	783,805	272	0.0347%
21	Code of Practice	997,228	344	0.0345%
22	Reports	979,170	332	0.0339%
23	Standards	136,024	42	0.0309%
24	About Us	647,013	194	0.0300%
25	Product Descriptions	611,549	169	0.0276%
26	Plans	4,173	1	0.0240%
27	Project Summaries	115,829	17	0.0147%
28	Media Releases	1,566,742	207	0.0132%
29	Fact Sheets	26,059	2	0.0077%
30	Ordinances	139,176	4	0.0029%

Tender Notices is the sub-corpus with the highest percentage of occurrence of *based on*. As shown in Table 6, it is a very small corpus with less than 5,000 words, but there are six instances of *based on*. Their concordance lines are shown in Figure 2.

```
1  bid or, where the assessment of the tenders is based on a marking scheme or formula approach, the
2  bid or, where the assessment of the tenders is based on a marking scheme, the tender has the highest
3  bid or, where the assessment of the tenders is based on a marking scheme or formula approach, the
4  bid or, where the assessment of the tenders is based on a marking scheme or formula approach, the
5  bid or, where the assessment of the tenders is based on a marking scheme or formula approach, the
6  bid or, where the assessment of the tenders is based on a marking scheme or formula approach, the
```
Figure 2: Concordance lines of *based on* in Tender Notices

At first glance, the concordance lines appear to be identical. Tracking the actual locations of these six instances in the sub-corpus shows that all instances are found in the remarks section of different tender notices. An extract is given below as an example:

Remarks The Government of the Hong Kong Special Administrative Region does not bind itself to accept any tender irrespective of whether the tender is the lowest bid or, where the assessment of the tenders is based on a marking scheme or formula approach, the tender has the highest overall mark.

It is this norm of providing such remarks in the tender notices which makes the phrasal verb *based on* so frequently used in the sub-corpus of Tender Notices.

In terms of colligation, there is always an auxiliary verb immediately in front of *based on*. This indicates that all the six instances are used in the passive voice. Following the phrasal verb, there is always a noun phrase. So the typical configuration of this phrasal verb in Tender Notices is:

auxiliary verb + based on + noun phrase

The collocates to the left of *based on* are "the" (6 times), "assessment" (6 times), "tenders" (6 times), and "is" (6 times). To the right, there are "a" (6 times), "marking" (6 times), "scheme" (6 times), "formula" (5 times), "approach" (5 times), etc.

The semantic preference of *based on* in this sub-corpus is 'evaluation of tenders' (100%) which is due to the co-occurrence of the phrase "the assessment of the tenders". The noun phrase "a marking scheme" refers to an instrument used for the evaluation and assessment of tenders.

Regarding the semantic prosody, the co-occurrence of "a marking scheme or formula approach" indicates that there is a sense here of 'rigorous' procedures (100%). "A marking scheme" or "formula approach" denotes a process which is designed to be rigorous, objective and widely-accepted. The co-selection of the semantic preference and

semantic prosody shows that the basis of the assessment of the tenders is rigorously and objectively benchmarked. The use of *based on* in these instances conveys sense of rigour.

There are only six instances of *based on* in this sub-corpus and it might be argued that the analysis of the phrasal verb is not representative based on such small data. However, it is found that the co-selections of *based on* are nearly the same in most of the instances across all the sub-corpora. Irrespective of which sub-corpus *based on* is used in, a strong semantic prosody of 'rigorous' is found to be associated with the phrasal verb. In fact, in the top three sub-corpora in which *based on* is found, namely Tender Notices, Transaction Discussions (HKIE) and Agreements, all of the instances of *based on* have very similar co-selections and the same semantic prosody of 'rigorous'. The instances in these sub-corpora are analysed below.

Transaction Discussions is the second sub-corpus in Table 6. Similar to Tender Notices, it is a very small sub-corpus of less than 8,000 words, but the percentage of occurrence of *based on* is high. There are eight instances, as shown in Figure 3 below.

```
1       than B; X is much more responsive than Y. (Based on Au, 1990) Response to Discussion by S.W.C.
2   of the second point, I have proposed Figure 1, based on statistics from over 1000 failures I have
3 at the study of the lock-up-stresses in the deck based on the selected column rigidities causing no
4       property of concrete that is fundamentally based on compressive strength. In Hong Kong it has
5  of modulus to be made using the same values but based on the actual (mean) strength of the concrete
6      analysis is a statistical estimation method based on regionalised variables. Regionalised
7   rarely exceed 10% even if numerous large voids based on the classification of construct on standard
8   of voice  Table 1 Summary of Core Test Results based on the recommendations given in Concrete
```

Figure 3: Concordance lines of *based on* in Transaction Discussions

All these instances are used in the passive voice. The pattern of colligation on the left side of the phrasal verb is not as obvious as that in Tender Notices. In lines 3, 6 and 7, there is a noun phrase (37.5%). In lines 1, 2 and 8, no other words precede *based on* in the clause (37.5%), for example in line 1, the phrasal verb is used inside brackets to provide extra information, or in line 8, the phrasal verb is used in a point form expression to

94

illustrate information in a table. In the remaining lines, there is an auxiliary verb followed by an adverb, and a conjunction. To the right of *based on*, there is always a noun phrase. Regarding collocation, there is no strong collocational pattern on either side of *based on*.

A semantic preference of 'assessment and evaluation' (75%) is observed in the use of *based on* in the sub-corpus. This semantic preference is made evident by the co-occurrence of phrases such as 'the study' in line 3, 'a statistical estimation method' in line 6 and 'test results' in line 8. They suggest that *based on* is mainly used in the contexts of discussing certain assessments or estimations. The words 'table' and 'figure', which occur in lines 1, 2 and 8, are also associated with this semantic preference as it is the usual case that test results or investigation findings are summarized and presented using tables or figures.

As for the semantic prosody, all the instances of *based on* convey the sense of 'rigorous' (100%). In fact, this semantic prosody is suggested by the noun phrases immediately following the phrasal verb. The phrase 'Au, 1990' and the word 'statistics' in lines 1 and 2 respectively denote the practice or study using a collection of data for analysis to achieve certain research or statistical purposes. These kinds of practice are of a scientific nature and are often rigorous and objective. In lines 3 to 6, the phrases 'the selected column rigidities', 'compressive strength', 'the actual (mean) strength' and 'regionalised variables' denote some values which are measured and obtained through particular well established or widely accepted methods of measurements. As for lines 7 and 8, the noun phrases are 'the classification of construct on standard CSI', in which CSI stands for the Construction Specifications Institute, and 'the recommendations given in Concrete Society Technical Report No 11'. In these two cases, the classification and recommendations are not arbitrarily made, but are established by institutions which are

renowned in the profession and to which community members look to for standardisation and authority. Therefore, the use of *based on*, together with its co-selections, create a meaning of grounding certain assessments and evaluations in rigorous and objective evidence or methods. This meaning is also shared by the use of *based on* in Tender Notices, but there is a difference in the contexts in which the phrasal verb is used. It is used in the discussion of assessments related to tenders in Tender Notices, but tends to be related to construction in Transaction Discussions.

Agreements is the sub-corpus with the third highest percentage of occurrence of the phrasal verb. It is a larger sub-corpus and there are 138 instances of *based on*. Figure 4 shows a sample of the concordance lines of *based on* in Agreements.

```
1    voluntary assessment of building performance based on LCA and LCC are discussed.The book concludes
2   why these steps are needed, the discussion is based on a hypothetical case of constructing, using
3 most widely used for determining the weights is based on subjective judgments on the relative
4           normalization and weighting, will be based on the quantities of the chemicals determined.
5         released, and further LCIA steps will be based on the quantified damages. Conceptually, the
6      releases (Cl- and Br- that destroys ozone) Based on chemical's reactivity lifetime MIDPOINT
7 should not simply ignore LCA in building design based on this reason, because the required data
8         to buildings Life cycle costing (LCC) is based on the theory of interest in economics, and
9     price growth, however, should be determined based on data provided by the US Department of Energy
10 calculation method Life cycle costing (LCC) is based on the concept of discounting future worth
```

Figure 4: Sample concordance lines of *based on* in Agreements

All the instances are used in the passive voice. To the left of the phrasal verb, there are a verb phrase (61 times, 44.2%), an auxiliary verb at the L1 or L2 position (51 times, 37.0%), and a noun phrase (12 times, 8.7%). In some instances, *based on* is used as the beginning of the sentence (14 times, 10.1%). To the right, there is always a noun phrase (100%). Thus, the most frequent configurations of *based on* in Agreements are:

verb phrase + based on + noun phrase

auxiliary verb + based on + noun phrase

The verb phrase is mostly realized in the form of either *auxiliary verb + verb*, as in line 9, or *verb + noun phrase*, as in line 7.

Collocates to the left include "the" (66 times), "of" (37 times), "and" (36 times), "is" (32 times), "be" (28 times), "in" (27 times), "to" (26 times), "for" (18 times), "a" (18 times), "building" (15 times), "are" (15 times), "energy" (14 times), "LCA" (13 times), "impacts" (12 times), "determined" (12 times), etc. To the right, top collocates are "the" (135 times), "of" (63 times), "and" (47 times), "data" (24 times), "to" (17 times), "in" (16 times), "for" (16 times), "LCA" (14 times), "that" (11 times), "from" (11 times), "energy" (11 times), "a" (11 times), "which" (10 times), "is" (10 times), "use" (9 times), etc. The use of words such as 'determined', 'estimated', 'calculate', 'assessment', and 'survey', which are found in 92 instances, contributes to a semantic preference of 'assessment and determination' (66.7%). This indicates that most of the instances of *based on* are used in the contexts of assessing or determining something. For example, in line 1, "benchmarking and voluntary assessment of building performance" is mentioned, and in line 3 "determining the weights" is used. More examples from the other lines include "an annual energy cost saving evaluated", "the survey", "the impacts of a building normalized", "the calculation", "the annual equivalent full-load hours were determined", etc.

In 26 other instances, *based on* has a semantic preference of 'development of tools' (18.8%). This semantic preference is made evident by the co-occurrence of words such as 'models', 'established', 'approaches', 'methodology', 'methods', 'develop', 'tool' etc. Examples are "system models may still be established", "energy systems", "these

approaches are frequently". Thus, these instances of *based on* are used in contexts of developing or establishing tools or models.

All instances of *based on* share the semantic prosody of 'rigorous'. Similar to those in the top two sub-corpora, this semantic prosody is mainly indicated by the noun phrases immediately following the phrasal verb. As shown in Figure 4, the noun phrases following *based on* are "LCA and LCC", "the quantities of the chemicals determined", "the quantified damages", "chemical's reactivity lifetime", "the theory of interest in economics", "data provided by the US Department of Energy", "the concept of discounting future worth of money to present value" etc. These scientifically developed analytical methods, objectively measured figures, well-established theory and concepts, and data and information obtained from authority, are all rigorously and objectively established and derived.

It may be argued that "a hypothetical case of constructing" and "subjective judgments", in lines 2 and 3 respectively, are not rigorously and objectively derived. However, the inspection of larger co-texts of these two instances shows otherwise. The "hypothetical case" was used as an illustrative example in a discussion of life cycle assessment. Although the case was not a real one, it is not created randomly and arbitrarily. In order to fulfill the purpose of illustration, the case, though hypothetical, was based on real, typical construction scenarios and circumstances. As in the case of line 3, the extract of the complete line is given below:

> *The method most widely used for determining the weights is based on subjective*
>
> *judgments on the relative importance of the impacts of individual categories,*

which are typically solicited from a penal of experts or from a representative

cross-section of stakeholders.

The noun phrase "subjective judgments" itself sounds subjective rather than rigorous. However, it is stated later in the line that the judgments are sought from "a panel of experts" or "a representative cross-section of stakeholders". These people are supposed to have insider knowledge and expertise in the profession. Therefore, determinations or decisions using their judgments are also considered as rigorously and thoroughly arrived at.

Almost all the instances of *based on* across the sub-corpora share the co-selections described above. As shown in Table 5, in two-thirds of the sub-corpora, all the instances of *based on* are associated with the semantic prosody 'rigorous'. Even in the remaining ten sub-corpora, more than 90% of the instances have this semantic prosody. There are only nine sub-corpora in which not all instances of *based on* have the semantic prosody 'rigorous'. However, due to the very limited frequency of those instances which do not fit into the typical co-selections of *based on*, it is difficult to determine whether there is a pattern to the other meanings.

In Abstracts, there are two instances which do not fit into the general pattern of co-selection. The extract of the two lines are given below:

… perform all the functions of a traditional joypad, mouse, keyboard, and joystick based on the programmable driver.

… subtypes of avian influenza based on the H5 and H7 strains have proven to be some of the most lethal...

Although sharing the common colligational patterns of *verb phrase + based on + noun phrase* and *noun phrase + based on + noun phrase*, these two instances do not have the same semantic preference or semantic prosody. The co-occurring words are not related to any kinds of assessment or model development, and the noun phrases following *based on* do not share the same semantic prosody. Unlike the instances described above in which *based on* is used to express that assessments or model development are rigorously done and evidence-based, the first instance here seems to mean that the different functions "of a traditional joypad, mouse, keyboard, and joystick" are performed depending on "the programmable driver". The second instance seems to mean that "subtypes of avian influenza" originated from "the H5 and H7 strains". It is obvious that these two instances do not share the typical co-selection of *based on* in engineering genres, yet no detailed pattern of co-selection can be discerned due to the lack of further instances of use.

In Transaction Proceedings, five instances do not fall into the typical pattern. Figure 5 shows the concordance lines of the five instances. The colligational pattern of these instances is similar to the typical configurations of *based on*, which are *auxiliary verb + based on + noun phrase*, as in lines 1, 4 and 5, and *noun phrase + based on + noun phrase*, as in lines 2 and 3. In lines 1 to 4, the co-occurring words "manufacturing", "composites", "rail industry" and "products" suggest a semantic preference of 'manufacturing and production'. Such semantic preference indicates that *based on* in these instances is used in the context of manufacturing and production.

With regard to the semantic prosody, based on in lines 1 to 4 has a sense of 'comprising'. In these four lines, the noun phrases immediately following based on are the major components or ingredients of the noun phrases occurring in front of the phrasal

verb. For example, in line 1, "cotton textiles" is the major segment forming "Hong Kong's manufacturing", or the "composites" referred to in line 2 comprises "polyester and phenolics". Thus, based on in these instances tends to relate to the composites used for manufacturing and production.

```
1  In the beginning, Hong Kong's manufacturing was based on cotton textiles, but gradually adding wool
2      there is growing evidence that composites based on polyester and phenolics are finding greater
3      wavelength of previous commercial products based on multimode fibres. For example with equal
4   as P1, P2, P3 and P4. These four products are based on Canadian linseed oil. Different admixture
5  in Hong Kong. The 1200 requirement obviously is based on intuitive judgement and is contentious. The
```

Figure 5: Concordance lines of *based on* in Transaction Proceedings which do not fall into the typical pattern

It is interesting to note that *based on* in line 5 has a different semantic prosody. In fact, it seems to convey a sense contrary to the typical semantic prosody of 'rigorous', as suggested by the use of "intuitive judgment". This means that the "requirement" is created using intuition as a basis, which may be assertive and subjective in nature and thus seems contrary to the concept of rigour. Also, the co-occurrence of "contentious" seems to imply that it is unacceptable for the decision or judgment to be made upon "intuitive judgment". This may support the idea that the association of the phrasal verb *based on* with the semantic prosody of 'rigorous' is an unmarked co-selection in engineering genres.

The viability of co-selection of *based on* with something not rigorous can be further illustrated by the following instance:

...in view of the representative data in the range of kT 0.85 were based on extremely small number of data, this range was ignored in the curve fitting process.

The decision of "the range of kT 0.85" is made with the support of "extremely small number of data" and is not considered as scientific and objective, while the larger co-texts

state that "the range was ignored" and, therefore, once again the co-selection of *based on* conveys a very strong semantic prosody of 'rigorous'.

In Publicity Materials, almost all of the instances have the semantic prosody of 'rigorous'. The only exception is the following:

… established strong long-term customer and supplier relationships based on trust and integrity…

The phrasal verb in this instance does not share the semantic preference of 'assessments and evaluations' or 'model development'. It is used in talking about establishing relationships with customers and suppliers. As for semantic prosody, "trust and integrity" cannot be rigorously or objectively measured, rather, they are subjective concepts. Therefore this instance is not regarded as having the typical semantic prosody of 'rigorous'.

In Conference Proceedings, a single instance does not fall into the typical patterns of co-selection. The line is:

Up to 5% Compliance Based on single sample and single test

This line appears to be in point form and co-texts show that it is one of the points listed in a table or figure describing the characteristics and quality of an item. Since the original formatting of the texts has been removed by the software ConcGram, there is a lack of information to confirm how the co-texts should be arranged. However, "single sample and single test" does not sound rigorous and any testing results generated from only one sample are not scientifically convincing.

In Handbooks, an instance does not share the semantic preference and semantic prosody of other instances. The extract is as follows:

… a coach can still be arranged but a fee should be charged to the members joining the visit at a pre-announced rate based on self finance basis.

Unlike the other instances in which the noun phrase following the phrasal verb *based on* is the evidence or foundation used to support how the evaluation is done, the noun phrase following *based on* in this instance is a way of financing showing how the "fee" would be settled. Thus, *based on* in this instance has a different co-selection.

In Guides, three instances do not have the semantic prosody of 'rigorous' and they are shown below:

the ventilation rates of 10 litres per second per person for office and restaurants are based on a totally non-smoking environment

the ventilation rates, which are the governing standard, are based on a totally non-smoking environment

the assessor's judgement on relevance of information will be based on limited knowledge of the organization

The first two lines have the semantic preference of 'assessment and determination', as deduced from the phrase "the ventilation rates", which are likely determined using assessment and calculation methods. However, no evidence for the semantic prosody of 'rigorous' is observed. The noun phrase following *based on*, "a totally non-smoking environment", refers to the scenario in which the determination of "the ventilation rates"

103

was made, rather than the process of how it was rigorously determined. Therefore, *based on* in these two instances does not have the semantic prosody of 'rigorous'. As for the third instance, the co-occurring words "limited knowledge of the organization" is not considered to be rigorous support for the "judgment".

In Reports, three instances do not fall into the typical pattern. They are extracted as follows:

...these variables were not significant for vehicles based on Hong Kong Island and in Kowloon...

...these variables were not significant for vehicles based on Hong Kong Island and in Kowloon...

...network expansion should be based on the East Kowloon Line connecting with...

These instances tend to have a semantic preference of 'transportation' which is suggested by the words "vehicles" in the first two lines. In the third line, the wider co-texts provide "network expansion", which refers to the expansion of "rail". In light of the occurrence of phrases which denote locations, for example, "Hong Kong Island", "Kowloon" and "the East Kowloon Line", *based on* here has the semantic prosody of 'locating' in all three instances. Thus, the phrasal verb *based on* in these cases tends to express where the transportation is located.

In About Us, five instances do not have the typical semantic prosody (see Figure 6). These instances tend to have the semantic preference of 'company achievement' because of the presence of phrases such as "continuing professional development", "corporate culture", "long-term customer relationships" and "our success". Regarding

semantic prosody, a sense of 'boosting' is observed due to the use of phrases such as "our belief", "its commitment to quality", "trust and integrity" and "hard work". In fact, these things are abstract and cannot be quantified. It is not known how exactly they are related to the achievement of the companies. Thus, these lines do not have the semantic prosody of 'rigorous', rather, there is a semantic prosody of 'boosting' as the companies seem to be promoting their company values and success.

```
1   members continuing professional development, based on our belief that lifelong learning becomes
2 corporate culture and long term objectives are based on its commitment to quality,  innovation and
3        customer and supplier relationships are based on  trust and integrity - the hallmark of all
4   strong  long-term customer relationships are based on trust and integrity - the hallmark of all
5   any good suggestions. "Our success is mainly based on the hard work of our staff and not just our
```

Figure 6: Concordance lines of *based on* in About Us which do not fit into the typical pattern

In Media Releases, 11 instances have a different semantic prosody. Figure 7 shows the concordance lines of these instances. In lines 1, 2, 4 and 5, a semantic preference of 'achievement' is observed, particularly in relation to the co-occurrence of the word "establishment" found in lines 1 and 4. In lines 2 and 5, the contexts are about the establishment of "the institution" and "the 'Youth Day'". These four instances tend to have the semantic prosody of 'promising' as suggested by the words or phrases "the understanding", "goals and targets", "vision" and "vision and mission". These instances of *based on* are used to express how promising their company achievement is.

```
1  Hong Kong. Found in 1991 with the establishment based on the understanding of "The young is our
2    out the goals and targets of the institution based on the particular vision and mission to share
3  development of health food and  pharmaceuticals based on Chinese medicine'. To take forward this
4   for the establishment of an entreprise cutlure based on  valuing people.  Second, we should serve
5 decades. The 'Youth Day' commissioned by ITC is  based on the vision that youth is the pillar of
6  development of health food  and pharmaceuticals based on Chinese medicine. It will support strategic
7   technology to improve life quality. It is also based on the innovation of our  businesses and
8         from a fossil fuel-based economy to one based on renewable  energy can spur economic growth
9   century from a fishing village into an economy based on light industry. From there, we  have
10 development of health  food and pharmaceuticals based on Chinese medicine. "The imminent task of
11 development of health food  and pharmaceuticals based on Chinese medicine. While the government
```

Figure 7: Concordance lines of *based on* in Media Releases which do not fit into the typical pattern

In lines 3, 6, 10 and 11, a semantic preference of 'development of food and medicine' due to the co-occurrence of the phrases "the development of health food and pharmaceuticals" and "Chinese medicine". In lines 8 and 9, there tends to be a semantic preference of 'economy' indicated by the presence of the noun phrases "a fossil fuel-based economy" and "an economy". The phrasal verb in these six instances seem to have a semantic prosody of 'comprising' as "Chinese medicine" is used to develop "health food and pharmaceuticals", and "renewable energy" and "light industry" are the main sectors forming the economies mentioned in the lines. In other words, the items coming before *based on* form part of the items following the phrasal verb.

Line 7 discusses the success of the business and it seems to be using the success for self-promotion as it mentions "the innovation of our business", which is not accurately measurable. Thus, this instance tends to have a semantic prosody of 'promoting'.

The above analysis shows that *based on* is used quite consistently in engineering genres. The non-contiguous forms *based…on* and *on…based* are examined below to determine whether they also share the same use. Non-contiguous instances are found in fifteen sub-corpora. The overall frequency of non-contiguous instances is low. It ranges from one to 15 in each sub-corpus. Despite the change in form, most of the non-contiguous instances share the most frequent patterns of co-selection of *based on*. The sub-corpora of Agreements and Transaction Proceedings are used as example for illustration as they have higher frequencies.

In Agreements, there are seven non-contiguous instances, five of which are *based…on*, and two are *on…based*. The concordance lines are shown in Figure 8.

```
1      the state of the material at the next time based solely on the known state values of the current
2          The solid waste product calculation is based simply on the total quantities of all kinds of
3     ans services equipment. These are, however, based directly on the performance of the individual
4      Y * Radioactive waste Y * Slags ashes Y ** Based directly on 30 emissions, without
5     the Study. The ranking process was conducted based as far as possible on LCI data from the IVAM
6     the specific chillers on which the model was based. However, deviations of the model predictions
7     on which design cooling load calculations are based, the simplest method is to assume the sky
```

Figure 8: Concordance lines of non-contiguous instances in Agreements

For the instances in lines 1 to 5, the words or phrase intervening between *based* and *on* are adverbials, namely "solely", "simply", "directly" and "as far as possible". The phrasal verb in these five instances have the semantic prosody of 'rigorous' as evidenced by the co-occurrence of phrases directly following *on*, such as "the known state values", "the total quantities", "30 emissions", "LCI data", etc. These phrases provide certain statistical evidence. The addition of an adverbial in between *based* and *on* does not affect the meaning of the phrasal verb, but it adds to the intensity or extent of the meaning of the phrasal verb. For example, in line 1, the meaning emphasized by the occurrence of "solely" is that "the known state values of the current time" is the only basis used for the estimation of "the state of the material at the next time".

Regarding the *on…based* instances in lines 6 and 7, the word sequence between *on* and *based* is a relative pronoun followed by a noun phrase and then an auxiliary verb. It is a preposed structure of the phrasal verb; the relative pronoun "which" refers to the noun phrase preceding the particle *on*, i.e. "the specific chillers" and "the condition" respectively. It can be seen that they have the same semantic preference, which is suggested by the occurrence of phrases such as "the model" and "calculations", and the same semantic prosody as well, which is suggested by the presence of "the specific chillers" and "the condition". Therefore, *on…based* share similar co-selection of typical patterns of *based on*.

Figure 9 shows the 15 non-contiguous instances in Transaction Proceedings. They are all constituency variations, i.e. *based…on* instances, except line 14 which displays an *on…based* instance. For the *based…on* instances, the words falling in the middle of *based* and *on* are adverbs, such as "also", "solely", "largely", "essentially". The addition of the intervening word does not change the meaning of the phrasal verb. These instances have the semantic prosody of 'rigour', for example, "the indicative results in Figure 13" in line 2 or "rigid prescriptive criteria" in line 3. They suggest that the determination made, the calculations or the system and model are supported by rigorous evidence.

```
1     be possible if the equivalent performance is based also on energy use rather than simply the
2   reduction factor applies for normal structures. Based again on the indicative results in Figure 13
3  the control could be adversely affected if it is based solely on rigid prescriptive criteria. In
4   selection of phase-mixed coating materials were based solely on the thermal characteristics.  Chung
5   identified. The problems with a scheme that is based solely on rigid prescriptive criteria and
6  the control could be seriously affected if it is based solely on rigid prescriptive criteria. For
7  data beyond those for compliance checks that are based solely on the prescriptive criteria. For
8     malaise point out the dangers of an economy based purely on finance and real estate, and
9     control over the secondary-loop pumps can be based solely on the total tow rate demand of the
10    the design of non-flexural components will be based largely on these earlier papers and other
11   amplification (FASA) model, as the concept is based essentially on treating the soil column as
12   derive ground motion attenuation relationships based specifically on earthquake characteristics
13  Figure 2, the shear demand Ve on the member is based not on the applied loads, but on the probable
14       and architecture on which the system is based. Section 3 presents the results the system's
15     But communication on the Internet is still based largely on the English language. This is only
```

Figure 9: Concordance lines of non-contiguous instances in Transaction Proceedings

In line 8, *based…on* tends to have a different semantic prosody from other instances. The noun phrase "finance and real estate" is an industry which forms part of the "economy". The rigorous sense is not observed. Rather, there seems to be a semantic prosody of 'comprising' derived from the notion of the economy formed by "finance and real estate". The adverbial "purely" reinforces the view that the "economy" is made up of "finance and real estate" only.

It is also noted that the phrasal verb in line 15 has a different semantic preference and semantic prosody. The co-occurring phrases "communication on the Internet" and "the English Language" do not share the semantic prosody of 'rigorous' or 'comprising'.

It might be a whole different semantic prosody which is not observed in any other instance across all the sub-corpora. However, this single instance does not provide enough data for the description of a pattern.

To summarise, *based on* has a very strong tendency for particular co-selections in engineering texts. It is always used in the passive. The typical configurations of *based on* are *auxiliary verb + based on + noun phrase* and *verb phrase + based on + noun phrase*. There are two major semantic preferences. One is 'assessment and making determination', in other words, the phrasal verb is used in the contexts of assessing or evaluating something or making decisions. The other semantic preference is 'model development' as the texts are about developing certain models, methods, systems, or theories. A very strong semantic prosody of 'rigorous' is observed in the vast majority of the instances. This semantic prosody indicates that the assessments done, decisions made or the models developed are rigorous and objective, and are grounded by data and evidence. The findings also show that *based on* may have other semantic preferences and prosodies. For example, some instances of *based on* are used in the context of manufacturing and production, and they have the semantic prosody of 'comprising'. In other words, *based on* is used to express what certain products are composed of. Other instances of *based on* convey a semantic preference of 'transportation' and a semantic prosody of 'locating'. Such instances show where the transporting vehicles are from but are few in number and so the description of those semantic prosodies may be incomplete.

The non-contiguous instances, whether they appear as *based…on* or *on…based*, share very similar co-selection of the typical *based on* instances as described above. For the *based…on* instances, the typical configuration is *auxiliary verb + based + adverb +*

on + noun phrase. The intervening words which serve as an adverbial to signal the extent to which the assessment or determination is rigorously conducted, or the amount of data or evidence that supports the assessment and determination. The semantic prosody of the *based…on* instances is the same as the prevailing semantic prosody of the most typical *based on* instances, which have the meaning of 'rigorous'. The intervening word adds a sense of intensity or emphasis. As for the *on…based* instances, they are a preposed structure of *based on*, and the textual environments also show that they share very similar co-selection. Therefore, *based…on* and *on…based* are phraseological variations of the phrasal verb *based on*.

Considering that the semantic preference and semantic prosody are highly specific to *based on*, regardless of the sub-corpus in which the phrasal verb is used, it is concluded that the phrasal verb *based on* is specifically used to provide evidence and support for assessments done and determinations made in all engineering texts. Since engineering involves calculations and assessments which are related to the quality and safety of products, they need to be credible and accurate. Therefore this use of *based on* with a rigorous foundation and rigorous evidence is very frequent in engineering texts so as to justify those assessments and calculations.

In a corpus study of intertextuality in business emails, Warren (2016) observed another genre-specific use of this phrasal verb. The phrasal verb is always followed by a noun phrase which frequently refers to a previous text in the email communication. The function considered for the phraseology *based on + noun phrase* in business emails is to signal intertextuality, or in other words, refer to prior discourses (Warren, 2016). The use of *based on* to signal intertextuality in business emails may seemingly be irrelevant to the

use and function of *based on* in engineering genres. However, it is argued that the semantic prosodies of *based on* in the two types of data are somehow related but have different perspectives. Although the context in which *based on* is used in business emails may not be about assessments and evaluation, or how new products or models are developed, using *based on* to introduce the prior texts or discourses, for example, "Axxxx's email", "C3 & C4 data", "earliest of END & CANCEL events", "the test scenario worksheet", "this comment", "your advice and suggestion", serves to be provide support or evidence for current decision and action. The fact that many of the noun phrases following *based on* in the current study identify previous texts or materials, such as "a marking scheme" in Tender Notices, "statistics from over 1000 failures" in Transaction Discussions and "data provided by the US Department of Energy" in Agreements, indicates that the sense of intertextuality is observed in the use of *based on* in engineering genres as well.

Correspondence of the specificity of *based on* to particular genres or registers is also observed in some previous studies. The phrasal verb is found to appear more frequently in academic writing (Biber et al., 1999; Liu, 2012) and abstracts of engineering journals (Nakano & Koyama, 2005). Academic writing, journal article abstracts and engineering texts are similar in their formal and technical style. Biber et al. (1999) reported that almost half of the phrasal verbs (they term them 'prepositional verbs') identified in academic writing are in the causative and existence semantic domains, and *based on* (Biber et al. present the phrasal verb as *be based on*) is categorised in the existence or relationship domain (pp. 418-419). This indicates that *based on* is commonly used to express the existence of something or a relationship between different items. Regarding Nakano and Koyama's (2005) studies, abstracts of engineering journals may

be considered as a combination which interestingly shares the nature of both engineering texts and academic writing. Thus, in academic writing, abstracts of journal articles, as well as engineering genres, *based on* is formally used to provide support for the existence of something or for decisions made.

4.2.2 *Base on*

Although *based on* is the most frequently used phrasal verb in the Hong Kong Engineering Corpus, other inflectional forms of BASE ON have very low occurrences. The phrasal verb *base on* has the second highest frequency among the inflectional forms, but its frequency in the whole corpus is only 49. Table 7 shows the 13 sub-corpora in which *base on* is found but only in five is the frequency greater than one.

Table 7: Sub-corpora in which *base on* occurs and the respective frequencies

	Sub-corpora	size	Freq	%	Rigour
1	Reports	979,170	25	0.0026%	100%
2	Notes	156,255	3	0.0019%	100%
3	Technical Papers	65,731	1	0.0015%	
4	Transaction Notes (HKIE)	79,058	1	0.0013%	100%
5	Project Summaries	115,829	1	0.0009%	100%
6	Media Releases	1,566,742	8	0.0005%	100%
7	Conference Proceedings	196,498	1	0.0005%	100%
8	Guides	783,805	3	0.0004%	100%
9	Product Descriptions	611,549	2	0.0003%	100%
10	Publicity Material	599,407	1	0.0002%	100%
11	About Us	647,013	1	0.0002%	100%
12	Code of Practice	997,228	1	0.0001%	
13	Transaction Proceedings (HKIE)	1,055,248	1	0.0001%	100%

Reports has both the highest frequency and percentage of occurrence of *base on*. Figure 10 shows a sample of concordance lines of the phrasal verb in Reports. In lines 1,

3 and 5, *base on* appears at the beginning of the clause. "Future Key Issues" is the title of the section. This structure is found in 60% of the lines (15 times). In 24% of the lines (6 times), *base on* immediately follows a noun phrase, such as lines 2, 4 and 6. In the remaining 16% of the lines (4 times), *base on* comes after a verb phrase. This shows that the patterns of colligation of *base on* is very similar to those of *based on*.

```
1   in this reporting month.    Future Key Issues Base on the site inspections and forecast of
2       0.13 mg L Total Phosphate In house method base on ASTM D 515-88 0.05 mg L pH (at 25oC) APHA 19
3  in this reporting month.    Future Key Issues Base on the site inspections and forecast of
4       0.13 mg L Total Phosphate In house method base on ASTM D 515-88 0.05 mg L pH (at 25oC) APHA 19
5  in this reporting month.    Future Key Issues Base on the site inspections and forecast of
6       0.13 mg L Total Phosphate In house method base on ASTM D 515-88 0.05 mg L pH (at 25oC) APHA 19
7  per share  Loss per share has been calculated base on the Group's unaudited loss attributable to
8  share  Earnings per share has been calculated base on the unaudited consolidated profit attributable
9  share  Earnings per share has been calculated base on the unaudited consolidated profit attributable
10 per share  Loss per share has been calculated base on the unaudited consolidated loss attributable
```

Figure 10: Sample concordance lines of *base on* in Reports

Collocates occurring to the left of *base on* include "in" (16 times), "issues" (15 times), "key" (15 times), "future" (15 times), "this" (10 times), "month" (10 times), "house" (6 times), "method" (6 times), "calculated" (4 times), "per" (4 times), "been" (4 times), "has" (4 times), "earnings" (2 times), "loss" (2 times), etc. To the right, there are "the" (19 times), "of" (17 times), "works" (15 times), "site" (15 times), "inspections" (15 times), "forecast" (15 times), "engineering" (15 times), "in" (15 times), "and" (15 times), "unaudited" (4 times), "to" (4 times), "attributable" (4 times), "consolidated" (3 times), "shareholders" (2 times), "profit" (2 times), etc. The semantic preference of 'evaluation and determination' is observed (100%) due to the occurrence of words "calculated" and "method". The texts are related to testing and calculation. In lines 1, 3 and 5, no words seem to be explicitly related to evaluation and determination, but the texts in that section are about determining future key issues. Thus the instances also fall into the semantic preference of 'evaluation and determination'.

Regarding the semantic prosody, all the instances have the sense of 'rigorous'. This is due to the noun phrases occurring to the right of *base on*. The phrases "the site inspections and forecast of engineering works", "ASTM D 515-88", and "the unaudited consolidated profit" are the supporting evidence justifying the evaluation or determination.

In Media Releases, there are eight instances of *base on*. Figure 11 shows the concordance lines of these eight instances. The pattern of colligation in lines 2 to 7 is *verb phrase + base on + noun phrase* (6 times, 75%), whereas the configuration in lines 1 and 8 is *base on + noun phrase* (2 times, 25%) in which *base on* is the beginning of the clause. For collocational pattern, collocates to the left of *base on* are "calculated" (4 times), "per" (4 times), "share" (4 times), and "loss" (3 times). To the right, there are "unaudited" (4 times), "attributable" (3 times), "loss" (2 times), and "consolidated" (2 times).

```
1    the Banks for a waiver of the above breach. Base on the longterm relationship and good track
2  per share Loss per share has been calculated base on the Group's unaudited loss attributable to
3  a demonstration. The projects were accessed base on creativity, innovation and whether they are
4  per share  Loss per share has been calculated base on the Group's unaudited loss attributable to
5         Increase in rental income is expected base on the current market conditions. With the
6  share Earnings per share has been calculated base on the unaudited consolidated profit attributable
7  per share Loss per share has been calculated base on the unaudited consolidated loss attributable
8    close twenty times to test its durability. Base on the expert's advice, we arranged remedial
```

Figure 11: Concordance lines of *base on* in Media Releases

The semantic preference of 'evaluation and determination' (75%) is observed due to the occurrence of words "calculated", "expected" and "arranged". They suggest that some evaluations and determinations are made. These instances also share the semantic prosody of 'rigorous' which is evident in the phrases "the Group's unaudited loss", "the current market conditions", "the unaudited consolidated profit" and "the expert's advice". It is therefore suggested that the calculation or the determination is not arbitrarily made, but is supported by objective and rigorously measured figures.

The instances in other sub-corpora also share this semantic prosody. There are a few instances in which a new colligational pattern is found. For example, the following are the concordance lines of the *base on* instances in Notes:

STAJDSD will base on the Finance Code shown in the account report from HyD to settle the XP fees…

The number of exhaust fan to be used would base on the capacity of the individual fan(s)…

The number of exhaust fan to be used would base on the capacity of the individual fan(s)…

The pattern of colligation is *modal verb + base on + noun phrase* in these three instances. This configuration is not found in *based on* instances due to the constraint of grammatical rules. Despite having different colligation, these three instances have the same semantic preference of 'evaluation and determination' as the first line is evaluating the settlement of "the XP fees", and the second and the third lines are about the determination of "the number of exhaust fan to be used". Since these evaluations and decisions are grounded by "the Finance Code", which is well-established and widely-accepted, and "the capacity of the individual fan(s)", which is objective, these instances of *base on* have the semantic prosody of 'rigorous'.

The findings show that the co-selection of most of the instances of *base on* are similar to those of *based on*. However, it is argued that many of the instances, such as the instances in Reports and Media Releases, should be used in the passive form as *based on*, rather than the bare form *base on*. These are examples of non-standard expressions of a

kind possibly common in Hong Kong. There are also examples which seem to call for the use of alternative forms to realize meanings in a standard way.

There are 13 non-contiguous instances found in seven sub-corpora, 11 of which are *base…on*, and two are *on…base*. The following are the concordance lines of the four *base…on* instances in About Us:

…Member States may base their distribution of allowances on "average emissions"…

Members may base their distribution of allowances on achievable progress…

…Member States may base their distribution of allowances on "average emissions"…

Members may base their distribution of allowances on achievable progress…

The colligational pattern of these instances are *modal verb + base + noun phrase + on + noun phrase*. They are about the determination of how to distribute allowances, thus these instances have the semantic preference of 'evaluation and determination'. Since the determination is grounded on "average emissions" and "achievable progress", it suggests that the instances share the semantic prosody of 'rigorous'.

In fact, these instances are paraphrases of *based on* because they can be re-structured to form passive sentences. For example:

…their distribution of allowances may be based on "average emissions"…

Such feature of transforming *base…on* to passive voice shows that these *base…on* instances are variations of *based on* instances, thus it is suggested that they can be regarded as the same phrasal verb but in different forms and phraseological variation.

As for *on…base* instances, there are only two times of occurrence. The extract of these two instances respectively from Reports and Transaction Proceedings are given below:

The COT also considered the primate study of Rier et al (1993) on which Health Council (1996) chose to base their revised TDI…

… provided MTRC with reliable information on which to base their assessment of tenders proposals.

As for the *on…based* instances, these two *on…base* instances are preposed structures. The relative pronoun "which" refers to the noun phrases "the primate study of Rier et al (1993)" and "reliable information" which share the semantic prosody of 'rigorous'. Thus, these two instances are paraphrases of *base on* and *base…on*, and share the same co-selection. For this reason, *on…base* is considered to be a phraseological variation of *base on*.

4.2.3 *Basing on*

The other inflectional form *basing on* occurs 16 times in total in the whole corpus, and these instances are spread across nine sub-corpora (see Table 8). For the sake of illustration, Conference Proceedings and Transaction Proceedings are used as examples as they have the highest number of instances of *basing on*.

Table 8: Sub-corpora in which *basing on* is found and its frequencies

	Sub-corpora	Size	Frequency	%
1	Conference Proceedings	196,498	3	0.0015%
2	Position Documents	75,660	1	0.0013%
3	Abstract	94,671	1	0.0011%
5	Product Descriptions	611,549	2	0.0003%
6	Transaction Proceedings (HKIE)	1,055,248	3	0.0003%
7	Guides	783,805	2	0.0003%
8	Code of Practice	997,228	2	0.0002%
9	Publicity Material	599,407	1	0.0002%

The three instances in Conference Proceedings are extracted as follows:

Fees are charged basing on the "Standard daily charges of project consultants".

The design performance of the BIPV system was calculated basing on the solar irradiance…

The design performance of the BIPV system was calculated basing on the solar irradiance…

In all three instances, *basing on* comes after a verb phrase and is then followed by a noun phrase. Thus, it has the configuration of *verb phrase + basing on + noun phrase*. Due to the use of the phrase and the phrases "fees are charged" and "calculated", these instances have the semantic preference of 'calculation' (100%). The phrases "the 'Standard daily charges of project consultants'" and "the solar irradiance" suggest that the calculations are not done arbitrarily but with reference to some standard or objective measurable index. Therefore, these three instances have the semantic prosody of 'rigorous' (100%).

The three instances of *basing on* in Transaction Proceedings are extracted as follows:

...the 3D-FEM of Tsing Ma Bridge was built by MSC/PATRAN [13] basing on as-built drawings...

... incorporated into the finite element model for normal mode analysis basing on Lanczos Method...

...sampled accelerometer data (basing on a sampling rate of 51.2 Hz) were stored...

The first two instances have the configuration of *verb phrase + basing on + noun phrase*. In the third instance, *basing on* followed by a noun phrase occurs in a bracket. They act as extra information for the preceding noun phrase. Thus, it could be considered as having the colligation of *noun phrase + basing on + noun phrase*. It is not obvious whether these instances share a particular semantic preference. But inspection of wider co-texts shows that the first and the second instances are used in the contexts of developing a model, i.e. "3D-FEM of Tsing Ma Bridge". Thus, they can be regarded as having the semantic preference of 'model development'. The third instance discusses certain analytical data. Also, the wider co-texts show that the data is used for the analytical study of the development of the "3D-FEM of Tsing Ma Bridge". Therefore, all these three instances have the semantic preference of 'model development'. As for the semantic prosody, the noun phrases following *basing on,* such as "Lanczos Method" in the second instance, suggest that the model and the data are developed with references and foundation. Thus, these instances of *basing on* have the semantic prosody of 'rigorous'.

The analysis of the other *basing on* instances shows that *basing on* has the semantic prosody of 'rigorous' across all the sub-corpora in which it is used. In fact, it is argued that all the instances of *basing on* in the data should for the sake of accuracy be replaced by *based on*. Such use of *basing on* may possibly be error or special style of writing of the writers of the texts.

There is a single occurrence of the non-contiguous instance and it occurs in the form of *basing…on*. The instance is found in Standards and is given below:

…(c) basing such talks on kits published by the Hong Kong Construction Association Ltd., …

This line is one of the points listing the characteristics of "item coverage". So, the configuration of this instance is *noun phrase + basing + noun phrase + on + noun phrase*. No pattern can be described as there is only this single instance. However, since the "talks" are designed using "kits published by the Hong Kong Construction Association Ltd." which are well-established references, this use of *basing…on* can be regarded as having the semantic prosody of 'rigorous'.

4.2.4 *Bases on*

The other inflectional form *bases on* is used the least. There are only three total instances in the corpus, and they are all from Codes of Practice. The instances are extracted as follows:

…the elastic moment as discussed in clause 8.1.1 of the Code, which bases on elastic behavior…

Compressive strength bases on buckling curve (c) (Table 8.7)

Compressive strength bases on buckling curve (c) (Table 8.7)

The phrasal verb *bases on* is used in the simple present tense with the preceding noun phrase in singular form. Thus the pattern of colligation of *bases on* is *noun phrase + bases on + noun phrase*. These instances have the semantic preference of 'calculations

and determinations' as the phrases "the elastic moment" and "compressive strength" suggest that these instances are discussing the calculation or determination of certain capacity and strength. And based on the co-texts, these capacity and strength are related to the design of some structural elements, such as steel. The noun phrases following *bases on* in the three instances serve to be the foundation or the reference for the calculation and determination. Thus, they have the semantic prosody of 'rigorous'.

Two non-contiguous instances in the form of *bases…on* are found in the corpus. They are found in Circular Letters. The extract of the lines is given below:

The Group bases its estimates on historical results…

The Group bases its estimates on historical results…

The two instances appear to be identical, but they are found in different places in the sub-corpus. The colligation of *bases…on* is *noun phrase + bases + noun phrase + on + noun phrase*. The word "estimates" suggests that the instances have the semantic preference of 'evaluation and estimation'. And the occurrence of "historical results" shows that the estimation is derived through a reasonably reliable data, rather than arbitrarily made.

These *bases…on* instances are used in the present tense active voice. If they are re-structured to be passive voice, they become typical *based on* instances. The intervening word between *bases* and *on* in these two instances would become the subject of the passive voice structure, in other words, the noun phrase to the left of typical *based on* instances. In fact, both *base…on* and *basing…on* can also be re-structured to be passive use, which is the typical use of *based on*. The intervening noun phrase between *base* and *on* in the case of *base…on*, and between *basing* and *on* in the case of *basing…on* is the noun phrase typically preceding *based on*.

4.2.5 Summary

To summarise, *based on* is the most frequently found phrasal verb in the Hong Kong Engineering Corpus. It is used in all the sub-corpora, except in Letters to Editor. Its frequency is not evenly distributed across the sub-corpora, the percentage of occurrence of *based on* in the sub-corpora ranges from 0.0029% to 0.1414%. The top three sub-corpora in which *based on* is found are Tender Notices, Transaction Discussions and Agreements. Despite that the relative frequencies in each sub-corpus vary, it is found that almost all of the instances of *based on* share the same co-selection. All the *based on* instances are used in the passive voice. They have a strong semantic preference of 'evaluation or determination' or 'assessment and evaluation', meaning that they are primarily used in the contexts of evaluating something or making certain determinations. More than 95% of the *based on* instances have the semantic prosody of 'rigorous'. This is because the noun phrases following *based on* always represent rigorous and objective evidence, support or foundation which justifies the evaluation or decision made. Thus, *based on* is used to express the meaning of 'justifying evaluations and determinations with rigorous evidence'.

The findings also indicate that the non-contiguous *based…on* and *on…based* share the same semantic preference and semantic prosody as those of *based on*. For *based…on*, the intervening words are adverbs or adverbials, such as "solely" and "directly". They emphasize the extent to which the evaluation or determination depends on the evidence. The addition of the intervening words does not change the core meaning of the phrasal verb but intensify it. As for *on…based* instances, they are simply displaying the preposed structure of the phrasal verb *based on* as this is one of the syntactic flexibility of the phrasal verb. It is shown that if the preposed structure *on…based* is re-written, it becomes

122

a typical *based on* instance. Therefore, *based…on* and *on…based* are considered as phraseological variations of the phrasal verb *based on* and they share the same co-selection. It should be noted that, although *based on* possesses the syntactic flexibility to demonstrate constituency variation or preposed structure, these might be considered to be marked configurations when we recall how limited their occurrence is compared to *based on*.

All three other inflectional forms *base on, basing on* and *bases on* are found in the corpus, but the frequencies of these other forms differ to a large extent from the frequency of *based on*. While *based on* is the most frequent phrasal verb in the corpus with more than 3,500 times of occurrence, the frequencies of the other three forms range only from three to 49 times. And there are only a few non-contiguous instances of the other inflectional forms.

BASE ON is a doubly transitive Type II phrasal verb, with a verb and a prepositional particle. The fact that most of the non-contiguous instances in the form of *base…on*, *basing…on* and *bases…on* can be transformed to the passive structure, which is the most frequent *based on* instances found in the corpus, is due to its own syntactic characteristic. An example is given below:

… bases + noun phrase 1 + on + noun phrase 2 …

… noun phrase 1 + (auxiliary verb) + based on + noun phrase 2 …

As described in Chapter 2, a doubly transitive Type II phrasal verb takes two objects. Noun phrase 1 in the first line is considered as the object of the verb *bases* and noun phrase 2 is the object taken by the prepositional particle *on*. When a passive structure is

used, *bases…on* becomes *based on*, probably with a preceding auxiliary verb; noun phrase 1 becomes the subject of *based on* and noun phrase 2 becomes the object.

All three other inflectional forms share the same semantic preference and semantic prosody, regardless of whether they occur as contiguous or non-contiguous instances. It is reasonable to argue that *based on, base on, basing on* and *bases on* are phraseological variation of the same phrasal verb as they share the same semantic prosody, however, they have such a large difference in their frequency of use. And such difference in frequency implies that although they all share the same meaning, it is the norm to use the phrasal verb in the passive form and contiguously as in *based on*. This high tendency for using *based on* in the passive voice, even though it can be used as the past tense active voice, corresponded to Biber, et al.'s (1999, p. 418) treatment to present the phrasal verb as the sequence of *be based on*. Thus if they were all lumped together under the lemma BASE ON, it would be confusing if information regarding frequency variation is omitted. It is recommended that more accurate and detailed description of the most typical configuration and genre-specific use be pointed out in dictionaries and grammar references.

4.3 Use of phrasal verbs: CARRY OUT

Among the inflectional forms of CARRY OUT, *carried out* and *carry out* are frequently used in engineering genres. This section examines and discusses the use and co-selections of *carried out, carry out, carrying out* and *carries out*.

4.3.1 *Carried out*

The form *carried out* is the second most frequent phrasal verb found in the Hong Kong Engineering Corpus. This phrasal verb has an overall frequency of 3,785 times (0.0410%) in the whole corpus, and is found in 29 sub-corpora, as shown in Table 9 below. The table shows the frequencies of *carried out* in each of the sub-corpora.

Table 9: Frequencies and percentage of occurrence of *carried out*

	Sub-corpora	Size	Frequency	%
1	Ordinances	139,176	128	0.0920%
2	Code of Practice	997,228	807	0.0809%
3	Notes	156,255	122	0.0781%
4	Speeches	2,822	2	0.0709%
5	Reports	979,170	631	0.0660%
6	Guides	783,805	458	0.0584%
7	Transaction Discussions (HKIE)	7,149	4	0.0560%
8	Manuals	296,299	162	0.0547%
9	Handbooks	67,284	32	0.0476%
10	Publicity Material	599,407	269	0.0449%
11	Abstracts	94,671	38	0.0401%
12	Transaction Proceedings (HKIE)	1,055,248	422	0.0400%
13	Fact Sheets	26,059	10	0.0384%
14	Conference Proceedings	196,498	75	0.0382%
15	Transaction Notes (HKIE)	79,058	29	0.0367%
16	Standards	136,024	44	0.0323%
17	Frequently Asked Questions	55,726	16	0.0287%
18	Product Descriptions	611,549	150	0.0245%
19	Tender Notices	4,242	1	0.0236%
20	Review Papers	106,506	24	0.0225%

21	Consultation Papers	111,494	25	0.0224%
22	Project Summaries	115,829	21	0.0181%
23	Q & A	27,703	5	0.0180%
24	Position Documents	75,660	11	0.0145%
25	Media Releases	1,566,742	227	0.0145%
26	Technical Papers	65,731	6	0.0091%
27	Circular Letters	143,313	13	0.0091%
28	About Us	647,013	46	0.0071%
29	Agreements	127,895	7	0.0055%

This phrasal verb is predominantly used as a past participle in the passive voice. The co-selections of *carried out* are similar across all sub-corpora in which instances of the verb are found. Ordinances is the sub-corpus with the highest percentage of occurrences of *carried out*. There are 128 instances (0.092% of the size of the sub-corpus). Figure 12 shows a sample of concordance lines of *carried out* in this sub-corpus.

```
1     requirements. 1.7.6 All site work to be carried out shall comply with:- a) The Construction Site
2   ferrules for identification which shall be carried out in a neat and systematic manner and finished
3 anual cleaning. ii) Manual cleaning shall be carried out by a power tool with rotary wire brush and
4   If blast cleaning is required, it shall be carried out in accordance with BS 5493:1977 Code of
5 be by hand brushing. Painting should only be carried out on dry surfaces when the relative humidity
6 raying or other approved method. It shall be carried out in a dry, draught-free atmosphere to avoid
7 Structural Steel 1.16.1 All welding shall be carried out only by welders of recognized proficiency,
8 carrying fluids. 1.16.2 Welding shall not be carried out under unfavourable conditions. The
9 or acceptance by the Contractor until he has carried out his own tests and has established that the
10 to be so used. (7) Any test and examination carried out before the commencement of the Ordinance on-
```

Figure 12: Sample concordance lines of *carried out* in Ordinances

Of the 128 instances, most are used in the passive voice (91.4%), with only 11 instances (8.6%) used in the active voice. This is reflected in the colligational patterns on the left side of the phrasal verb. In 94 instances (73.4%), there is an auxiliary verb in either L1 or L2 position. The auxiliary verb is realized as "be" (54 times, 57.4%), "been" (9 times, 9.6%), "is" (9 times, 9.6%), "being" (8 times, 8.5%), "was" (5 times, 5.3%), "has" (5 times, 5.3%), "are" (2 times, 2.1%) and "have" (2 times, 2.1%). Except for "has" and "have", the use of the other auxiliary verbs indicates that *carried out* is used in the passive voice. For those instances with "be", the auxiliary verb "be" is preceded by a

126

modal verb (31 times, 57.4%) or an infinitive marker (23 times, 42.6%). Another pattern observed in this sub-corpus is that *carried out* comes after a noun phrase in 30 instances (23.4%). The phrasal verb is not used in the past tense in any of these instances but rather as a past participle functioning as an adjective phrase. For example, in line 10 of Figure 12 above, *carried out* functions as an adjective phrase modifying the preceding noun phrase, i.e. "Any test and examination". With regard to the instances which are used in the active voice, the items preceding the phrasal verb are auxiliary verb (7 times, 5.5%) which is realized as "has" or "have", relative pronoun (3 times, 2.3%) and a single occurrence of a pronoun.

To the right of *carried out*, a prepositional phrase is typically found (80 times, 62.5%). In some instances, *carried out* is the last item in the clause (26 times, 20.3%). And in some cases, a noun phrase comes after the phrasal verb (13 times, 10.2%). Therefore, the most frequent configuration of *carried out* in Ordinances is:

(modal verb) + auxiliary verb + carried out + prepositional phrase

Top collocates to the left of *carried out* are "be" (68 times), "the" (64 times), "work" (38 times), "shall" (38 times), "and" (37 times), "to" (36 times), "of" (33 times), "or" (28 times), "a" (26 times), "is" (23 times), "test" (21 times), "such" (21 times), "any" (18 times), "for" (17 times), "examination" (17 times), "works" (8 times), "contractor" (5 times), "inspections" (5 times), "lift" (5 times), etc. On the right side, top collocates include "the" (92 times), "a" (47 times), "of" (46 times), "in" (38 times), "and" (35 times), "to" (29 times), "by" (25 times), "or" (23 times), "on" (21 times), "such" (19 times), "shall" (19 times), "any" (18 times), "under" (15 times), "with" (14 times), "be" (13 times), "gas" (12 times), "for" (12 times), "is" (11 times), "authority" (11 times), "which"

127

(10 times), etc. In terms of the semantic preference, all of the instances share the semantic set of 'engineering-related works'. It is found that, *carried out* is used in close proximity with words or phrases which are related to different aspects of engineering, such as "construction work", "welding", "repairs", "test and examination", "lift work", "gas installation work", "periodic inspections", "pipe", "meter", "trench works". These words or phrases represent different kinds of construction or installation works, repairs and maintenance, and testing and inspections which are typical tasks engineers perform or encounter in their daily lives.

Regarding the semantic prosody, the majority of the instances are used with a sense of 'requirement' (126 times, 98.4%). This is due to the occurrence of words or phrases such as "shall comply with" as in line 1 in Figure 12, and "in accordance with BS 5493:1977 Code of Practice" as in line 4. In other instances, "shall" or "should" are found and convey deontic modality. There are also phrases such as "pursuant to subregulation", "as is specified in subregulation" and "required by this Ordinance". The data suggests that requirements are set to regulate work or behaviour in different ways; for example, how works are performed, as in "Manual cleaning shall be carried out by a power tool with rotary wire brush…", in what environment the works are performed, as in "It shall be carried out in a dry, draught-free atmosphere…", to what standard the works are conducted, as in "…to ensure it has been carried out in accordance with the requirements under the Ordinance…", or whose obligation to conduct the works as in "Every examination under subregulation (1) shall be carried out by a competent person…". This semantic prosody can be explained by the fact that engineering works are always governed by strict rules, standard procedures, and generally-accepted practices in order to assure quality of work and the safety of people.

There are only two instances which do not share this semantic prosody. Both instances convey the sense of 'defining' rather than 'requirement'. One of the instances is shown here:

"charge" means any charge for water, any fee, the cost of repairs or other works carried out by the Water Authority under section 17, and any other charge, including a surcharge, which is payable under this Ordinance;

The phrasal verb *carried out*, in this case, is used in the definition of terminology which has specialised meanings in the context. Although the phrasal verb in these two instances also co-occurs with words or phrases denoting requirements, it does not function to state and specify the requirements for performing engineering works, rather, they are used to provide clarifications to describe clearly the nature, scope and meaning of a term in a particular context.

Table 9 shows the frequency and percentage of occurrence of the major semantic prosody for *carried out* in the sub-corpora. In 21 out of 29 sub-corpora (72.4%), there is over 70% of occurrences of the prosody of 'requirement', and 14 of which have an occurrence of over 90%. While most of the *carried out* instances across the sub-corpora carry a strong sense of 'requirement', not all the instances necessarily have the semantic prosody as explicitly as described above. The instances in the sub-corpus of Reports are used for illustration with further details.

Table 10: Frequencies and percentage of occurrence of *carried out* having the semantic prosody of 'requirement'

	Sub-corpora	Size	Freq.	%	Requirement	%
1	Ordinances	139,176	128	0.0920%	126	98.4%
2	Code of Practice	997,228	807	0.0809%	796	98.6%
3	Notes	156,255	122	0.0781%	121	99.2%
4	Speeches	2,822	2	0.0709%	1	50%
5	Reports	979,170	631	0.0660%	501	79.4%
6	Guides	783,805	458	0.0584%	418	91.3%
7	Transaction Discussions (HKIE)	7,149	4	0.0560%	2	50%
8	Manuals	296,299	162	0.0547%	162	100%
9	Handbooks	67,284	32	0.0476%	31	96.9%
10	Publicity Material	599,407	269	0.0449%	257	95.5%
11	Abstracts	94,671	38	0.0401%	12	31.6%
12	Transaction Proceedings (HKIE)	1,055,248	422	0.0400%	176	41.7%
13	Fact Sheets	26,059	10	0.0384%	10	100%
14	Conference Proceedings	196,498	75	0.0382%	70	93.3%
15	Transaction Notes (HKIE)	79,058	29	0.0367%	3	10.3%
16	Standards	136,024	44	0.0323%	43	97.7%
17	Frequently Asked Questions	55,726	16	0.0287%	14	87.5%
18	Product Descriptions	611,549	150	0.0245%	114	76%
19	Tender Notices	4,242	1	0.0236%	1	100%
20	Review Papers	106,506	24	0.0225%	20	83.3%
21	Consultation Papers	111,494	25	0.0224%	21	84%
22	Project Summaries	115,829	21	0.0181%	19	90.5%
23	Q & A	27,703	5	0.0180%	5	100%
24	Position Documents	75,660	11	0.0145%	6	54.5%
25	Media Releases	1,566,742	227	0.0145%	153	67.4%
26	Technical Papers	65,731	6	0.0091%	4	66.7%
27	Circular Letters	143,313	13	0.0091%	13	100%
28	About Us	647,013	46	0.0071%	35	76.1%
29	Agreements	127,895	7	0.0055%	5	71.4%

In total, there are 631 instances of *carried out* in Reports. Figure 13 shows a sample of concordance lines in the sub-corpus. With respect to colligation, in the L1 or L2 positions, there is an auxiliary verb in 540 of the lines (85.6%). The auxiliary verb is realized as "was" (221 times, 40.9%), "be" (121 times, 22.2%), "were" (113 times, 20.9%), "been" (36 times, 6.7%), "are" (13 times, 2.4%), "is" (11 times, 2.0%), "being"

(10 times, 1.9%), "have" (8 times, 1.5%) and "has" (2 times, 0.4%). In four instances, the word "wasere" appears. It is assumed that it is miswriting of the word "was" or "were". The predominant occurrence of an auxiliary verb immediately in front of *carried out*, except for "have" and "has" indicates that the phrasal verb is used in passive voice. For those instances with "be" at the L1 position, a modal or semi-modal verb (64 times, 52.9%) and an infinitive marker (57 times, 47.1%) is found at the L2 position. In 87 instances (13.8%), a noun phrase comes in front of *carried out*. The phrasal verb is used in past tense form in 13 of these cases (14.9%), whereas in the other 74 instances (85.1%), *carried out* is used in the passive form functioning to modify its preceding noun phrase, such as those in lines 3 and 6 in Figure 13.

Regarding colligation to the right, it can be seen that *carried out* is mostly followed a prepositional phrase (457 times, 72.4%). In the other 50 instances (7.9%), *carried out* is the last item in the clause. There is a noun phrase following *carried out* in 48 other instances (7.6%). In the remaining lines, there are to-infinitive clauses (28 times, 4.4%), adverb (25 times, 4.0%), conjunction (21 times, 3.3%) and verb phrase (2 times, 0.3%). The most frequent configuration of *carried out* in Report is:

auxiliary verb + carried out + prepositional phrase

```
1    using non-oxidizing biocides. The  dosing is carried out over a period at intervals of several days,
2     study and assessment. Subsequent tests were carried out with power  quality improvement devices to
3   shape due to the harmonic current suppression carried out by the Active Harmonic Filter. RTCB only a)
4         layout and  services design, which are carried out by different professionals. These various
5  electrical scheme and installation is normally carried out using the  tender drawings. Flowchart of
6        the vast amount of construction projects carried out in Hong  Kong over the past 25 years,
7        in Hong Kong relies heavily on the works carried out by the biggest and most consistent
8  uality in swimming pools in various camps were carried out. Complaint of low water pressure in
9  ong Kong, most research works on BSE have been carried out in the four universities offering degree
10        The whole refuse collection process was carried out in a fully enclosed manner such that
```

Figure 13: Sample concordance lines of *carried out* in Reports

Top collocates to the left of *carried out*, top collocates are "the" (293 times), "monitoring" (260 times), "was" (244 times), "and" (166 times), "be" (159 times), "of" (139 times), "wastewater" (135 times), "were" (131 times), "to" (126 times), "no" (109 times), "site" (107 times), "waste" (93 times), "audit" (91 times), "month" (85 times), "works" (83 times), etc. Top collocates to the right include "the" (463 times), "at" (210 times), "in" (184 times), "of" (181 times), "and" (162 times), "by" (145 times), "this" (131 times), "on" (114 times), "month" (113 times) "to" (108 times), "a" (103 times), "monitoring" (99 times), "reporting" (88 times), "site" (71 times), "during" (64 times), etc. The semantic preference of *carried out* is 'engineering-related works' which is observed in every instance, for example, "the harmonic current suppression" in line 3, "construction projects" in line 6, "a technical study", "baseline marine water monitoring", "dredging activities", "audit checks" and "site inspections".

In this sub-corpus, the semantic prosody of 'requirement' is observed in 501 instances (79.4%). Although the percentage is high, it is noted that the semantic prosody is not as explicitly represented as in the instances in Ordinances. This is somehow reflected in the collocation because modal verbs such as "shall" and "should", or words or phrases such as "regulations", "in accordance with", "pursuant to" and "require" are not found among the top collocates listed above. However, most instances of *carried out* in Reports have a sense of 'requirement' because, for example in line 1 of Figure 13 above, it is highly likely that there are standard procedures and required steps for how "the dosing" (for cooling tower operation) is conducted. The phrase "over a period at intervals of several days" suggests that there is strict guidance to be followed, rather than conducting dosings randomly. Also, referring to wider co-texts shows that this instance of *carried out* occurs in the following extracts:

... The Prevention of Legionnaites' Disease Committee has published a Code of Practice and information pamphlet which recommended the good practices to be followed in the design, operation and maintenance of cooling towers for the effective control and prevention of the Legionnaites' Disease ...

... CIBSE recommends that microbiological control can be done by continuous or intermittent chemical addition or by non-chemical means ...

This further proves that this instance of *carried out* has the semantic prosody of 'requirement' as the carrying out of the process is guided by a code of practice and recommendations of CIBSE, the Chartered Institution of Building Services Engineers. Although CIBSE is not an authority with legal power, it promotes and recommends good practices in the building services engineering industry.

Another example is line 4, the co-text of *carried out* is extracted as follows:

A building project starts with inception and feasibility study. When the client decides to proceed, the design stage commences. The design team, usually led by the architect appointed by the client, is responsible to design a building which fulfills the requirements of the client set out in his brief. Typically, the design of a building includes architectural design, structural layout and services design, which are carried out by different professionals

This extract demonstrates that there are standard procedures for the early stages of working on a building project, from the initial and preliminary study examining the feasibility to the structural and services design. This shows how standardised and organised the engineering works are. And in this instance, *carried out* is used to express the people who perform "architectural design, structural layout and services design", i.e. "different professionals". Again, there may not be any co-occurring words or phrases which explicitly state the obligation for the works to be performed by "different

professionals". However, in this kind of context, the architectural and structural design of a building requires specific professional knowledge and skills, thus it is customary and required practice for works to be done by professionals or people with qualifications in the field.

In Figure 13, lines 1, 2, 4, 5 and 10 have the semantic prosody of 'requirement'. The other instances do not have this semantic prosody because there is no evidence indicating any obligation to conduct the works. For example, in line 6, "the vast amount of construction projects" and "in Hong Kong over the past 25 years" attempt to show that the development of building services engineering in Hong Kong was quick and impressive. The co-occurring words by no means convey a sense of obligation to conduct the construction projects. Thus, this instance of *carried out* does not share the semantic prosody of 'requirement'. In line 9, as another example, it is uncertain whether the "research works" conducted by "the four universities offering degree programmes in BSE and the Institute of Vocational Education" are required or not. While some research works are required in the sense that they are commissioned and directed by government bodies to study the feasibility, potential effects and possible solutions to new policies or proposed works, some are done voluntarily by the industry or universities to contribute to improvement and innovation in the field.

It is noted that a number of sub-corpora with a strong semantic prosody of 'requirement' have very low frequencies of *carried out*. For example, Fact Sheets and Questions and Answers both have 100% of the occurrence of the semantic prosody, but the total frequency of *carried out* in these two sub-corpora are ten and five respectively. In Tender Notices, there is only a single instance of the phrasal verb. Due to the limited

occurrences, the findings reported for these sub-corpora are suggestive rather than in any way definitive and need further investigation.

The other sub-corpora have relatively lower percentages of occurrences of the semantic prosody of 'requirement'. These are Speeches, Transaction Discussions, Abstracts, Transaction Proceedings, Transaction Notes, Position Documents, Media Releases, and Technical Papers. Speeches, Transaction Discussions, and Technical Papers have less than ten instances of *carried out*. For some of the sub-corpora, the low percentage of occurrence of the semantic prosody may be due to the limited frequencies of the phrasal verb. For example, 50% of the instances in Speeches and Transaction Discussions have the semantic prosody of 'requirement', however, there are only two and four instances respectively of the phrasal verb in these two sub-corpora. Such limited frequencies may result in a misrepresentation of the actual pattern of the use of *carried out* in the sub-corpora. For other sub-corpora, the low occurrence may be due to the nature of the texts in which the phrasal verb is used.

In Abstracts, there are 38 instances of *carried out* but only 31.6% have the semantic prosody of 'requirement'. Figure 14 shows a sample of concordance lines of *carried out* in this sub-corpus. In fact, the phrasal verb used in this sub-corpus share very similar co-selections to those described above for Ordinances and Reports, such as the dominant use of passive voice, the colligational pattern of *auxiliary verb + carried out + prepositional phrase*, and the strong semantic preference for 'engineering-related works'. With regard to the semantic prosody, however, only 12 instances share the sense of 'requirement'. For example, in line 1, the "inspection" is conducted for "manufacturing license", and the "Therapeutic Goods Administration" is a regulatory body in Australia

for therapeutic goods (Therapeutic Goods Administration, 2015). This suggests that *carried out* in this instance carries a sense of requirement due to the nature and purpose of the inspection. In line 4, the instance deals with the terms of the division of assets among members of a company as a result of the liquidation of a business. This kind of decision is regulated by some company ordinances or documented in articles of association of the company. Therefore, the carrying out of the "division" has the semantic prosody of 'requirement'. Lines 1, 2, 4 and 8 share this semantic prosody.

```
1         Therapeutic Goods Administration (TGA) carried out an inspection of our facility for GMP
2 October 2004 and a similar inspection was also carried out by the Department of Health, HKSAR in May
3         customers and local companies have been carried out since 2004  The ITF Project on Wireless
4   and may determine how such division shall be carried out as between the members or different
5    dies with high productivity. This study was carried out to investigate the aerated and tapped bulk
6   1990 to July 1994. Further analysis has been carried out to accertain the accuracy of prediction of
7      tests on the two types of aggregates are carried out to investigate the effects of using the
8   rise. This paper describes the pilot projects carried out by the Hong Kong Housing Department in the
9  STL of sandwich panel constructions have been carried out to investigate the dependence of the sound
10   problems of landfill in Hong Kong has been carried out. This paper summarises results of the
```

Figure 14: Sample concordance lines of *carried out* in Abstracts

Lines 3, 5, 6, 7, 9 and 10 are unclassifiable. For example, in line 3, the co-text is extracted as below:

… Successful commercial deployments with internationally well known consumer electronics customers and local companies have been carried out since 2004…

The carrying out of "deployments" tends to be companies' own decisions. It is not likely to be a requirement. In line 5, the phrasal verb is used in the proximity of some research or technical study, as suggested by the co-occurring words "this study" and "to investigate". As was noted earlier, research can be done voluntarily or commissioned by government bodies. It is undetermined in this case whether the study is required to be conducted or not. It is the same case for lines 6, 7, 9 and 10. This sub-corpus is a collection of abstracts which summarises the content of research papers or studies

conducted in different areas of engineering. Thus there are many instances similar to these ones which are not exactly classified as having the semantic prosody of 'requirement'.

The sub-corpus of Transaction Proceedings consists of full length research papers describing research projects or developments in the engineering industry. Thus, for the same reason that it is uncertain whether the research projects are required or not, only 41.7% (176 times) of the instances of *carried out* have the semantic prosody of 'requirement'. Figure 15 shows a sample of concordance lines of *carried out* in Transaction Proceedings. Only lines 5 and 8 have the semantic prosody of 'requirement'. The co-text of line 5 is extracted as follows:

All test powders were first conditioned in a room of 23 ℃ and 55% relative humidity for at least 24 hours. Blending was carried out by thoroughly mixing the two pre-weighed blending partners by a domestic blender before bulk densities were measured. Care was taken to ensure that the mixing blades of the blender were kept at low speed to avoid particle attrition.

The extract shows that there are strict procedures, control and care taken for the materials and conditions related to the carrying out of the work "blending". In line 8, *carried out* is used to express the compliance of "the bleeding test" with "ASTM Standard C232-71(5)". Thus, in Figure 15, these two instances of *carried out* have the semantic prosody of 'requirement'

```
1     study of Concrete pumping in UK and Germany carried out in 1986 (Anson et al, 1989), There were a
2  suggested. Such studies are more likely to be carried out by Research Institutes serving an industry
3     One of the early studies on the subject was carried out by McGeary 10, who investigated the
4 prescribed sizes and pro portions. Later, work carried out by Adler 1, Grey and Beddow 6, Hausner 8,
5   humidity for at least 24 hours. Blending was carried out by thoroughly mixing the two pre-weighed
6       Ratio". Wong 15 and Geldart et a1 5 have carried out extensive studies on gas fluidization of
7 ese racks and the weathering profiles has been carried out (Strange & Shaw, 1986) and much experience
8      106(4) respectively. The bleeding test was carried out in accordance with the ASTM Standard
9 an areas where huge housing projects are being carried out; thus, the present paper should provide
10 strength concrete mixes to 35% by volume were carried out at the University of Hong Kong [8,9]. Very
```

Figure 15: Sample concordance lines of *carried out* in Transaction Proceedings

It is noted that citations or references, are found in lines 1 to 4, and lines 6, 7 and 10. They come right after *carried out* and are mostly abbreviated alphanumeric expressions as in, for example, "Anson et al, 1989" in line 1 or citation numbers in square brackets, as in "[8, 9]" in line 10. The use of citations is a conventional practice in intellectual research to attribute ideas to the correct sources and acknowledge their contributions. Therefore, such use of *carried out* is considered as having the semantic preference of 'author and year of publication' and having a semantic prosody of 'acknowledging' in terms of when the earlier research contributions were made and who the contributors were.

In summary, *carried out* has fairly similar co-selections across the engineering genres. The phrasal verb is predominantly used in the passive voice, in the form of *auxiliary verb + carried out + prepositional phrase*. It has a very strong semantic preference of 'engineering-related works' because the co-occurring words mostly denote technical procedures of engineering works such as building and installation, testing and inspections, quality monitoring and safety measures, and research projects. These are all kinds of core works engineers deal with in the engineering workplace. Regarding the semantic prosody, *carried out* is frequently used to express the requirement of conducting the engineering-related works, particularly in terms of the way, the time and the environment of how the works are conducted, the people who are responsible for performing the works, and the standards to which the carrying out of the works complies with.

Regarding the non-contiguous form of *carried out*, it is surprising to find that there is only a single instance in the whole corpus. The only instance is found in

Ordinances which is the sub-corpus with the highest percentage of occurrence of *carried out*. The instance displays a constituency variation and is extracted below:

(2) Where a registered examiner or registered contractor carries out lift work in accordance with the relevant portions of codes of practice established under this Part, he shall be deemed to have carried them out to the satisfaction of the Director.

Immediately to the left of *carried…out*, the auxiliary verb "have" occurs. Therefore, this instance is used in the active voice. To the right of the instance, there is a prepositional phrase "to the satisfaction of the Director". The intervening word between *carried* and *out* is the pronoun "them" which refers to "lift work". This phrase confirms that the instance shares the same semantic preference of 'engineering-related works' despite the difference in form. And the instance is used in the textual environment with the occurrence of "codes of practice established under this Part", "shall be deemed to" and "to the satisfaction of the Director". These phrases indicate that this instance of *carried…out* conveys the sense of 'requirement'.

4.3.2 *Carry out*

The second most frequent form of CARRY OUT is *carry out* which is also among the top most commonly found phrasal verbs in the whole corpus. Among the 31 lists of the top ten most frequent phrasal verbs, *carry out* is on 11 sub-corpora's lists. Overall, there are 1,513 instances (0.016%) in the corpus and they are distributed across 27 sub-corpora. Ordinances is the sub-corpus with the highest percentage of occurrence of this phrasal

verb. There are 90 instances of *carry out* in Ordinances. A sample of concordance lines are shown in Figure 16.

```
1      a Registered Electrical Contractor Worker to carry out all electrical installations General
2 test and inspection. 1.22.5 The Contractor shall carry out tests as stated in the current appropriate
3          of the Engineer, the Contractor shall carry out the commissioning tests for the complete
4      a Registered Electrical Contractor Worker to carry out all electrical installations within the
5   by the Contract or ordered by the Engineer to carry out any examination, measurement or test, then
6   from the Engineer. 1.21.4 The Contractor shall carry out tests as stated in the current appropriate
7          of the Engineer, the Contractor shall carry out the commissioning tests for the complete
8  2.6 Testing The Contractor shall be required to carry out tests to demonstrate that the equipment
9          for in the Contract, the Contractor shall carry out the maintenance work provided for in
10          in the following sub-paragraphs :- (a) to carry out all repairs necessary to maintain the
```

Figure 16: Sample concordance lines of *carry out* in Ordinances

To the left, a modal verb (43 times, 47.8%) or an infinitive marker (41 times, 45.6%) is often found at the L1 or L2 position of *carry out*. To the right of the phrasal verb, there is always a noun phrase (87 times, 96.7%). In a few occasions, there is a prepositional phrase (3 times, 3.3%). Thus, the most frequent configurations of *carry out* in the sub-corpus are:

modal verb + carry out + noun phrase

infinitive marker + carry out + noun phrase

This colligational pattern shows that *carry out* is used in the active voice, and as an infinitive rather the simple present tense form. Only in a few instances is the phrasal verb used in the simple present tense form when it follows a relative pronoun (5 times, 5.6%).

In terms of collocation, top collocates to the left are "to" (62 times), "the" (58 times), "shall" (39 times), "of" (32 times), "a" (31 times), "in" (19 times), "contractor" (19 times), "person" (16 times), "and" (16 times), "registered" (15 times), "gas" (14 times), "who" (13 times), "by" (13 times), "personally" (12 times), "no" (11 times), etc. To the right, there are "the" (67 times), "work" (47 times), "to" (37 times), "in" (36 times),

"a" (36 times), "or" (33 times), "such" (23 times), "any" (23 times), "and" (20 times), "of" (19 times), "gas" (17 times), "installation" (17 times), "under" (11 times), "that" (11 times), "relation" (11 times), etc.

There is a strong semantic preference for 'engineering-related works' (88 times, 97.8%). As the same as *carried out*, *carry out* has this semantic preference due to the occurrence of words or phrases denoting the core tasks which engineers perform in their workplace, for example, "all electrical installations" as in lines 1 and 4, "tests" as in lines 2, 6 and 8, "the commissioning tests" as in lines 3 and 7, and "the maintenance work" in line 9 and "all repairs" in line 10. The phrasal verb has a very strong semantic prosody of 'requirement' (100%) in this sub-corpus. It is created, in some instances, by the use of "shall" which conveys a sense of directives and obligation, and "may". The modal verb *may* is sometimes used to express possibility and permission (Biber et al., 1999). Although *may* generally has a higher frequency of use expressing possibility than permission, it is concerned with permission in the current context. In other instances, phrases constituting this prosody include "required to", "referred to in paragraph (i)", "permit in writing a person other than a person referred to in subsection (1)", "required under this Ordinance", "on receipt of a notice under section 16", and "in contravention of section 14(1)", etc. Thus, *carry out* is used to express the meaning of 'requiring to perform some engineering-related works' in Ordinances.

This semantic prosody is not only found in Ordinances, but in most of the engineering genres. Table 11 shows the sub-corpora in which *carry out* is found and the percentage of occurrence of the semantic prosody of 'requirement' in each sub-corpus.

Table 11: Frequency and percentage of *carry out*

	Sub-corpora	Size	Freq.	%	Requirement	%
1	Ordinances	139,176	90	0.0647%	90	100.0%
2	Handbooks	67,284	31	0.0461%	13	41.9%
3	Publicity Material	599,407	264	0.0440%	235	89.0%
4	Notes	156,255	57	0.0365%	54	94.7%
5	Guides	783,805	272	0.0347%	257	94.5%
6	Plans	4,173	1	0.0240%	1	100.0%
7	Manuals	296,299	60	0.0202%	57	95.0%
8	Review Papers	106,506	20	0.0188%	19	95.0%
9	Frequently Asked Questions	55,726	10	0.0179%	9	90.0%
10	Code of Practice	997,228	166	0.0166%	163	98.2%
11	Reports	979,170	150	0.0153%	136	90.7%
12	Consultation Papers	111,494	17	0.0152%	15	88.2%
13	Standards	136,024	20	0.0147%	18	90.0%
14	Product Descriptions	611,549	77	0.0126%	60	77.9%
15	Technical Papers	65,731	8	0.0122%	8	100.0%
16	Fact Sheets	26,059	3	0.0115%	2	66.7%
17	Project Summaries	115,829	10	0.0086%	4	40.0%
18	Media Releases	1,566,742	129	0.0082%	74	57.4%
19	Q & A	27,703	2	0.0072%	2	100.0%
20	Conference Proceedings	196,498	13	0.0066%	10	76.9%
21	Abstract	94,671	6	0.0063%	5	83.3%
22	Transaction Notes (HKIE)	79,058	5	0.0063%	4	80.0%
23	About Us	647,013	39	0.0060%	26	66.7%
24	Transaction Proceedings (HKIE)	1,055,248	51	0.0048%	27	52.9%
25	Agreements	127,895	6	0.0047%	2	33.3%
26	Position Documents	75,660	3	0.0040%	2	66.7%
27	Circular Letters	143,313	3	0.0021%	3	100.0%

Of the 27 sub-corpora in which *carry out* is found, 19 of them have over 70% of the instances with the semantic prosody of 'requirement', and only in six of which, the phrasal verb has less than ten times of occurrence. For example, there is only a single instance in Plans, and eight instances in Technical Papers. Therefore, the 100% of the occurrence of the semantic prosody of 'requirement' in these two sub-corpora may be due to limited data.

In the remaining eight sub-corpora, the percentage of occurrence of the 'requirement' semantic prosody is lower. For example, in Handbooks which is the sub-corpus with the second highest percentage of occurrence of *carry out*, only 38.7% of the instances (12 times) share this semantic prosody. Figure 17 shows a sample of concordance lines from Handbooks. In this sub-corpus, there is only a single instance of the use of the modal verb "shall" in front of *carry out*. In the other instances, an infinitive marker is found at the L1 position (14 times, 45.2%), or *carry out* seems to be the first item in the clause (13 times, 41.9%). To the right, there is a noun phrase immediately following *carry out* in all instances.

```
1 Departments, all of which employ consultants to carry out varying portions of  their work.  The names
2  Where it is necessary for the consultants  to carry out additional Services, these should be
3 Contract.  The consultants must be  allowed to carry out the duties of the Engineer independently
4  (vi) Where Contract Advisers are intending to carry out a technical audit on  completed contracts,
5  given by the EACSB to appoint  consultants to carry out a specific assignment and as soon as the
6  not be employed under an existing Agreement to carry out  duties not authorised as part of that
7 to properly estimate the resources necessary to carry out the  Assignment and to quote competitive
8      in the supervision on site of works.(DMS) carry out field work investigation and to assist in
9  utilities or laying of new services  etc.(DMS) carry out field investigations on the works and to
10 fenced.(DMS)  perform routine site tests.(DMS) carry out investigation into public complaints.(DMS)
```

Figure 17: Sample concordance lines of *carry out* in Handbooks

Top collocates occurring on the left side of the phrasal verb are "to" (18 times), "the" (9 times), "of" (9 times), "consultants" (6 times), "and" (6 times), "registered" (3 times), "employ" (3 times), "electrical" (3 times), "contractors" (3 times), "underwater" (2 times), "structures" (2 times), "should" (2 times), "shall" (2 times), etc. To the right, there are "the" (16 times), "and" (15 times), "of" (12 times), "to" (11 times), "work" (9 times), "survey" (7 times), "field" (6 times), "as" (5 times), "a" (5 times), "surveys" (4 times), "staff" (3 times), "on" (3 times), "investigation" (3 times), "in" (3 times), "hydrographic" (3 times), etc. Many of these collocates and phrases co-occurring with *carry out* contribute to the strong semantic preference of 'engineering-related works' (100%), for example, "the duties of the Engineer" as in line 3, "a technical audit" as in

line 4, and "field work investigation" as in line 8. Even in some instances the words or phrases may seem general and ambiguous, such as "varying portions of their work" and "additional services" in lines 1 and 2 respectively, it is shown in the co-text that they actually refer to engineering-related works, for example, "work" refers to public works and "services" refers to consultants' services. 38.7% of the instances (12 times) carry the semantic prosody of 'requirement'. In Figure 17, *carry out* in lines 1 to 7 have the requiring semantic prosody. For example, in line 2, the phrase "where it is necessary" and the modal verb "should" indicate that there are restrictions and requirements about the carrying out of "additional services". In line 5, as another example, the appointment of consultants is under the permission of "the EACSB", which stands for Engineering & Associated Consultants Selection Board. This is an advisory board for approving the selection and appointment of engineering and associated consultants for government projects, so the board has authoritative power.

The phrasal verb in lines 8 to 10 looks different from other instances. In terms of colligation, these instances are those where *carry out* appears to be the first item in the clause. These instances are shown in the following extract from the sub-corpus:

RAIOW

May include duties, as appropriate, to :To include duties to:

assist in the supervision on site of works.(DMS)

carry out field work investigation and to assist in problem investigation.(DMS)

Under the supervision of a RE/RSIOW/RIOW, ensure that works are carried out

according to Specifications and drawings and identifying faulty materials. workmanship

etc. This involves liaising with the relevant contractors.(DMS)

This extract shows that the instance of *carry out* is listed under the phrase "To include duties to:" and it is one of the items in point form describing the duties of "RAIOW", which stands for Resident Assistant Inspector of Works. Thus, it expresses that carrying out field work investigation and assisting in problem investigation is one of the duties expected to be performed by "RAIOW". Since the phrasal verb is used in the context of describing job duties, as characterised by the occurrence of the phrase "to include duties to" and the use of *carry out* in a short phrase as one of the bullet items under the job title, it is considered as having the semantic prosody of 'describing job duties'. This semantic prosody is shared by the instances in lines 8 to 10 in Figure 17, and also other instances in the sub-corpus which do not have the semantic prosody of 'requirement'.

Regarding the non-contiguous form, there are only three instances in the corpus. They are found in Publicity Materials, Guides and Transaction Proceedings, and all display constituency variation in the form of *carry…out*. The three instances from the respective sub-corpora are extracted below:

Under the Buildings Ordinance, the carrying out of large-scale building works of a very simple nature (such as the erection of drying racks on the external walls of household apartments) are governed by the same set of controls, including the requirements to obtain prior approval and consent from the Building Authority (i.e. Buildings Department) before commencement of works and to appoint Authorized Persons(i.e., architects, engineers and surveyors registered under the Buildings Ordinance) and registered professionals to design and supervise the works, as well as registered contractors to carry them out.

Cleaning a grease trap is not a very pleasant job and staff members responsible for this task should be encouraged to carry it out promptly as required and thoroughly.

Without detracting from the necessity for reasonable and meaningful engineering calculations and from the rewards to those who can carry them out, at least equal professional prestige and responsibility should be accorded men of judgment even when that judgment is not expressed in numerical form

The first two instances are preceded by the infinitive marker "to". As for the third instance, *carry…out* is preceded by the modal verb "can". In the second instance, the phrasal verb is followed by an adverb. And in all the three cases, *carry* and *out* are intervened by a pronoun. Based on the co-texts, the intervening pronoun refers to "large-scale building works", "cleaning a grease trap" and "engineering calculations" respectively. These tasks are all core engineering activities. Therefore, these three non-contiguous instances share the semantic preference of 'engineering-related works'. The typical semantic prosody of 'requirement' is clearly observed in the first two instances. It is particularly evident in the occurrence of "Under the Buildings Ordinance", "governed by the same set of controls" and "the requirements to obtain prior approval and consent" in the first instance, and "responsible for", "should" and "promptly as required" in the second instance. These phrases create a strong obligatory sense about how the engineering works are to be performed. As for the third instance, the modal verb "can" preceding *carry…out* expresses ability, but the phrase "at least equal professional prestige and responsibility should be accorded men of judgment" expresses a recommended attitude towards engineering judgments and decisions. Thus, this instance seems to carry a sense of 'recommendation'. Overall, *carry…out* shares a very similar co-selection to the

contiguous form *carry out*. They are considered to be phraseological variations of the same phrasal verb.

4.3.3 *Carrying out*

Another form *carrying out* is found in 23 sub-corpora. Table 12 shows the sub-corpora in which *carrying out* is found and the percentage of occurrence of the semantic prosody of 'requirement'. It is shown that the semantic prosody is found in over 70% of the instances in 15 out of 23 sub-corpora.

Table 12: Sub-corpora in which *carrying out* is found

	Sub-corpora	Size	Freq	%	Requirement	%
1	Ordinances	139,176	64	0.0460%	62	96.9%
2	Code of Practice	997,228	204	0.0205%	179	87.7%
3	Frequently Asked Questions	55,726	11	0.0197%	10	90.9%
4	Guides	783,805	154	0.0196%	132	85.7%
5	Standards	136,024	25	0.0184%	25	100.0%
6	Publicity Material	599,407	104	0.0174%	81	77.9%
7	Notes	156,255	16	0.0102%	11	68.8%
8	Consultation Papers	111,494	11	0.0099%	8	72.7%
9	Handbooks	67,284	6	0.0089%	5	83.3%
10	Reports	979,170	82	0.0084%	72	87.8%
11	Conference Proceedings	196,498	16	0.0081%	11	68.8%
12	Review Papers	106,506	8	0.0075%	4	50.0%
13	Product Descriptions	611,549	43	0.0070%	32	74.4%
14	Media Releases	1,566,742	79	0.0050%	43	54.4%
15	Circular Letters	143,313	7	0.0049%	6	85.7%
16	Technical Papers	65,731	3	0.0046%	1	33.3%
17	Manuals	296,299	13	0.0044%	11	84.6%
18	About Us	647,013	25	0.0039%	13	52.0%
19	Fact Sheets	26,059	1	0.0038%	0	0.0%
20	Abstract	94,671	3	0.0032%	3	100.0%
21	Transaction Proceedings (HKIE)	1,055,248	26	0.0025%	8	30.8%
22	Project Summaries	115,829	2	0.0017%	2	100.0%
23	Agreements	127,895	1	0.0008%	1	100.0%

Ordinances, again, is the sub-corpus with the highest percentage of occurrence of this phrasal verb. There are 64 times of occurrence and a sample of concordance lines are shown in Figure 18. On the left side of *carrying out*, there are prepositions (20 times, 31.3%), determiners (17 times, 26.6%), nouns (15 times, 23.4%), conjunctions (4 times, 6.3%), verbs (4 times, 6.3%) and adjectives (2 times, 3.1%). To the right, there are mainly noun phrases (41 times, 64.1%) and prepositional phrases (21 times, 32.8%). Such a colligational pattern indicates that the phrasal verb *carrying out* is normally used as a gerund (63 times, 98.4%). Some instances function as noun phrases together with a subsequent noun phrase or prepositional phrase (37 times, 57.8%). For example, in line 1, *carrying out* forms a noun phrase with its preceding determiner "the" and its following prepositional phrase "of the Works" as "the carrying out of the Works", or in line 3, *carrying out* together with its preceding and following noun phrases to form a larger noun phrase, i.e. "description of methods of carrying out such commissioning tests". Some other instances function as adjectives modifying the preceding noun phrase, for example, in line 8, "carrying out such test" modifies "the persons".

```
1    supply of the necessary material, etc. for the carrying out of the Works and for the removal of all
2    the Secretary for Transport and Housing of the carrying out of work on the tramway; (Amended L.N.
3  detailed programme and description of methods of carrying out such commissioning tests for approval.
4  detailed programme and description of methods of carrying out such commissioning tests for approval.
5       cables. 5.1.5. The PABX shall be capable of carrying out routine background diagnostics
6    (ii) an addition to the installation; and (c) carrying out construction work in accordance with
7  with approved plans includes a reference to the carrying out of construction work in accordance with
8     have been taken to protect-  (a) the persons carrying out such test; (b) any persons working in
9    there is reasonable access for the purpose of carrying out maintenance on the installation; (c)
10  reasonably practicable, that-  (a) the persons carrying out such purging, any persons working in
```

Figure 18: Sample concordance lines of *carrying out* in Ordinances

In this sub-corpus, there is only a single instance in which *carrying out* is used in a verb phrase:

… the premises of any contractor or subcontractor who is carrying out or has carried out any work on the tramway…

It is a present participle used with its preceding auxiliary verb "is" to express the continuous aspect.

Top collocates to the left of *carrying out* in Ordinances are "the" (50 times), "of" (37 times), "and" (20 times), "in" (18 times), "or" (17 times), "a" (17 times), "for" (16 times), "person" (12 times), "registered" (11 times), "gas" (10 times), "any" (9 times), "to" (7 times), "examiner" (7 times), "personally" (6 times), "no" (6 times), etc. To the right, there are "the" (51 times), "of" (47 times), "work" (29 times), "and" (20 times), "in" (18 times), "a" (18 times), "gas" (17 times), "any" (17 times), "to" (12 times), "such" (12 times), "or" (11 times), "lift" (10 times), "installation" (10 times), "examination" (9 times), "is" (9 times), etc. The semantic preference of 'engineering-related works' (62 times, 96.9%) is also observed in the use of *carrying out*. This is suggested by phrases such as "work on the tramway" in line 2, "routine background diagnostics" in line 5, "construction work" in lines 6 and 7, and "maintenance on the installation" in line 9.

Throughout the concordance, a sense of 'requirement' is observed in the use of *carrying out*, for example, in line 1, for "the carrying out of the works" to be possible, the Contractor needs to "make all arrangements for the daily supply of the necessary material". In line 3, the Contractor needs to seek approval with respect to commissioning tests by notifying the Engineer three weeks in advance and furnish a detailed programme and description of methods of conducting the tests. In line 8, the protection for the people is a prerequisite of the performing of the pressure test. The phrasal verb *carrying out* has a strong semantic prosody of 'requirement' (62 times, 96.9%) in this sub-corpus.

The only two instances which do not share this semantic prosody are considered as having a sense of 'defining', as in:

… (2) In this section and section 19, "test and examination" means carrying out the following safety checks- …

The use of the phrasal verb is to provide a clarification of the particular meaning of certain terminology used within a certain context.

In Codes of Practice, which is the sub-corpus with the second highest percentage of occurrence of *carrying out*, there are 204 instances. Of which, 87.7% of them have the semantic prosody of 'requirement', whereas the other instances tend to share a sense other than requiring or regulating the performing of the engineering-related works. Figure 19 shows a sample of concordance lines of *carrying out* in the sub-corpus. From the perspective of colligation, *carrying out* in Codes of Practice is mainly preceded by a preposition (78 times, 38.2%), a determiner (49 times, 24.0%), a conjunction (32 times, 15.7%), and a noun phrase (25 times, 12.3%). And it is sometimes the first item in the clause (16 times, 7.8%), such as in line 10 in Figure 19. Following the phrasal verb, there is either a noun phrase (154 times, 75.5%) or a prepositional phrase (50 times, 24.5%). This colligation indicates that *carrying out* is used as a gerund functioning as a noun phrase in all the instances in this sub-corpus.

```
1   also on the accessibility of the steelwork for carrying out the maintenance. E5.5.2 Types of
2    be made over the full length of welds before  carrying out any further non destructive testing
3 of the skilled nature of the process, operators carrying out final ultrasonic examination of the weld
4     for preparing the method statement  and for carrying out the erection safely, it is good practice
5Section 16 of the Code gives recommendations for carrying out tests on  structures or components of
6    16.1.3 of the Code gives recommendations for carrying out tests and is  generally self explanatory.
7The following safety points should be noted when carrying out surveys and  inspections. They are not
8        for the preparation of Supervision Plans, carrying out  their respective supervision duties and
9  each stream; (d) the specific tasks of TCPs in carrying out site supervision; (e) the division of
10       for safety and the RC fails to rectify. ? Carrying out site inspections as necessary.
```

Figure 19: Sample concordance lines of *carrying out* in Codes of Practice

The top collocates to the left include "the" (167 times), "and" (66 times), "of" (56 times), "for" (52 times), "in" (46 times), "be" (43 times), "to" (39 times), "shall" (33 times), "or" (32 times), "by" (27 times), "a" (26 times), "than" (20 times), "is" (14 times), "other" (13 times), "when" (11 times), etc. To the right of *carrying out*, top collocates are "the" (171 times), "of" (96 times), "and" (65 times), "in" (51 times), "works" (45 times), "work" (34 times), "to" (33 times), "as" (30 times), "a" (28 times), "be" (25 times), "PCW" (24 times), "test" (23 times), "any" (23 times), "construction" (21 times), "shall" (20 times), etc . All the instances share the semantic preference of 'engineering-related works' as characterised by the occurrence of phrases such as "the maintenance" in line 1, "final ultrasonic examination" in line 3, "the erection" in line 4, and "tests on structures and components" in line 5. Similar to Ordinances, *carrying out* in Codes of Practice has a strong semantic prosody of 'requirement'. For example, in line 1, the carrying out of "the maintenance" requires "the accessibility of the steelwork" to be possible. In line 2, "visual inspection" of welds is a required procedure prior to performing "any further non-destructive testing". In line 3, only people who "hold a valid certificate of competence from a nationally recognised authority" are qualified for performing "final ultrasonic examination of the weld". Or in line 6, "the Code" provides guidance on the recommended practice for conducting tests. The co-occurring words and phrases of *carrying out* in lines 1 to 9 create the semantic prosody of 'requirement' (179 times, 87.7%). As for line 10, *carrying out* is one of those instances in which it is the first item in the clause. With the subsequent noun phrase, it is listed as one of the bullet items under "Responsibilities and Duties under AP's Stream". Although the original formatting has been removed, all these responsibilities and duties are presented in table form as suggested by the occurrence of "Table 4.1" in the text. Thus, this instance of *carrying out*,

rather than emphasizing and expressing the requirement, functions to describe engineering-related job duties. In the sub-corpus, 19 instances share these co-selections and the semantic prosody of 'describing job duties' (9.3%).

For non-contiguous instances, there are only two occurrences in the whole corpus. Both instances occur in the sub-corpus of Ordinances, and they are extracted below:

... he shall, before carrying it out, submit details of the proposed lift work to the Director and obtain the approval ...

...to have had sufficient experience or training in such work to be capable of carrying it out competently without supervision;

In both cases, the pronoun "it" is dropped between *carrying* and *out*. This pronoun refers to "lift work" and "builder's lift work or tower working platform work" respectively in the two instances. They share the typical semantic preference of 'engineering-related work', and in these cases particularly, the work is related to building lifts. The typical semantic prosody is observed in the first instance. Based on the wider co-text, *carrying…out* is used in describing the required procedures for "lift work that is not in accordance with the relevant portions of codes of practice", and the use of the modal verb "shall" also indicates that seeking official approval is a requisite step before the work can be performed. Therefore, there is a strong sense of 'requirement' in the first instance. As for the other instance, the co-texts show that this *carrying...out* is used in a defining paragraph for the specific term "competent worker". So, having adequate related experience or training and being able to perform the work skillfully is part of the characteristics of a "competent worker" in the particular context. This instance is considered as having the semantic prosody of 'defining'.

4.3.4 *Carries out*

The last inflectional form *carries out* has the lowest frequency in the corpus as compared to the other three forms of the lemma CARRY OUT. Overall, there are 99 occurrences in the whole corpus and these are distributed among 17 sub-corpora. Table 13 shows the sub-corpora in which *carries out* is found.

Table 13: Sub-corpora in which *carries out* occurs

	Sub-corpora	Size	Freq	%	Requirement	%
1	Ordinances	139,176	16	0.0115%	14	87.5%
2	Fact Sheets	26,059	2	0.0077%		
3	Notes	156,255	9	0.0058%	9	100%
4	Project Summaries	115,829	3	0.0026%	3	100%
5	Code of Practice	997,228	22	0.0022%	11	50.0%
6	Frequently Asked Questions	55,726	1	0.0018%	1	100.0%
7	About Us	647,013	11	0.0017%	9	81.8%
8	Manuals	296,299	4	0.0013%	4	100.0%
9	Publicity Material	599,407	8	0.0013%	5	62.5%
10	Product Descriptions	611,549	8	0.0013%	7	87.5%
11	Conference Proceedings	196,498	2	0.0010%	2	100.0%
12	Review Papers	106,506	1	0.0009%	1	100.0%
13	Standards	136,024	1	0.0007%	1	100.0%
14	Guides	783,805	4	0.0005%	4	100%
15	Media Releases	1,566,742	4	0.0003%	3	75%
16	Transaction Proceedings (HKIE)	1,055,248	2	0.0002%	1	50.0%
17	Reports	979,170	1	0.0001%	0	0.0%

As is the case for *carried out, carry out* and *carrying out*, Ordinances is the sub-corpus with the highest percentage of occurrence of *carries out*. There are 16 occurrences. Figure 20 displays a sample of concordance lines. On the left side of the phrasal verb, there is mostly a noun phrase (12 times, 75%). In some cases, there is a relative pronoun (3 times, 18.8%) at the L1 or L2 position. Since this phrasal verb is a third person singular,

the noun phrase preceding the phrasal verb and the referent of the relative pronoun are all singular nouns. To the right of *carries out*, there is always a noun phrase.

```
1         sound and gastight. (2) A person who carries out a pressure test required under subregulation
2   position of the gas pipe. (2) A person who carries out, or permits to be carried out, any works in
3   any source of ignition. (6) Where a person carries out work in relation to a gas fitting which
4  deposit. (Enacted 1990)   (1) Where a person carries out work in relation to an installation pipe
5   installation pipe is installed and a person carries out work in relation to such pipe, he shall
6   in proper working order. (2) Where a person carries out testing and examination specified in
7          means a person who as a business carries out gas installation work, whether or not-  (a)
8          means an individual who personally carries out gas installation work; "prescribed
9  installer, and-  (a) where the individual so carries out such work in relation to such plant, he
10 registered examiner or registered contractor carries out lift work in accordance with the relevant
```

Figure 20: Sample concordance lines of *carries out* in Ordinances

Regarding collocation, collocates to the left of *carries out* are "a" (12 times), "where" (8 times), "the" (8 times), "person" (7 times), "who" (4 times), "water" (3 times), "registered" (3 times), "authority" (3 times), "to" (2 times), "subsection" (2 times), "so" (2 times), "or" (2 times), etc. To the right, there are "work" (13 times), "in" (11 times), "the" (7 times), "a" (7 times), "to" (6 times), "relation" (5 times), "under" (4 times), "that" (4 times), "or" (4 times), "of" (4 times), "lift" (4 times), "gas" (4 times), "any" (4 times), "installation" (3 times), "shall" (3 times), etc. The semantic preference of 'engineering-related works' is found in all the instances, as suggested by co-occurring phrases such as "a pressure test" as in line 1, "work in relation to a gas fitting" as in line 3 and "lift work" in line 10.

In 87.5% of the instances (14 times), phrases denoting requirements or obligation are found, for example, "required under subregulation (1)" in line 1 or "in accordance with the relevant portions of codes of practice" in line 10. Such co-selection creates a semantic prosody of 'requirement'. It is noticed that 75% of the *carries out* (12 times) in this sub-corpus is used in a clause beginning with the word "where", such as lines 3 and 4 in Figure 20. The word "where" used in such context is to provide a possible case or a scenario. An example of this is found in line 9 of the extract below:

… (a) where the individual so carries out such work in relation to such plant, he shall be deemed not to have contravened regulation 3(1) thereby …

It introduces a case of "the individual" conducting "such work in relation to such plant", and if this case happens or is fulfilled, the individual "shall be deemed not to have contravened regulation 3(1)". Although *carries out* is not directly used in stating a requirement or regulation, it is used in setting a scenario or parameters in relation to the regulation of the performing of the works in some way.

In Frequently Asked Questions, Manuals, Conference Proceedings, Review Papers and Standards, all the instances of *carries out* have the semantic prosody of 'requirement'. However, the frequencies of *carries out* in these sub-corpora are generally low, ranging from one to four only. Therefore, the findings in these sub-corpora need to be hedged.

There are two instances of *carries out* in Fact Sheets. Due to its small size, a frequency of two makes it the sub-corpus with the second highest percentage of occurrence of *carries out*. However, these two instances draw attention because they appear at first sight to be problematic:

… Inspections are carries out by HyD staff in a daily cycle for safety inspections …

… all maintenance works are carries out under extremely tight working programmes …

The phrasal verb is preceded by an auxiliary verb "are", and followed by a prepositional phrase. This pattern of colligation is typical for *carried out* in engineering genres. However, it is considered ungrammatical to use *carries out* immediately after the auxiliary verb "are". These two instances are clearly errors and they are acceptable if they are replaced by *carried out*.

Notes is the sub-corpus with the third highest percentage of occurrence of the phrasal verb. There are 9 instances and they are shown in Figure 21. To the left of *carries out*, there is mostly a noun phrase (6 times, 66.7%). In the remaining cases, there is a pronoun (2 times, 22.2%) at the L1 or L2 position, or a conjunction (1 time, 11.1%). Following the phrasal verb, there is always a noun phrase (100%).

```
1  HKAS on matters relating to  accreditation and carries out audits of the PWL and term contract
2    The Landslip  Preventive Measures Division 1 carries out an overall diagnosis of the landslide
3  Special LPM Studies From time to time, the GEO carries out special LPM studies. These studies
4Maintenance Audit Unit under the Island Division carries out the maintenance audit work on behalf of
5        slopes. The Slope Maintenance Audit Unit carries out slope maintenance audits on these seven
6   of the Research and Development Division, HyD carries out audit inspections of all excavation sites
7        the railway development of Hong Kong. It carries out railway studies and is responsible for
8  construct waterworks projects.  The department carries out design and supervision of construction of
9    the quality of water supplied. It regularly carries out chemical, bacteriological, limnological,
```

Figure 21: Concordance lines of *carries out* in Notes

Collocates to the left are "the" (6 times), "division" (3 times), "it" (2 times), and "unit" (2 times), "to" (2 times), "time" (2 times), "maintenance" (2 times), "on" (2 times), "of" (2 times), "and" (2 times), etc. To the right, collocates are "of" (7 times), "and" (5 times), "the" (4 times), "maintenance" (3 times), "audit" (2 times), "audits" (2 times), and "studies" (2 times), "on" (2 times), "in" (2 times), etc. There are not many collocates due to the low overall frequency of the phrasal verb. The semantic preference of 'engineering-related works' (100%) is observed in this sub-corpus. This is created by the co-occurring words or phrases "PWL" which stands for the Public Works Laboratories, "diagnosis of the landslide data", "LPM" which stands for Landslip Preventive Measures, "the maintenance audit work", "slope maintenance", "audit inspections", "railway studies", "construction of waterworks project", and "chemical, bacteriological, limnological, biological and radiological examination of samples". These phrases denote different types of works which are done in the vicinity of the engineering industry. The noun phrases following *carries out* in every instance share a sense of inspecting and monitoring. For

example, "audits of the PWL" in line 1 denotes reviewing and inspecting the works of the Public Works Laboratories, "special LPM studies" in line 3 is the investigation and evaluation of the Landslip Preventive Measures, and "design and supervision of construction of waterworks projects" in line 8 refers to planning and directing the construction projects. It means that *carries out* used in these contexts is not expressing simply performing engineering-related works, but it is more specifically related to the monitoring and regulating of the works. Therefore, all the *carries out* in this sub-corpus carry the semantic prosody of 'requirement'.

In total, the semantic prosody of 'requirement' is found in over 70% of the instances of *carries out* in 12 out of the 17 sub-corpora (70.6%). In Publicity Materials, 62.5% (5 times) of the instances have this 'requirement' prosody. There are eight instances of *carries out* in the sub-corpus, which are shown in Figure 22. Lines 1, 2 and 4 are the instances which do not have the semantic prosody of 'requirement'. They share similar colligational features, with a noun or a relative pronoun preceding the phrasal verb and a noun phrase following it. However, they do not share the same semantic preference and semantic prosody with other instances.

```
1         services for carparks (P 06) Chevalier carries out caring program to the elderly residents
2   business units  of the Group. Chevalier also carries out trading activities throughout the world,
3   fice are described below: Land Formation CEO carries out planning and engineering feasibility
4    channels. Introduction Works contractor who carries out works, such as water works, drainage works,
5     the contractor and or person who actually carries out the works is responsible for ensuring where
6  he public. Who Should Register Any person who carries out gas installation work, for example, a
7        The contractor & or person who actually carries out the works is responsible for ensuring that
8  cted new regulations which require anyone who carries out construction work in the vicinity of a gas
```

Figure 22: Concordance lines of *carries out* in Publicity Materials

The phrasal verb co-occurs with "caring program" and "trading activities" in lines 1 and 2. They do not share the typical semantic preference of 'engineering-related works'. Although Chevalier is a construction and engineering company, the two co-occurring

phrases do not denote activities directly related to engineering-related works. Rather, "caring program" represent activities the company does as a socially responsible corporate, and "trading activities" are non-engineering-related activities but are parts of the business which are resulted from the company practicing diversification. These two instances are considered as not sharing the semantic prosody of 'requirement' because the carrying out of the "caring program" and "trading activities" in lines 1 and 2 are not about fulfilling any requirements or recommended standards, but are rather voluntary actions that contribute to society and diversify risks which the company faces. In line 4, the word "works" denotes different fundamental engineering activities, such as "water works" and "drainage works". Although sharing the typical semantic preference of 'engineering-related works', *carries out* in line 4 is used to discuss possible engineering activities performed by the contractor, and the possible consequences of the activities. No hints of requirement or obligation are found.

There is no occurrence of non-contiguous instances of *carries out* in the corpus.

4.3.5 Summary

To summarise, the co-selections of *carried out, carry out, carrying out* and *carries out* are rather similar. The most frequent one *carried out* is mostly used in the passive voice in the form of a past participle. It collocates with words or phrases that denote different types and areas of engineering activities, thus contributing a strong semantic preference of 'engineering-related works'. The second most frequent phrasal verb *carry out* is mostly used in infinitive form, preceded by an infinitive marker or modal verb. It also frequently

co-occurs with words or phrases sharing the semantic feature of 'engineering-related works'. For the form *carrying out*, its frequency of being used as a present participle in the continuous tense is surprisingly low. Instead, it is used predominantly as a gerund functioning as a modifier. The last phrasal verb *carries out* is rather infrequent in the engineering corpus. It is used in the simple present tense, thus the phrasal verb always comes after a singular noun phrase, pronoun or relative pronoun. Despite the difference in form and patterns of colligation, the four phrasal verbs share a strong semantic preference of 'engineering-related works'.

CARRY OUT is reported to have a semantic preference for performing an activity or a task in the BNC with frequent collocates such as "investigation, examination, survey, operation, research, duty, task, etc." (Chen, 2013). While findings in this study largely correspond with those in the BNC, the phrasal verbs are used more specifically in the contexts of performing a wide variety of engineering-related works in the HKEC. However, it is surprisingly found that most of the instances are not simply dealing with the fundamental tasks of engineering, such as building, installation and operating of machines, rather, CARRY OUT are often used to discuss the performing of monitoring and inspecting tasks such as auditing and testing, and planning and design. These tasks often involve taking references to or complying with legally binding regulations or recommended standards and practices which are particularly concerned with quality and safety assurance (Ho, Ahmed, Kwan & Ming, 2000). This creates the semantic prosody of 'requirement'. Such co-selection is found in most of the instances of all the four phrasal verbs, and across most of the sub-corpora. Minor semantic prosodies observed in the use of the four phrasal verbs include defining specific terminology, describing engineering-related job duties, and acknowledging the others' research work.

159

There are limited occurrences of *carried…out*, *carry…out* and *carrying…out*. The findings that a pronoun dropping in between the phrasal verb in all the non-contiguous instances conform to the description in Chapter 2 with regard to the syntactic flexibility of phrasal verbs. CARRY OUT is a Type I phrasal verb (combinations of a verb and an adverbial particle). Inversion of the particle and the direct object is allowed when the direct object is realized as a pronoun. It is also found that the non-contiguous instances share the typical meaning of the contiguous instances, i.e. to express the requirements in relation to the performing of engineering-related works. Even in a single instance in which this typical semantic prosody is absent, it is used to define a specific term used in the texts, and this is one of the minor semantic prosodies observed in the contiguous instances. Thus, the non-contiguous instances are simply phraseological variations of the same phrasal verbs. It is indicated that irrespective of the difference in inflected form and phraseological variation, the four phrasal verbs *carried out, carry out, carrying out* and *carries out* share the same meaning of performing required engineering-related works, irrespective of whether the requirement is legally-enforced, or self-imposed as an ethically good practitioner of engineering-related works.

It is commonly accepted that Type I phrasal verbs are used more frequently in spoken contexts and are relatively infrequent in written English. However, a number of studies show that CARRY OUT is an exceptional case. In the Longman Spoken and Written English Corpus, some commonly used Type I phrasal verbs were identified, and while most of them have much higher occurrences in the registers of conversation and fiction, CARRY OUT presents a different picture. It has an occurrence of over 100 times per million words in both the registers of news and academic prose, while showing less than 20 times per million words in conversation and fiction (Biber et al., 1999, p. 410). Indeed, CARRY

OUT is considered as academic vocabulary as it is more frequently found in the academic sub-corpora than in other sub-corpora of the BNC and COCA (Chen, 2013). The number of occurrence of CARRY OUT in the academic sub-corpus in the BNC and COCA is similar to that as found by Biber et al. (1999), Liu (2012, pp. 33-34) classified CARRY OUT as a Band 2 phrasal verbs, which means that there are 50-99 tokens per million words in academic writing. While these studies pointed out important findings as to what phrasal verb lemma are more commonly used in academic writing or other registers, the potential significance of particular inflected forms is neglected. Language users relying on the information may be misled. Thus, more specific information in respect to particular inflected forms is valued. In this regard, *carried out* is found to be among the most frequently used multi-word constructions in the abstracts of engineering journal articles (Nakano & Koyama, 2005). The data used by Nakano and Koyama, i.e. abstracts of engineering journal articles, has features shared by engineering texts and academic writing. From the findings of the current study, where *carried out* is the most frequently used form among other forms comprising the lemma and the second most frequent phrasal verb in the whole corpus of engineering texts, it can be concluded that *carried out* is highly specific to engineering genres in terms of frequency of use and expression of a particular meaning of performing engineering-related works as required.

Unlike BASE ON, two inflectional forms of CARRY OUT, *carried out* and *carry out*, are found among the lists of the top ten most frequent phrasal verbs in the sub-corpora. Although *based on* is the most frequent phrasal verb in the corpus, its inflectional counterparts are far different in terms of frequency. The facts that both *carried out* and *carry out* are frequent and that they share similar extended units of meaning reflect their specificity to the context of engineering writing.

Biber et al. (1999) pointed out that the particle *out* rarely has a literal meaning in the phrasal verb CARRY OUT. The findings in the current study conform to such description since none of the instances *of carried out, carry out, carrying out* and *carries out* is used with the literal spatial sense of moving something away from a particular place. Depending on contexts, people can regard CARRY OUT as a compositional construction or an idiomatic construction. The present findings, as well as those of previous studies, show that it is used primarily in its idiomatic sense, i.e. to perform or to do (Sinclair et al., 1989; Rundell & Fox, 2005). This can be attributed to the preference for a configuration with the two elements of the phrasal verbs being adjacent to each other (Biber et al., 1999; Gries & Stefanowitsch, 2004; Gilquin, 2015). Even in the very limited occurrences in which the direct object of the phrasal verb drops between the two constituents of the phrasal verb, the direct object is always realized as a pronoun. Semantic influence aside, the length of the object also seems to be a key factor for its position. It is observed that in the canonical configuration, the objects of the phrasal verbs are normally noun phrases which normally consist of more than one word, whereas in the very limited occurrences in which the direct object is placed between the two constituents of the phrasal verb, the direct object is realized by a single-word pronoun. This conforms to the principle of end-weight which suggests placing longer objects after the adverbial particle (Biber et al., 1999).

4.4 Use of phrasal verbs: COMPLY WITH

This section analyses and discusses the third most frequent phrasal verb in the corpus, *comply with*, and its other inflectional forms.

4.4.1 *Comply with*

Overall, there are 1,801 instances of *comply with* in the Hong Kong Engineering Corpus. They are distributed in 26 sub-corpora, in 16 of which, the phrasal verb occurs among the top ten most frequent phrasal verbs. Table 14 shows the sub-corpora in which *comply with* is found and the respective frequencies and percentage of occurrence.

Table 14: Frequencies and percentage of occurrence of *comply with* in the sub-corpora

	Sub-corpus	Size	Frequency of *comply with*	%
1	Ordinances	139,176	102	0.0733%
2	Publicity Material	599,407	286	0.0477%
3	Code of Practice	997,228	432	0.0433%
4	Frequently Asked Questions	55,726	22	0.0395%
5	Guides	783,805	304	0.0388%
6	Standards	136,024	52	0.0382%
7	Manuals	296,299	100	0.0337%
8	Handbooks	67,284	21	0.0312%
9	Consultation Papers	111,494	34	0.0305%
10	Review Papers	106,506	29	0.0272%
11	Q & A	27,703	7	0.0253%
12	Tender Notices	4,242	1	0.0236%
13	Notes	156,255	22	0.0141%
14	Product Descriptions	611,549	83	0.0136%
15	Transaction Notes (HKIE)	79,058	10	0.0126%
16	Reports	979,170	117	0.0119%
17	Conference Proceedings	196,498	21	0.0107%

18	Circular Letters	143,313	11	0.0077%
19	Technical Papers	65,731	5	0.0076%
20	Abstract	94,671	6	0.0063%
21	About Us	647,013	32	0.0049%
22	Project Summaries	115,829	5	0.0043%
23	Media Releases	1,566,742	66	0.0042%
24	Transaction Proceedings (HKIE)	1,055,248	29	0.0027%
25	Position Documents	75,660	2	0.0026%
26	Agreements	127,895	2	0.0016%

Unsurprisingly, the Ordinances sub-corpus recorded the highest percentage of occurrences of *comply with*. There are 102 instances of *comply with*, which accounts for about 0.0733% of the sub-corpus. Figure 23 below displays a sample of concordance lines of *comply with* in Ordinances. Some instances are used in the base infinitive form, some in to-infinitive form and some carry the present tense. These different uses of the phrasal verb are reflected in the colligation. To the left, there is a tendency for the phrasal verb to colligate with a modal verb (66 times, 64.7%). In most of these cases, the modal verb occurs at the N-1 position, or sometimes N-2 with an adverb found between the modal verb and the phrasal verb. There are also a number of instances when a longer phrase or clause is found between these elements, for example in line 10 in Figure 23, a short phrase "in all respects" is found between the modal verb and *comply with*. Apart from modal verbs, there may also be infinitive markers (34 times, 33.3%) to the left of the phrasal verb. Only in two instances (2.0%), was *comply with* used in present tense preceded by a noun phrase. On the right side of the phrasal verb, there is always a noun phrase. Therefore, the most frequent configuration of *comply with* in this sub-corpus is:

modal verb + comply with + noun phrase

164

```
1    1.7.4 All mechanical and electrical works shall comply with the technical details stated in the
2       1.7.6 All site work to be carried out shall comply with:- a) The Construction Site (Safety)
3          the Works and the execution thereof shall comply with all relevant Ordinances, Regulations or
4    in writing. All costs resulting from failure to comply with this requirement shall be borne by the
5    fitted. 1.13.10 Lubrication grease points shall comply with BS 1486-1:1959 Lubricating nipples.
6       and the armoured wires. 1.13.19 Motors shall comply with British Standards BS 4999-0:1987
7          1.13.20 Each starter for the motor shall comply with BS EN 60470:2001, IEC 60470:2000
8    of persons upon whom there is any obligation to comply with any requirements under this Ordinance,
9     are occasioned by failure of the Contractor to comply with the above testing and inspecting
10     the testing of the Works shall in all respects comply with the appropriate safety regulations.
```

Figure 23: Sample concordances of *comply with* in Ordinances

Collocates found to the left of the phrasal verb include "shall" (58 times, 56.9%), "to" (36 times, 35.3%), "the" (31 times, 30.4%), "and" (22 times, 21.6%), "of" (17 times, 16.7%), "fails" (12 times, 11.8%), "a" (10 times, 9.8%), "be" (9 time, 8.8%), "fittings" (9 times, 8.8%), "not" (8 times, 7.8%), "or" (8 times, 7.8%), "pipes" (8 times, 7.8%), "failure" (7 times, 6.9%), "in" (7 times, 6.9%), "for" (6 times, 5.9%), etc. To the right, there are "the" (68 times, 66.7%), "BS" (34 times, 33.3%), "for" (25 times, 24.5%), "of" (23 times, 22.5%), "and" (22 times, 21.6%), "requirements" (20 times, 19.6%), "a" (15 times, 14.7%), "requirement" (14 times, 13.7%), "or" (13 times, 12.7%), "under" (12 times, 11.8%), "in" (10 times, 9.8%), "part" (10 times, 9.8%), "any" (9 times, 8.8%), "regulations" (9 times, 8.8%), "this" (8 times, 7.8%), etc.

The semantic preference of 'requirements' is observed in all the 102 lines. Such preference is realized by the phrases such as 'all relevant Ordinances', 'British Standards BS 4999-0:1987', 'regulation 10', 'any instructions issued under regulation 12', 'the provisions of this Ordinance' and 'the requirements of IEC 60332'. Regarding the semantic prosody, *comply with* has a strong semantic prosody of 'obligation' (98 instances, 96.1%). The phrasal verb is frequently used with the modal verb "shall" which carries deontic force and words such as "obligation" and "require". The modal verb "shall" is employed to denote a sense of order or to issue authoritative instructions in more formal contexts (Sinclair, 1990; Carter & McCarthy, 2006). In some other instances, the use of the phrases "fails to", "failed to" or "failure to" with *comply with* describes a

scenario of non-compliance with requirements. A sense of 'obligation' is also observed in these instances as penalties or legal consequences may result from non-compliances; for example, in line 4 in Figure 23, Contractors bear the costs resulting from non-compliance with the requirement, or in line 19, the person not conforming to the requirement is regarded as committing "an offence and is liable to a fine of $5000 and imprisonment for 3 months". The co-selection of *comply with* and this semantic preference and semantic prosody expresses the obligation to conform to imposed requirements or standards. Ordinances set out the stipulated requirements to regulate the behaviour of parties or conditions of products related to engineering activities. It is crucial to emphasise the obligation to meet and conform to the requirements.

Apart from the 'obligatory' meaning, some instances of *comply with* (26 times, 25.5%) is associated with the phrases 'fails to', 'failure to' or 'failed to', contributing to a semantic preference of 'non-compliance', for example, line 4 in Figure 23. These lines have a semantic prosody of 'describing consequences' as suggested by the occurrence of phrases such as 'commits an offence', 'borne by', 'liable to a fine' and 'for hearing by a disciplinary tribunal'. They describe the consequences or penalties for people or activities which do not meet the stipulated requirements. Thus, in these 26 lines, the function of *comply with* finds its place in an extended unit of meaning which is 'describing consequences of non-compliance of stipulated requirements'.

Table 15: Meaning distribution of *comply with*

	Sub-corpus	size	Freq	%	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
1	Ordinances	139,176	102	0.0733%	62.7%	2.0%	25.5%	1.0%	3.9%	1.0%	1.0%	1.0%	1.0%	
2	Publicity Material	599,407	286	0.0477%	47.6%	5.6%	7.0%	12.9%	21.0%	1.0%	2.1%		1.4%	1.4%
3	Code of Practice	997,228	432	0.0433%	71.8%	7.4%	4.9%	3.7%	4.9%	4.2%	1.6%	0.5%	0.9%	0.9%
4	Frequently Asked Questions	55,726	22	0.0395%	22.7%	18.2%	13.6%	18.2%	0.9%		0.9%			
5	Guides	783,805	304	0.0388%	54.9%	4.6%	13.5%	6.9%	7.6%	4.3%	2.3%			5.6%
6	Standards	136,024	52	0.0382%	76.9%		5.8%	9.6%	1.9%	1.9%				3.8%
7	Manuals	296,299	100	0.0337%	76.0%	4%	7%	2.0%	3.0%	2.0%	4.0%	1.0%		
8	Handbooks	67,284	21	0.0312%	61.9%	9.5%		14.3%	4.8%	9.5%				
9	Consultation Papers	111,494	34	0.0305%	44.1%		17.6%	14.7%	11.8%			2.9%		5.9%
10	Review Papers	106,506	29	0.0272%	41.4%	3.4%	13.8%	10.3%	6.9%	6.9%	3.4%	6.9%		3.4%
11	Q & A	27,703	7	0.0253%	14.3%	14.3%	42.9%	14.3%	14.3%					
12	Tender Notices	4,242	1	0.0236%	100%									
13	Notes	156,255	22	0.0141%	54.5%	4.5%	18.2%	9.1%	9.1%					4.5%
14	Product Descriptions	611,549	83	0.0136%	48.2%	9.6%	13.3%	19.3%	6.0%	1.2%	1.2%		1.2%	
15	Transaction Notes (HKIE)	79,058	10	0.0126%	20%	50%		10%	10%	10%				
16	Reports	979,170	117	0.0119%	8.5%		0.9%	12.8%	6.8%	5.1%	0.9%	1.7%	0.9%	
17	Conference Proceedings	196,498	21	0.0107%	28.6%	19.0%	9.5%	9.5%	23.8%	4.8%				
18	Circular Letters	143,313	11	0.0077%	36.4%	18.2%	9.1%		27.3%			9.1%		
19	Technical Papers	65,731	5	0.0076%	60%			20%						20%
20	Abstract	94,671	6	0.0063%	16.7%	16.7%			50%	16.7%				
21	About Us	647,013	32	0.0049%	18.8%			6.3%	43.8%	3.10%			28.1%	
22	Project Summaries	115,829	5	0.0043%	40%				60%					
23	Media Releases	1,566,742	66	0.0042%	34.8%	7.6%	28.8%	12.1%	7.6%	6.1%	1.5%		1.5%	
24	Transaction Proceedings (HKIE)	1,055,248	29	0.0027%	17.2%	27.6%	3.4%	13.8%	13.8%	20.7%				
25	Position Documents	75,660	2	0.0026%				50%	50%					
26	Agreements	127,895	2	0.0016%		100%								

Note: M1=obligatory, M2=evaluating, M3=describing consequences, M4=making certain,
M5=intention/purpose, M6=characterising, M7=condition, M8=difficulty, M9=reassuring, M10=guidance

Table 15 shows the semantic prosodies found in the instances of *comply with* in each sub-corpus. Based on the table, *comply with* tends to be more specific to the semantic prosody of 'obligatory' as this meaning is found in every sub-corpus in which the phrasal verb is used. Particularly, it is associated with 60% or more of the instances in Ordinances, Codes of Practice, Standards, Manuals, Handbooks, Tender Notices and Technical Papers. The common feature of these sub-corpora, except for Tender Notices and Technical Papers, is that they are rather highly regulatory in nature in terms of regulating people's behaviour and guiding the operation of engineering activities. In Tender Notices and Technical Papers, there are only single occurrence and five instances of *comply with* respectively. In another six sub-corpora, there are about 40% to 60% of the instances associated with the meaning of 'obligatory'. They are Publicity Materials, Guides, Consultation Papers, Review Papers, Notes and Product Descriptions. Despite that these sub-corpora, except for Guides, do not function to be regulatory, it is reasonably important for the phrasal verb to be associated with 'obligatory' sense in engineering genres due to the safety concern. Therefore, the need and obligation to follow stipulated requirements and standards is emphasized to ensure quality and safety.

In fact, in all the sub-corpora, *comply with* has a strong semantic preference of 'requirements' as suggested by the co-occurring words or phrases such as 'requirements', 'standards', 'regulations', 'provisions of the code', 'conditions', etc. Thus, the fact that *comply with* has a high tendency to be used to talk about following different kinds of regulatory requirements explains why it tends to be specific to the semantic prosody of 'obligatory'.

It may be argued that *comply with* necessarily has the obligatory sense because it expresses acting in accordance with or meeting specified requirements. However, a different semantic prosody is found to be associated with most of the instances in the Reports sub-corpus. Figure 24 shows a sample of concordance lines from Reports. In this genre, there are 117 instances of *comply with*, constituting 0.0119% of the whole sub-corpus. In terms of colligation, the item coming in front of *comply with* is mostly an auxiliary verb followed by an adverb (41 times, 35.0%). In other cases, there are pronouns (22 times, 18.8%), infinitive markers (20 times, 17.1%), noun phrases (16 times, 13.7%), modal verbs (13 times, 11.1%), and relative pronouns (2 times, 1.7%). Based on such co-selection, it can be seen that *comply with* is used as a base infinitive (57 times, 48.7%), in present tense (39 times, 33.3%) and as a to-infinitive (21 times, 17.9%). On the other side of *comply with*, there is always a noun phrase. The most frequent configuration of *comply with* in the Reports sub-corpus is:

auxiliary verb + adverb + comply with + noun phrase

In fact, when this configuration occurs, the auxiliary verb is "does" and the adverb is "not". And the sequence is always *noun phrase + complies + does no + comply witht*, such as lines 3 to 8 in Figure 24. It is noted that ConcGram may not be able to display certain symbols in the concordance. In such cases, it is important to refer back to the original texts and it is found that the actual sequence in the text is *noun phrase + complies / does not + comply with*. The symbol '/' denotes a choice between 'complies with' and 'does not comply with'.

```
1  ENVIRONMENT QUALITY TARGETS STATUS Achieved  To comply with relevant environmental legislation with
2    in 2007, resulting in a higher DISR. 36 37 To comply with relevant environmental legislation with
3       The high volume sampler complies  does not comply With the specified requirements end is deemed
4  titration: 5 % The equipment complies  does not comply with the specified requirements and is deemed
5         % The salinity meter complies * does not comply with the specified requirements and is deemed
6  titration: 5 % The equipment complies  does not comply with the specified requirements and is deemed
7          % The salinity meter complies * does not comply with the specified requirements and is deemed
8       The high volition sample complies  does not comply with the specified requirements and in deemed
9   and percentile sound pressure level (Lx). They comply with International Electro technical
10  and percentile sound pressure level (Lx). They comply with International Electro technical
```

Figure 24: Sample concordance lines of *comply with* in Reports

Top collocates to the left of *comply with* include "complies" (41 times, 35.0%), "does" (41 times, 35.0%), "not" (41 times, 35.0%), "calibration" (27 times, 23.1%), "to" (27 times, 23.1%), "they" (22 times, 18.8%), "the" (20 times, 17.1%), "sampler" (19 times, 16.2%), "volume" (17 times, 14.5%), "that" (15 times, 12.8%), "and" (13 times, 11.1%), "sample" (11 times, 9.4%), "of" (10 times, 8.5%), "volition" (10 times, 8.5%), "contractors" (6 times, 5.1%), etc. To the right, top collocates include "the" (87 times, 74.4%), "requirements" (60 times, 51.3%), "acceptable" (41 times, 35.0%), "deemed" (41 times, 35.0%), "specified" (41 times, 35.0%), "unacceptable" (41 times, 35.0%), "and" (26 times, 22.2%), "is" (25 times, 21.4%), "in" (22 times, 18.8%), "international" (20 times, 17.1%), "end" (19 times, 16.2%), "commission" (18 times, 15.4%), "electro" (18 times, 15.4%), "publications" (18 times, 15.4%), "technical" (18 times, 15.4%), etc. There is a strong semantic preference of 'requirements stipulated by authorities' in 110 out of 112 lines (98.2%), for example, 'the specified requirements', 'the MEPS standard', 'International Electrotechnical Commission Publications', 'relevant environmental legislation', and 'minimum fresh air requirements'. Regarding the semantic prosody, 73 instances (62.4%) carry the meaning of 'evaluating'. This is because of the occurrence of the sequence *noun phrase + complies / does not*, which indicates something either 'complies with' or 'does not comply with' the stipulated requirements. Or in some instances, such as lines 9 and 10 in Figure 24, the pattern of a pronoun or a noun phrase

immediately preceding *comply with* indicates that the phrasal verb is used in the present tense and it also denotes the semantic prosody. Modality does not figure in these instances since they are simply stating the fact as to whether something meets the requirements or not. Rather than expressing the obligation to conform to requirements, *comply with* when possessing this semantic prosody carries the meaning of 'evaluating', whether requirements are met or not.

The main function of reports is to evaluate and present information. The Reports sub-corpus contains texts which report on inspections, monitoring, or calibrations of some aspects of company performance or engineering products, such as corporate sustainability, social and environmental impact, baseline monitoring of water quality, and assessments of different engineering products and equipment. There are statements reporting whether companies or engineering products meet the stipulated requirements or not.

The semantic prosody of 'obligation' is also observed in this genre, but only in 13 instances (11.1%). There is the use of modal verbs denoting deontic modality or the use of words such as "obligation". Occasionally, this semantic prosody is created by describing the consequences of non-compliance of the requirements. These instances share very similar co-selections to those described in the Ordinances sub-corpus.

For the meaning of 'describing consequences', Position Documents and Q & A sub-corpora have the highest percentages of occurrence of this semantic prosody. There are respectively 50% and 42.9% of the instances associated with the meaning of 'describing consequences' in the sub-corpora. But indeed, the actual frequency of *comply with* in Position Documents is two and that in Q & A is seven. Figure 25 shows the concordance lines of *comply with* in Q & A. The three instances having the semantic

prosody of 'describing consequences' are lines 1 to 3. They have the semantic preference of 'stipulated requirements and standards' as the noun phrases immediately following *comply with* denote different regulations or requirements. The collocation of these three instances with the phrase "fails to" and "shall be guilty of an offence" contributes to the semantic prosody of 'describing consequences'. The use of the phrasal verb in these instances is to inform what would happen for the non-compliance with the requirements.

```
1  ncillary purpose. (3) Any person who fails to comply with paragraph (1) shall be guilty of an
2    of that vehicle. (2) A person who fails to comply with a requirement under paragraph (1) shall be
3  ame of that person. (2) A person who fails to comply with a requirement to furnish information under
4 50 million in replacement of our dispensers to comply with the Vapour Recovery Phase 2 requirements.
5   construction and equipment provision need to comply with the BEC.  3. The choice of complying
6 help disadvantaged existing building owners to comply with these measures if necessary. 6. HKCA
7 e disinfection of the HATS effluent would also comply with the existing policy of providing
```

Figure 25: Concordance lines of *comply with* in Q & A

In some other sub-corpora, there are about 10% to 30% of the instances associated with this meaning. They are Ordinances, Frequently Asked Questions, Guides, Consultation Papers, Review Papers, Notes, Product Descriptions, and Media Releases. Although the percentage of occurrence of this meaning is not high across different sub-corpora, it is fairly common for *comply with* to be used with this meaning. In addition to emphasizing people's obligation to conform to the stipulated requirements, some of the texts provide information about the consequences of not meeting or following the requirements so as to create a deterrent effect.

Apart from the three meanings 'obliged to conform to stipulated requirements', 'describing consequences of non-compliance with requirements' and 'evaluating whether requirements are met', two other meanings are also associated with the use of *comply with* in most sub-corpora, namely 'making certain the requirements are satisfied' and 'intending to conform to the requirements'.

According to Table 14, 50% of the instances in Position Documents have the meaning of 'making certain the requirements are satisfied'. Nevertheless, there are only two instances in total in this sub-corpus. Product Descriptions has the second highest percentage of occurrence of this meaning. Figure 26 shows a sample of concordance lines from this sub-corpus.

```
1 strict quality control ensure that the  products comply with the highest standards of the industry.
2    Order for Removal of UBW If the owners do not comply with this Order, the Buildings Department
3   Built-in validation support functions help you comply with GLP GMP regulations.  A record of each
4 required for analytical instruments that have to comply with regulations such as GLP, OMP, and ISO.
5          criteria in order for the instrument to comply with FDA inspections.  In the case of
6   Built-in validation support functions help you comply with GLP GMP regulations. A record of each
7   Built-in validation support functions help you comply with GLP GMP regulations. A record of each
8   Built-in validation support functions help you comply with GLP GMP regulations. A record of each
9  Humidity Recorder built-in  Salt Spray Tester   Comply with CNS JIS ASTM standardsChamber temp: room
10However, the corresponding test report might not comply with each purchasing order and thus the
```

Figure 26: Sample concordance lines of *comply with* in Product Descriptions

In Product Descriptions, there are 83 instances of *comply with*. There is no dominant pattern of colligation. A modal verb, an infinitive marker or a noun phrase may occur in front of *comply with*. To the right, there is always a noun phrase. As the same as in other sub-corpora, there is a strong semantic preference of 'stipulated requirements and standards' (76 times, 91.6%) which is evident by the occurrence of words such as 'requirements', 'regulations', 'order', 'standards, etc. 40 instances of *comply with* in this sub-corpus (48.2%) are associated with the meaning of 'obliged to conform to the stipulated requirements' which was described earlier. In 16 instances (19.3%), *comply with* collocates with words such as "ensure", "help", "facilitate" and "check", as in the following:

...strict quality control ensure that the products comply with the highest standards of the industry...

...Built-in validation support functions help you comply with GLP GMP regulations...

…These solutions can facilitate financial institutions to comply with the AML guidelines and regulations….

…VPLD should be checked and approved to comply with Waterworks requirements…

The use of these words adds a sense of support and certainty. They express that something is done to support or make certain the compliance with requirements. Therefore, these instances have the semantic prosody of 'making certain'.

Regarding the meaning of 'intending to conform to the requirements', 40% to 60% of the instances in Abstracts, About Us and Project Summaries are associated with this meaning. In About Us, there are 32 instances of *comply with*. Figure 27 shows a sample of concordance lines.

```
1  Directors and Senior Management are obliged to comply with the standard of the Model Code set  out
2  operation.  CLP Respects Laws & Standards ? We comply with all applicable laws and regulations. ? We
3   recognises this possibility: "[a]nd worse, to comply with the emission caps stipulated in the SPLs
4  achieving emissions reductions: "And worse, to comply with the emission caps stipulated in the SPLs
5  business ethics. The Code  states that we must comply with all applicable legal requirements and
6  this Code. As a result, we must all review and comply with its provisions. We must also report
7      value. CLP Respects Laws and Standards We comply with all applicable laws and regulations. We
8  achieving emissions reductions: "And worse, to comply with the emission caps stipulated in the SPLs
9    who operate under an open access regime. To comply with the regime, market participants must
10     Social Responsibility Policy  Bribery  We comply with ICAC regulations of HKSAR, no matter the
```

Figure 27: Sample concordance lines of *comply with* in About Us

There is a tendency to find an infinitive marker at the N-1 position (20 times, 62.5%) of *comply with* in this sub-corpus. And to the right, there is always a noun phrase. Thus, the most frequent configuration is:

infinitive marker + comply with + noun phrase

The infinitive marker is 'to'. Other collocates to the left include "we" (7 times), "and" (7 times), "the" (6 times), "of" (6 times), "that" (5 times), "environmental" (5 times), "a" (5 times), "worse" (4 times), "management" (4 times), "has" (4 times), "control" (4 times),

"with" (3 times), "this" (3 times), "system" (3 times), etc. To the right of *comply with*, there are "the" (25 times), "and" (15 times), "of" (11 times), "requirements" (7 times), "regulations" (7 times), "in" (7 times), "standards" (6 times), "caps" (6 times), "to" (5 times), "emission" (5 times), "applicable" (5 times), "all" (5 times), "stipulated" (4 times), "environmental" (4 times), "individual" (4 times), etc. The semantic preference of 'requirements' (100%) is also observed in this sub-corpus. The semantic prosody of 'intention' is observed in 14 instances (43.8%). It is mainly suggested by the use of the word 'to' as in lines 3, 4, 8 and 9 in Figure X or in some cases the phrase 'in order to'. The use of the infinitive marker 'to' expresses an intention or a purpose. Thus, these instances of *comply with* together with its co-selections is used to express people or companies' intention or the purpose of their behaviour which is to conform to the requirements.

Although there is a small number of instances associated with other meanings, most instances of *comply with* are used with the five meanings discussed above. And in fact, in the first sixteen sub-corpora, except for Tender Notices, most instances of *comply with* are associated with the semantic prosody of 'obligatory'. Therefore, it suggests that the phrasal verb *comply with* is specific to engineering genres in expressing and emphasizing the obligation for people or products to conform to the stipulated requirements.

There are 34 non-contiguous instances of *comply with* across the sub-corpora. Some of them display the constituency variation as *comply…with*, some displays as the positional variation *with…comply*. Overall, the variation in its form does not have an influence on meanings. In Ordinances, for example, there are six non-contiguous

instances which are shown in Figure 28. All of them are associated with the semantic prosody of 'obligatory'.

```
1    decided by the Engineer. 1.1.3 The tender shall comply fully with the requirements stated in the
2   the Ordinance with which it is his obligation to comply, in particular in relation to the safety of
3   the Ordinance with which it is its obligation to comply, in particular in relation to the safety of
4   the Ordinance with which it is its obligation to comply. (3) Disciplinary action referred to in
5   the Ordinance with which it is his obligation to comply; or (b) a registered gas installer who it
6   the Ordinance with which it is his obligation to comply, in particular in relation to the safety of
```

Figure 28: Concordance lines of non-contiguous *comply with* in Ordinances

In line 1, the use of the modal verb 'shall' and the word 'requirements' suggest that it is obligatory to follow the requirements stated in the tender. Thus, it shares the semantic prosody of 'obligatory'. The intervening word between *comply* and *with* is an adverb 'fully'. It does not make the meaning of *comply…with* any different from *comply with* but it serves to emphasize the extent to which the requirements should be met.

In the other five instances, the phrasal verb displays a preposed construction as it is a Type II phrasal verb. The particle *with* together with a relative pronoun *which* is fronted. The relative pronoun refers to the noun phrase preceding the particle, i.e. "the Ordinance" in lines 2 to 5. In other words, the lines can be considered as '…it is his obligation to comply with the Ordinance…'. These five lines do not colligate with a modal verb. However, the collocation and semantic preference resemble those of the other lines as suggested by the associated noun phrase "the Ordinance". With respect to the semantic prosody, the use of 'obligation' creates the prosody of 'obligatory'. Therefore, although the phrasal verb in these five lines exhibit positional variation, they have essentially the same patterns of co-selection, and hence the same extended unit of meaning as the other lines.

Other non-contiguous instances in other sub-corpora are used with different meanings described above, such as 'describing consequences of non-compliance with requirements' and 'intending to conform to requirements', but the variation in its form, irrespective of whether it is *comply…with* or *with…comply*, does not create new meanings. Therefore, it can be argued that *comply with*, *comply…with* and *with…comply* are the phraseological variations of the same phrasal verb.

4.4.2 *Complying with*

Among the inflectional forms of COMPLY WITH, *complying with* is the form with the second highest number of overall frequencies across the sub-corpora. It is found in 23 sub-corpora. Table 16 shows the 23 sub-corpora in which *complying with* is found and the meanings that are associated with it.

There are eleven meanings which *complying with* may denote in engineering genres. Based on the table, it appears that *complying with* does not have a strong specificity to any one of the meanings, but there is a tendency for it to be used with the meaning of featuring. Review Papers is the sub-corpus with the highest percentage of occurrence of *complying with*. This sub-corpus is used for illustration.

Table 16: Sub-corpora in which complying with occurs

	Sub-corpus	size	Freq	% in the sub-corpora	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11
1	Review Papers	106,506	18	0.0169%				16.7%		55.6%		16.7%	11.1%		
2	Ordinances	139,176	17	0.0122%		11.8%				58.8%	23.5%				
3	Q & A	27,703	3	0.0108%								100%			
4	Code of Practice	997,228	92	0.0092%	3.1%	4.2%	1.0%		4.2%	83.3%	1.0%	1.0%		2.1%	
5	Manuals	296,299	27	0.0091%			3.7%	3.7%	7.4%	85.2%					
6	Guides	783,805	69	0.0088%	1.4%	1.4%		7.2%	4.3%	56.5%	4.3%	1.4%	2.9%	7.2%	2.9%
7	Handbooks	67,284	5	0.0074%			20%	20%		60%					
8	Consultation Papers	111,494	7	0.0063%					14.3%	28.6%		14.3%		28.6%	14.3%
9	Notes	156,255	9	0.0058%						100%					
10	Frequently Asked Questions	55,726	3	0.0054%				33.3%		66.7%					
11	Standards	136,024	6	0.0044%	16.7%		16.7%			50%			16.7%		
12	Publicity Material	599,407	26	0.0043%				7.7%	7.7%	76.9%		3.8%	3.8%		
13	Fact Sheets	26,059	1	0.0038%				100%							
14	Product Descriptions	611,549	19	0.0031%		5.3%	10.5%			73.7%			10.5%		
15	Position Documents	75,660	2	0.0026%			100%								
16	Conference Proceedings	196,498	5	0.0025%		20%				60%			20%		
17	Abstract	94,671	2	0.0021%		50%						50%			
18	Reports	979,170	16	0.0016%	6.3%					31.3%		6.3%	25%		31.3%
19	Transaction Proceedings (HKIE)#	1,055,248	17	0.0016%		17.6%				58.8%		11.8%			5.9%
20	Circular Letters	143,313	2	0.0014%						100%					
21	About Us	647,013	9	0.0014%		11.1%				11.1%		22.2%	55.6%		
22	Media Releases	1,566,742	15	0.0010%				6.7%	13.3%	73.3%					6.7%
23	Project Summaries	115,829	1	0.0009%		100%									

Note: M1=obligatory, M2=evaluating, M3=describing consequences, M4=making certain,
M5=intention/purpose, M6=characterising, M7=condition, M8=difficulty, M9=reassuring, M10=guidance,
M11=achieving

```
1        models that have been type-approved to be complying with the qualifying standards for
2  The Government will give priority to vehicles complying with the  specified qualifying standards
3  The trade has raised no major difficulties in complying with such a requirement. Nonetheless, in
4        which some trades would encounter in complying with the control on idling vehicles, other
5    a total of 6 924 public transport vehicles complying with the National III emission standard in
6  005. A recommended catalogue of motor vehicles complying with the National III emission standard has
7  2007. Shenzhen - A catalogue of motor vehicles complying with the National III emission standard has
8   Starting from 1 January 2007, motor vehicles complying with the National III emission standard will
9          promote the sale of motor vehicle fuel complying with National III standard in the province.
10 ecessary assistance to the affected trades in complying with the Scheme.  8. The trades and
```

Figure 29: Sample concordance lines of *complying with* in Review Papers

Figure 29 shows the sample concordance lines of *complying with* in Review Papers. To the left of *complying with*, the pattern of colligation is not strong. There is either a noun phrase (10 times, 55.6%) or a preposition (7 times, 38.9%) at the N-1 position. There is also a single instance of the verb *to be*. To the right, there is always a noun phrase immediately following *complying with*. Collocational patterns on both sides are also not particularly strong. To the left of *complying with*, there are 'vehicles' (6 times, 33.3%), 'fuel' (4 times, 22.2%), 'trades' (3 times, 16.7%), 'assistance' (2 times, 11.1%), and 'difficulties' (2 times, 11.1%). To the right, collocates are 'standard' (8 times, 44.4%), 'standards' (3 times, 16.7%), and 'requirement' (2 times, 11.1%).

Although there are no strong collocates, the collocates to the right of *complying with* share the semantic feature of regulatory criteria and requirements. Thus, the semantic preference is requirements. Regarding the semantic prosody, there is not a strong one in this sub-corpus. In ten of the lines (55.6%), the semantic prosody is 'characterising and setting parameters'. These lines are mostly characterised by the pattern of having a noun phrase at the N-1 position, such as line 2 and lines 5 to 9. The phrasal verb *complying with* is used to provide specification of the noun phrase it follows. For example, in line 2, the noun phrase at the N-1 position is 'vehicles', but instead of talking about simply all vehicles, the phrasal verb is used to limit the kinds of vehicles to which "the Government will give priority", i.e. only those vehicles which comply with "the specified qualifying

standards". Thus, there is a sense of 'characterising and setting parameters' in the use of *complying with* in these instances.

The remaining lines in this sub-corpus do not share this semantic prosody, but they have other meanings. However, due to the low frequencies, other meanings will be discussed and illustrated using other sub-corpora as example. As shown in Table 15, the semantic prosody of 'characterising and setting parameters' is found dominant in, or even specific to, some sub-corpora, such as Codes of Practice (83.3%), Manuals (85.2%), Notes (100%), Publicity Materials (76.9%), Product Descriptions (73.3%), Circular Letters (100%) and Media Releases (73.3%).

Q & A ranks the third in the table, but the actual frequency is only three. All these three lines have very neat pattern. The three lines are extracted below:

operators of the following transport trades may have difficulty complying with the ban-

As for RMB operators having difficulty complying with the ban at RMB stands …

Regarding the concern of tourist coach operators that tourist coaches with sealed windows will have problem complying with the ban …

The colligation is the occurrence of a noun phrase immediately on both sides of *complying with*, so the apparent configuration of *complying with* in this sub-corpus is:

noun phrase + complying with + noun phrase

Collocates to the left of *complying with* are 'difficulty' (2 times, 66.7%) and 'have' (2 times, 66.7%), to the right are 'the ban' (3 times, 100%). Because of the consistent use of

'the ban', the semantic preference of *complying with* in this sub-corpus is 'prohibition' (100%).

In terms of the semantic prosody, a sense of 'difficulty' is found (100%). This prosody is suggested by the nouns immediately to the left of the phrasal verb, i.e. 'difficulty' and 'problem'. Looking more closely, we can see the sequence right before *complying with* is a human operator followed by "have or having" and then the noun "difficulty" or "problem". Therefore, in this sub-corpus, *complying with* is used to discuss the difficulty of a human operator in conforming to the ban.

This semantic prosody is also observed in other sub-corpora, such as Review Papers (16.7%), Consultation Papers (14.3%) and About Us (22.2%). However, the percentage of occurrence of this semantic prosody is rather low. This indicates that the use of *complying with* in relation to the discussion of difficulty in conforming to requirements is not as common across the sub-corpora.

Regarding non-contiguous forms, two instances are found. They are displayed in constituency variation, i.e. *complying...with*. One instance is from Review Papers and the other is from Codes of Practice:

After carefully considering its request to incentivise also private cars complying only with the emission requirement…
This requirement does not apply to motors incorporated in an item of current using equipment complying as a whole with an appropriate recognised standard.

In these two cases, a noun phrase immediately occurs to the left of *complying* and to the right of *with*. In between *complying* and *with*, there is an adverb in both instances. Thus, the configuration of *complying…with* is:

noun phrase + complying + adverb + with + noun phrase

The noun phrases directly following the word *with* are 'the emission requirement' and 'an appropriate recognized standard' respectively. The occurrence of 'requirement' and 'standard' suggests that *complying…with* has the same semantic preference as most *complying with* instances, that is 'requirements'. The pattern of a noun phrase immediately occurring to the left of *complying* suggests that the semantic prosody is 'characterising and setting parameters'. This means that *complying…with* is used to provide extra information so as to characterise what items is exactly being mentioned. For example in the first instance, rather than any 'private cars', the kinds of 'private cars' being discussed in the text are those which meet the emission requirement.

The intervening words, i.e. 'only' in the first instance and 'as a whole' in the second instance, have no influence on the meaning of the phrasal verb, yet they provide a parameter as to the compliance with the requirements. For example in the first instance, the 'private cars' which are being discussed are those which conform to solely the emission requirements but not the other requirements else.

There is no distinction in the meaning of *complying with* and *complying…with*. Thus, the two forms are phraseological variations of the same phrasal verb. And the findings suggest that this phrasal verb, regardless of its form, is specific to the function of 'characterising and setting parameters'.

4.4.3 *Complied with*

The form *complied with* is found in 19 engineering sub-corpora. Table 17 shows the sub-corpora in which *complied with* occurs. It can be seen that the number of meanings *complied with* has in engineering genres is less than that of *comply with* and *complying with*. In most of the sub-corpora in which *complied with* occurs, the phrasal verb has the meaning of 'obligatory', 'making certain' and 'introducing a condition'.

Tender Notices is the sub-corpus with the highest percentage of occurrence of *complied with*. However, the actual frequency is in fact only two times. The concordance lines of these two instances are extracted and given below:

… the Tenderer and its nominated staff have fully complied with the mandatory requirements.

… the Tenderer and its nominated staff have fully complied with those mandatory requirements,

To the left of *complied with*, there is a noun phrase followed by an auxiliary verb "have" and an adverb. To the right, there is a noun phrase. The words occurring on both sides are the same for the two lines. The phrasal verb as the past participle and is used in the present perfect tense. They appear to express the meaning of 'evaluating'. But inspection on the wider co-text shows that *complied with* tends to be expressing the meaning of 'making certain' in line 1 and 'describing consequences' in line 2.

The completed version of the first line is extracted from the sub-corpus:

The Tenderer shall include all the relevant document proof in the Technical Proposal to demonstrate that the Tenderer and its nominated staff have fully complied with the mandatory requirements.

The words 'demonstrate' and 'proof' carries a sense of giving proof or evidence of something to show the existence or truth of a statement. Therefore *complied with* in this line is about making certain "the Tenderer and its nominated staff" have followed the requirements.

The other line is also extracted as follows:

Any Tenderer who fails to demonstrate to the satisfaction of the Tender Assessment Panel that the Tenderer and its nominated staff have fully complied with those mandatory requirements, its tender will not be considered.

In this instance, the word 'demonstrate' also appears, however, it follows the word 'fails'. So unlike the first line in which *complied with* is used to express how "the Tenderer" needs to have evidence to prove that the requirements are met, *complied with* in this line is used to talk about the consequences of failing to follow the requirements.

Table 17: -corpora in which *complied with* is found and meaning distribution

	Sub-corpus	size	Freq	%	M1	M2	M3	M4	M6	M7	M10
1	Tender Notices	4,242	2	0.0471%			50%	50%			
2	Codes of Practice	997,228	92	0.0092%	32.6%	2.2%		43.5%		10.9%	7.6%
3	Ordinances	139,176	12	0.0086%	25%			16.7%	16.7%	41.7%	
4	Guides	783,805	59	0.0075%	22.0%			47.5%	3.4%	22.0%	1.7%
5	Manuals	296,299	18	0.0061%	27.8%			16.7%	5.6%	44.4%	
6	Reports	979,170	55	0.0056%	25.5%	69.1%		3.6%			
7	Frequently Asked Questions	55,726	3	0.0054%	33.3%			33.3%		33.3%	
8	Notes	156,255	6	0.0038%				66.7%	16.7%	16.7%	
9	Publicity Material	599,407	19	0.0032%	36.8%	15.8%		31.6%	10.5%	5.3%	
10	Abstract	94,671	3	0.0032%	33.3%					66.7%	
11	Handbooks	67,284	2	0.0030%						100%	
12	Standards	136,024	4	0.0029%						75%	
13	Consultation Papers	111,494	3	0.0027%	66.7%				33.3%		
14	Media Releases	1,566,742	33	0.0021%		69.7%		18.2%	6.1%	3.0%	3.0%
15	About Us	647,013	13	0.0020%		53.8%		15.4%		30.8%	
16	Review Papers	106,506	2	0.0019%					50%	50%	
17	Product Descriptions	611,549	11	0.0018%	9.1%	27.3%		36.4%	18.2%	9.1%	
18	Transaction Proceedings (HKIE)	1,055,248	9	0.0009%	22.2%	33.3%		11.1%	11.1%		11.1%
19	Conference Proceedings	196,498	1	0.0005%		100%					

Note: M1=obligatory, M2=evaluating, M3=describing consequences, M4=making certain,
M5=intention/purpose, M6=characterising, M7=condition, M10=guidance

The sub-corpus with the second highest percentage of occurrence is Codes of Practice. In this sub-corpus, there are 92 instances of *complied with*, yet it does not have a

particular strong semantic prosody. A number of semantic prosody is observed and none of which has an occurrence of over half of the total frequency in the sub-corpus. Figure 30 shows a sample of the concordance lines of *complied with* in Codes of Practice.

```
1 (Planning) Regulation 41C and this Code are also complied with. 7. Firefighting and Rescue Stairway a
2   Code of Practice for Means of Escape are also complied with.  Part IV - FIREMAN'S LIFT 8.
3     wall of a lobby to a required staircase are complied with. 13.3 Where only one or some of the
4 10m such that the requirement in note 1 above is complied with, i.e. all the acute angles £\ between
5  has been achieved and the capacity of each pile complied with approved plans. F6 (vi) Proof test
6         in Section 8 of this Code shall also be complied with during critical stages of work. 2) If
7 of this Supervision Plan is  prepared by me and complied with the Technical Memorandum, the Code of
8 of this  Supervision Plan is prepared by me and complied with the Technical Memorandum, the Code of
9  of this  Supervision Plan is prepared by me and complied with the Technical Memorandum, the Code of
10  of this Supervision Plan is prepared by me and complied with  the Technical Memorandum, the Code of
```

Figure 30: Sample concordance lines of *complied with* in Codes of Practice

Regarding the colligation, an auxiliary verb is found at either N-1 or N-2 position in 88.0% of the lines (81 times). Preceding the auxiliary verb, there is mostly either a noun phrase or a noun phrase followed by a modal verb. To the right of *complied with*, a punctuation mark signaling the end of the clause is found in 69.6% of the instances (64 times). In 19.6% of the lines (18 times), there is a noun phrase. The most frequent configurations of *complied with* in this sub-corpus are therefore:

noun phrase + auxiliary verb + complied with

noun phrase + modal verb + auxiliary verb + complied with

Most of the use of the auxiliary verb is realized by the verb *be*, and they appear in the form of 'be' (36 times), 'are' (19 times), 'being' (11 times), 'is' (5 times) and 'been' (4 times). This indicates that *complied with* in these lines is used as the passive structure. Other collocates to the left are 'shall' (17 times), 'requirements' (13 times), 'should' (10 times), 'ordinance' (7 times), 'conditions' (6 times), 'code' (5 times), etc. To the right, there are 'code' (7 times), 'requirements' (7 times), 'technical' (6 times), 'memorandum' (4 times), 'practice' (4 times), etc. No collocate is particularly strong. The occurrence of

words such as 'requirements', 'ordinance', 'conditions' and 'code' suggests the semantic preference of 'requirements' (95.7%, 88 times).

There is no strong semantic prosody in the use of *complied with* in this sub-corpus. In almost half of the instances (43.5%, 40 times), the sense of 'making certain' is observed. This is suggested by the use of words such as 'ensure', 'check', 'confirm', etc. For example,

...to confirm that the requirement of chemical composition and mechanical properties appropriate to the type of steel have been complied with...

The occurrence of these words indicates that something is done to prove or make it certain that the requirements are followed. Another semantic prosody 'obligatory' is associated with 29 instances (32.6%). This is mainly suggested by the use of modal verbs 'shall' and 'should', such as line 6 in Figure 30. These two modal verbs have the modality sense of signaling obligation. In about one-tenth of the lines (10.9%, 10 lines), the word 'if' or 'unless', or the phrase 'provided that' are found. These words, together with complied with, are used to introduce a condition of compliance with requirements. These lines, therefore, have the semantic prosody of 'describing the effect on the condition of compliance with requirements', For example,

Openings for lighting and/or ventilation may be made in an external wall enclosing the lobby provided that the same requirements as specified in the Code of Practice for Fire Resisting Construction for openings in the external wall of a lobby to a required staircase are complied with.

This instance is about what happens, i.e. 'Openings for lighting and/or ventilation may be made in an external wall enclosing the lobby', on the condition that 'the same requirements as specified in the Code of Practice for Fire Resisting Construction for openings in the external wall of a lobby to a required staircase are complied with'. Another semantic prosody found is 'providing guidance' (7.6%, 7 times). This means that the instances are related to how the requirements can be satisfied. These instances are mainly characterised by the occurrence of words such as 'guidance', as in

This Code of Practice provides full guidance on how the requirements of the Buildings Ordinance may be complied with.

In some instances, the word 'guidance' is not found, but other patterns, such as *considered as + complied with + if*, as in the following:

The requirement will be considered as having complied with if the total fan motor power has to exceed 1.6 W per L/s of supply air quantity in order to…

Alternatively, the pattern of *can be + complied with + by* also constitute to the same semantic prosody, as in the following:

The storage area requirements stipulated above can be complied with by the provision of a cupboard or cabinet of a robust design for small chemical waste producers

The use of these patterns is to illustrate how or by what means the requirements can be conformed to. Thus, rather than emphasizing how important it is to meet the requirements, the use of *complied with* in these instances is to provide guidance on how to meet the requirements. Apart from these semantic prosodies, the prosody of 'evaluating' is also observed in this sub-corpus, but the frequency is very low (2.2%, 2 times). The use of

complied with in these two lines reports on whether the requirements are conformed to or not.

As shown in Table 16, the use of *complied with* is quite commonly found in expressing the meaning of 'obligatory', 'making certain', 'characterising', and 'introducing a condition', however, the percentage of occurrence distributes across these meanings. Even in one or two cases in which there is a 100% occurrence of a particular semantic prosody, the actual frequency of *complied with* in the sub-corpus is only one or two times, such as Handbooks and Conference Proceedings. Therefore, *complied with* in engineering sub-corpora is not particularly specific to one of the meanings.

No non-contiguous instance is found.

4.4.4 *Complies with*

Among the four inflectional forms of COMPLY WITH, *complies with* has the lowest total frequency in the whole engineering corpus, and it is also found in the least number of sub-corpora. Table 18 shows the sub-corpora in which *complies with* occurs. Unlike *complying with* and *complied with* that may be used to express seven or more meanings, there are only four meanings which are found to be associated with the use of *complies with* in engineering sub-corpora.

Table 18: Sub-corpora in which *complies with* is found

	Sub-corpus	size	Freq	%	M2	M4	M6	M7
1	Handbooks	67,284	9	0.0134%	22.2%	44.4%	33.3%	
2	Ordinances	139,176	15	0.0108%	13.3%	13.3%	20%	53.3%
3	Frequently Asked Questions	55,726	6	0.0108%		66.7%		33.3%
4	Publicity Material	599,407	47	0.0078%	19.1%	44.7%	23.4%	12.8%
5	Guides	783,805	39	0.0050%	12.8%	20.5%	64.1%	2.6%
6	Code of Practice	997,228	45	0.0045%	4.4%	60%	22.2%	13.3%
7	Notes	156,255	4	0.0026%		100%		
8	Product Descriptions	611,549	14	0.0023%	50%	21.4%	28.6%	
9	Reports ** ?	979,170	22	0.0022%	90.9%	9.1%		
10	Conference Proceedings	196,498	3	0.0015%	33.3%	66.7%		
11	Technical Papers	65,731	1	0.0015%				100%
12	Circular Letters	143,313	2	0.0014%	100%			
13	Manuals	296,299	4	0.0013%		50%	50%	
14	Review Papers	106,506	1	0.0009%		100%		
15	About Us	647,013	6	0.0009%	33.3%		66.7%	
16	Transaction Proceedings (HKIE)	1,055,248	8	0.0008%	50%		25%	25%
17	Standards	136,024	1	0.0007%		100%		
18	Media Releases	1,566,742	8	0.0005%	25%	62.5%	12.5%	

Handbooks is the sub-corpus with the highest percentage of occurrence of *complies with*. There are 9 instances and they are shown in Figure 31. They are found to be associated with the meanings of 'making certain', 'characterising and setting parameters' and 'evaluating'. To the left of *complies with*, there are relative pronoun (33.3%, 3 times), conjunction (33.3%, 3 times), pronoun (22.2%, 2 times) and a noun phrase. To the right, there is always a noun phrase (100%). Collocates to the left are 'which' (3 times), 'plug' (3 times), 'check' (2 times), 'it' (2 times), 'safe' (2 times), etc. To the right, there are 'requirements' (4 times) and 'drawings' (2 times). The semantic

preference tends to be 'requirements' due to the occurrence of words such as 'requirements', 'standards' and 'standard'.

```
1    Engineer  We certify that : (i) Our design complies with the standards set out in the Agreement
2    done by the Contractor and check whether it complies with Specification drawings and approved
3 f work done by contractor and check whether it complies with rectification drawings and approved
4    that the electrical installation is safe and complies with the statutory safety requirements.
5    that the electrical installation is safe and complies with the statutory safety requirements.
6    and design, quality control. setting out and complies with all survey requests of the engineer in
7 pliance must be fitted with a 3-pin plug which complies with the safety requirements, and its parts
8 pliance must be fitted with a 3-pin plug which complies with the safety requirements.    2.1.4 Do not
9        is fitted with a 2-round-pin plug which complies with safety standard BS4573 or EN50075 and is
```

Figure 31: Concordance lines of *complies with* in Handbooks

Regarding the semantic prosody, lines 2 to 5 are associated with the meaning of 'making certain (44.4%, 4 times), as characterised by the use of the words 'check' as in lines 2 and 3, and 'confirm' as viewed in the expanded version of lines 4 and 5. These two words suggest that the readers, or the parties concerned, need to make sure that the requirements are followed. The semantic prosody of 'evaluating' (22.2%, 2 times) is associated with the use of the phrasal verb in lines 1 and 6. In these two cases, the use of *complies with* is to report on whether the item, i.e. the noun phrase to the left of *complies with*, conforms to the requirements or not. The remaining three lines, i.e. lines 7 to 9, are associated with the meaning of 'characterising and setting parameters' (33.3%, 3 times). This semantic prosody is mainly suggested by the pattern of *relative pronoun + complies with + noun phrase*. The relative pronoun 'which' refers to the preceding noun phrase, i.e. 'a 3-pin plug' in lines 7 and 8, and 'a 2-round-pin plug' in line 9. It might appear that these lines share the pattern and meaning of 'evaluating'. However, it is argued that the focus in these lines is not on whether the noun phrase conforms to the requirements, rather, *complies with* together with its following noun phrase, i.e. 'complies with the safety requirements' as in lines 7 and 8, and 'complies with safety standard BS4573 or EN50075' in line 9, serve to be a piece of supplementary information characterising what

the noun phrase exactly denotes. Thus, there is a difference between the semantic prosodies of 'evaluating' and 'characterising and setting parameters'.

Ordinances is the sub-corpus with the second highest percentage of occurrence of *complies with*. There are 15 instances in total. They are also associated with the meaning of 'evaluating', 'making certain' and 'characterising and setting parameters'. In addition to these meanings, most instances in this sub-corpus are associated with the meaning of 'introducing a condition'

```
1        (a) to the owner of a road tanker which complies with the conditions specified in Part 1 of
2 or (b) to the owner of a cylinder wagon which complies with the conditions specified in Part 2 of
3the Authority determining whether that vehicle complies with the relevant conditions specified in
4 ) shall be to ascertain whether a gas vehicle complies with these regulations and any conditions of
5    installation unless a safety device which complies with subregulation (2)is fitted to the tanker
6 of which has an electrical  resistance which complies with British Standard 2050; or(b)an antistatic
7      heating or cooling unless the appliance complies with regulation 3B(1)(d) of the Gas Safety
8  applicable, include- (a)details as to how it complies with its duty under regulation 9;(b)details as
9  tower working platform, as the case may be, complies with the relevant requirements set out in
10  tower working platform, as the case may be, complies with the relevant requirements set out in
```

Figure 32: Sample concordance lines of *complies with* in Ordinances

Figure 32 shows a sample of concordance lines from Ordinances. Regarding the pattern of colligation, there is a relative pronoun (26.7%, 4 times), noun phrase (20%, 3 times) or pronoun (13.3%, 2 times) at the N-1 position. In six instances, the phrasal verb appears to be the first item of the clause, as in lines 9 and 10. But the phrase 'as the case may be' acts as an adverbial phrase, and so, it can be said that a noun phrase comes before *complies with* in these six instances. Putting these six instances into consideration, a noun phrase typically occurs to the left of *complies with* in this sub-corpus (60%, 9 times). To the right of the phrasal verb, there is always a noun phrase (100%). The most frequent configuration of this phrasal verb in the sub-corpus is:

noun phrase + complies with + noun phrase

Collocates to the left include 'case' (6 times), 'lift' (6 times), 'which' (4 times), 'ascertain' (2 times), 'unless' (2 times), 'vehicle' (2 times), etc. To the right, there are 'relevant' (7 times), 'requirements' (6 times), 'section' (6 times), 'conditions' (4 times), 'regulation' (2 times) etc. There is a strong semantic preference of 'requirements' (100%) as postulated by the collocates occurring to the right of the phrasal verb.

As for the semantic prosody, 53.3% of the instances are associated with the meaning of 'introducing a condition'. This is suggested by the occurrence of the word 'unless', as in lines 5 and 7. The word 'unless' is used to introduce a condition or circumstance. In another six instances, there is the word 'where', for example:

Where the registered examiner, after carrying out the test and examination, is satisfied that the builder's lift or tower working platform, as the case may be, complies with the relevant requirements set out in section 9 and is in safe working order, he shall forthwith deliver to the registered contractor, in the form and containing the information specified by the Director, a certificate of test and examination in duplicate.

The word 'where', being used this way, has similar function as 'if'. Therefore, in this example, 'where' is used to introduce a condition, which is 'the builder's lift or tower working platform' conforming to 'the relevant requirements set out in section 9 and is in safe working order'. Only when this condition exists, then 'a certificate of test and examination in duplicate' can be delivered to 'the registered contractor'.

Based on the findings on Table 18, *complies with* is not highly specific to a particular meaning, but fourteen out of eighteen of the sub-corpora in which *complies with* is found have instances associating with the meaning of 'making certain'. Also, in most of the sub-corpora in which this meaning is found, it has the highest percentage of

occurrence as compared to other meanings, for example, Handbooks, Frequently Asked Questions, Publicity Materials, Codes of Practice, Conference Proceedings, etc. Therefore, there is a tendency for the phrasal verb *complies with* to be used with the sense of 'making certain' in engineering genres.

Non-contiguous instances are found in four sub-corpora. They are used in the form of *complies…with*. It is found that the difference in the form does not make a different meaning between *complies with* and *complies…with*. In total, there are 12 instances of *complies…with* in the whole engineering corpus, half of them have the meaning of 'making certain', and the other half expresses the meaning of 'evaluating'.

Figure 33 shows the concordance lines of *complies…with* in Publicity Materials and Guides. Lines 1 and 2 are from Publicity Materials and lines 3 to 6 are from Guides. These six instances are associated with the meaning of 'making certain'. Regarding the colligation, a noun phrase occurs immediately to the left of *complies* and to the right of *with*. In the middle of *complies* and *with*, there is an adverb. So, the typical configuration of *complies…with* in these two sub-corpora is:

noun phrase + complies + adverb + with + noun phrase

Collocates to the left of *complies* are 'demonstrate' (4 times), 'device' (4 times), 'application' (4 times), 'ensure' (2 times), and 'routing' (2 times). To the right of with, there are 'environmental' (4 times), 'performance' (4 times), 'current' (2 times) and 'Fire Services Department's' (2 times). In between *complies* and *with*, there are 'in full' (4 times) and 'fully' (2 times).

```
1 Customer is required to ensure that the routing complies fully with current Fire Services Department
2 Customer is required to ensure that the routing complies fully with current Fire Services Department
3    to demonstrate  the device under application complies in full with the environmental performance
4    to demonstrate  the device under application complies in full with the environmental performance
5    to demonstrate  the device under application complies in full with the environmental performance
6    to demonstrate  the device under application complies in full with the environmental performance
```

Figure 33: Concordance lines of *complies…with* in Publicity Materials and Guides

The concordance lines do not show the complete information following the word *with*. Referring back to the original texts reveals that the noun phrase occurring to the right of *with* is 'current Fire Services Department's regulations' in lines 1 and 2, and 'the environmental performance specifications' in lines 3 to 6. These noun phrases contribute to the semantic preference of 'requirements' (100%) as regulations and specifications are some kinds of directives and precise criteria requiring people to follow. As for the semantic prosody, all the lines are associated with a sense of 'making certain' as evident by the words 'ensure' as in lines 1 and 2, and 'demonstrate' as in lines 3 to 6. This means that the people concerned need to be certain and show clearly that they conform to the requirements. In these non-contiguous instances, the words dropping in between *complies* and *with* do not change the meaning. However, it serves to emphasize the extent to which the requirements need to be conformed to, which is in full compliance. Therefore, the use of *complies…with* in these instances is used to express that people concerned need to be certain and show clearly that they are behaving in full compliance with the requirements.

There are another six instances of *complies…with*. Figure 34 shows the concordance lines of these instances. Lines 1 to 4 are from Product Descriptions, and lines 5 and 6 are from Circular Letters. The colligation of these six instances is the same as that of the other *complies…with* instances described above, that is:

noun phrase + complies + adverb + with + noun phrase

Following the word *with*, the noun phrase is 'the drinking water standard'. The semantic preference of 'requirements' is also observed in these lines. In line 4, the noun phrase 'BS 142-3.2 and IEC 255-4' represents the codes of some standards or regulations. They are some levels of quality which are required to be followed.

```
1 water provided by the Water Supplies Department complies fully with the drinking water standard acc
2 water provided by the Water Supplies Department complies fully with the drinking water standard acc
3 water provided by the Water Supplies Department complies fully with the drinking water standard acc
4    the need to use non-standard C.T. The relat complies in part with BS 142-3.2 and IEC 255-4.  Lay
5 water provided by the Water Supplies Department complies fully with the drinking water standard acc
6 water provided by the Water Supplies Department complies fully with the drinking water standard acc
```

Figure 34: Concordance lines of *complies…with* in Product Descriptions and Circular Letters

Regarding the semantic prosody, these lines are associated with the meaning of 'evaluating'. The phrasal verb in these instances are used to evaluate and report whether the noun phrases occurring to the left, i.e. 'the treated fresh water provided by the Water Supplies Department' as in lines 1 to 3 and lines 5 and 6, and 'the relay' as in line 4, follow the required standard. The word 'fully' and the phrase 'in full' intervening the phrasal verb do not affect the semantic prosody of the phrasal verb. They describe more specifically the extent of the compliance. Thus, *complies…with* in these instances is used to evaluate the extent to which the items conform to the required standard. As the same as the other inflectional forms, the use and meaning of *complies with* and *complies…with* do not vary, and thus, they are the phraseological variation of the same phrasal verb.

4.4.5 Summary

To conclude, the analysis of the use of the four inflectional forms comprising COMPLY WITH, *comply with, complying with, complied with* and *complies with*, shows that inflectional forms of the same lemma do not necessarily share similar co-selections and

196

usage. They tend to be associated with different meanings within the engineering context. The most frequent form *comply with* is more typically used with the meaning of 'obligatory', *complying with* tends to be used for 'characterising and setting parameters', and *complies with* is more frequently used with the meaning of 'making certain'. As for *complied with*, it does not seem to have specificity to any one of the meanings. In other words, they may be regarded as four distinct phrasal verbs based on the variation in their frequency of use and meanings. It may be true that the use of certain phrasal verbs does not differ from their inflectional forms, such as BASE ON and CARRY OUT. However, unless proved otherwise, it is inappropriate to lump the inflectional forms together and present them based on the lemma considering the varied frequencies of use and patterns of co-selections. And in this case, the four forms of COMPLY WITH demonstrate such characteristic. These findings may have implications for the presentation of phrasal verbs in dictionaries and teaching materials.

4.5 Use of phrasal verbs: REFER TO

This section analyses and discusses the use of *refer to* and other related inflectional forms.

4.5.1 *Refer to*

The fourth most commonly occurring phrasal verb in the HKEC is *refer to*. It is found in 25 of the sub-corpora, and in twelve of them, *refer to* ranks among the ten most frequent phrasal verbs. Table 19 shows the sub-corpora in which this phrasal verb is found. In this table, the sub-corpora are listed based on the percentage of occurrence of *refer to*. Since the sizes of the sub-corpora vary, there is possibility that sub-corpora with the highest percentage of occurrence may have a lower actual frequency.

Table 19: Sub-corpora in which *refer to* is found and its frequencies

No.	Sub-corpora	Freq of the PV	%
1	Letters to Editor	4	0.1145%
2	Frequently Asked Questions	27	0.0485%
3	Consultation Papers	42	0.0377%
4	Transaction Discussions (HKIE)	2	0.0280%
5	Handbooks	16	0.0238%
6	Manuals	68	0.0229%
7	Guides	173	0.0221%
8	Media Releases	339	0.0216%
9	Code of Practice	186	0.0187%
10	Position Documents	13	0.0172%
11	Reports	133	0.0136%
12	Notes	21	0.0134%
13	Publicity Material	79	0.0132%
14	Standards	17	0.0125%
15	Q & A	3	0.0108%
16	Product Descriptions #	51	0.0083%
17	Technical Papers	5	0.0076%
18	Conference Proceedings	9	0.0046%
19	Transaction Proceedings (HKIE)#	44	0.0042%

20	About Us	26	0.0040%
21	Ordinances	5	0.0036%
22	Review Papers	3	0.0028%
23	Abstract	2	0.0021%
24	Circular Letters	3	0.0021%
25	Agreements	1	0.0008%

Examination of the concordance lines across the sub-corpora shows that four meanings are associated with *refer to* in engineering genres. Four sub-corpora are selected to illustrate the four respective meanings. Letters to Editor is the sub-corpus with the highest percentage of occurrence of *refer to*. However, there are only four instances because it is one of the smallest sub-corpora. Figure 35 shows the four concordance lines of *refer to* in Letters to Editor.

```
1  article about environmental impact assessment  I refer to Ms Christine Loh's article "Law of the La
2         day for crucial meeting on air quality"  I refer to Ms Claire Garner's article "Wrong day for
3      project not a victory if the people lose  We refer to the letter of Ms Karen Winton, entitled
4    - Response to letter about amendment of APCO  I refer to David Renton's letter of 19 April 2008
```

Figure 35: Concordance lines of *refer to* in Letters to Editor

A strong colligational pattern can be observed from the concordance. In all the four lines, there is a personal pronoun immediately to the left of *refer to,* and a noun phrase to the right, with a personal name in the noun phrase. Thus, the regular configuration of this phrasal verb in Letters to Editor is:

personal pronoun +refer to + noun phrase (with a personal name)

Most of the personal pronouns are realised as "I" (3 times, 75%). The remaining instance is the plural personal pronoun "we". To the right of the phrasal verb, the collocates include "article" (2 times, 50%) and "letter" (2 times, 50%). This collocation contributes to a semantic preference of 'a written form of document or communication' (100%). If

the concordance lines are expanded, in addition to a personal name, a date and an article title or a topic are stated in all four instances as in, for example, "I refer to Ms Christine Loh's article 'Law of the Land' on 8 May 2008" in line 1. The presence of the personal name, title and date suggests that the semantic prosody is 'precise' (100%). This information makes it easier to identify which article the writer is talking about. The co-selections, therefore, suggest that *refer to* in Letters to Editor is used to make a precise reference to an article or a letter on a specific topic published on a specific date.

In fact, these lines demonstrate intertextuality. Intertextuality considers discourse as "text in context" and how the texts are overlapped and interconnected to other texts, past or present, so as to seek their meaning in relation to other texts (Tekin, 2008, p. 733). This use of *refer to* serves the function of situating the current text relative to a prior text (Warren, 2013). In this case, the letter to editor is the current text, and the prior text is the "article" discussed in the lines. The writers precisely signal which prior article they want to discuss, follow up or comment on. This extended unit of meaning is also dominant in Transaction Discussions (HKIE) and Q & A.

In other sub-corpora, *refer to* has very different co-selections. In Consultation Papers, for example, there are 42 instances of *refer to* in total. Figure 36 shows a sample of concordance lines.

```
1   initiative in the 2006-07 Policy Agenda. Please refer to paragraph 2.16 above. Environment,
2      the implementation of the programme (please refer to para.16 for details). Participating
3   initiative in the 2007-08 Policy Agenda. Please refer to paragraph 2.16 above. 4.12 Initiative
4            of buildings. The regulations normally refer to codes and standards developed by
5       of tungsten light bulbs with CFLs, please refer to our response to Greenpeace China above.
6   initiative in the 2007-08 Policy Agenda. Please refer to paragraph 2.12 above.  Environmentally
7   initiative in the 2007-08 Policy Agenda. Please refer to paragraph 2.25 above. 4.2 Initiative
8   initiative in the 2007-08 Policy Agenda. Please refer to paragraph 2.24 above. 4.3 Initiative
9   initiative in the 2007-08 Policy Agenda. Please refer to paragraph 2.26 above. 4.4 Initiative
10  initiative in the 2007-08 Policy Agenda. Please refer to paragraph 2.20 above. 4.5 Initiative
```

Figure 36: Sample concordance lines of *refer to* in Consultation Papers

Strong collocates occur on both sides of the phrasal verb. To the left of *refer to*, 90.5% of the instances (38 times) has the word "please". On the right hand side, there are 'paragraph' (37 times, 88.1%) and "above" (37 times, 88.1%). So in terms of colligation, an adverb typically occurs at the N-1 position. There is typically a noun phrase with a numeral followed by an adverb immediately following *refer to*. The configuration is:

adverb + refer to + noun phrase with numeral + adverb

The semantic preference of 'inter-text source' (90.5% of the instances) is observed based on the occurrence of word sequence following *refer to*. The numeral representing the particular paragraph such as "paragraph 2.25", and the adverb "above" indicate that the source of the information can be found in the same piece of text. As for the semantic prosody, the frequent use of "please" tends to suggest an attitudinal meaning of 'recommending' (90.5% of the instances). Overall, *refer to* in the Consultation Papers is used to express the meaning of recommending readers to inter-text sources.

Another meaning associated with *refer to* is to directly ask readers to read some source of texts. In Codes of Practice, there are 186 instances of this phrasal verb. 96.8% of which (180 times) are contiguous instances. Figure 37 shows a sample of concordance lines of *refer to* in Codes of Practice.

```
1 loads on erected elements at construction stage (refer to clause 2.5.1); and  demoulding, storage ,
2    where effective factor under no load effect, refer to Table E2.2.is the permissible stress
3  stress amplitude with cycle number n = 2 x 106, refer to Table E2.3. Table E2.2 - Effective Factor
4  load factor may be taken as 1.2 instead of 1.4. Refer to clause 2.2.4 of BS5959-1: 2000. E2.5.5
5    yield strength of 1100 MPa. Engineers should refer to supplier's documents for details and QA.
6  from Equation 6.1 of the Code. P-delta effects :refer to the second-order effects. There are two
7     of plastic and compact beam sections should refer to the plastic behaviour as zone ?in Figure
8    capacity of semi-compact beam section should refer to the elastic moment as discussed in clause
9 lateraltorsional and flexural-torsional buckling refer to beam buckling and they are the same.
10    girders with web stiffeners and users should refer to a bridge code such as BS5400 for more types
```

Figure 37: Sample concordance lines of *refer to* in Codes of Practice

To the left of *refer to*, 12.2% (22 times) of the lines have modal verbs, such as "should", "shall" and "may". A noun phrase immediately precedes the phrasal verb in 10% of the lines (18 times). In more than two-thirds of the other instances (124 times, 68.9%), *refer to* is the first item in the clause, for example, "Refer to Eurocode 3 for details" and "For other steel grades, refer to Appendix 8.4". The phrasal verb *refer to* is predominantly followed by a noun phrase, with no particular strong collocates. Collocates on the left of *refer to* include "please" (11 times, 6.1%), "should" (11 times, 6.1%) and "shall" (8 times, 4.4%). To the right, frequent collocates include "appendix" (32 times, 17.7%), "clause" (29 times, 16.1%), "code" (19 times, 10.6%), "table" (13 times, 7.2%), "section" (11 times, 6.1%), and "details" (10 times, 5.6%). Most of the collocates to the right are text sources providing information, such as "appendix", or parts of some text sources such as "clause" or "section". Thus, the semantic preference is 'source of information' (146 times, 81.1%). Regarding the semantic prosody, a sense of 'directive' is observed in 79.4% of the instances (143 times). Unlike Consultation Papers in which there is a frequent occurrence of "please", the dominant imperative use of the phrasal verb without any politeness marker or in some cases the presence of "should" and "shall" attribute to the semantic sense of giving directives. Therefore, as suggested by the semantic preference and semantic prosody, *refer to* in Codes of Practice tends to be used to directly request readers to read some source of information.

There are six other instances of *refer to* in this sub-corpus, but they are non-contiguous and are discussed later.

Another meaning that is associated with *refer to* is defining. This meaning seems to be specific to Agreements and Review Papers. However, in these two sub-corpora, there is only a single occurrence of *refer to*. Therefore, Transaction Proceedings (HKIE)

is selected instead to illustrate this meaning. Figure 38 shows the sample concordance lines. There are 43 instances in this sub-corpus.

```
1     5. The Legends HK, UK, WG and BJ on Figure 5 refer to Hong Kong, United Kingdom. West Germany and
2  Concrete Research, Vol. 36, No. 12, June 1984. *Refer to Cement and Concrete Association Cement
3     The power consumptions quoted in the figures refer to the rate of thermal dynamic energy either
4     saturation at humidifier Q, Qe: air flow rate refer to Figure 1 in m3 s Qep, Qp: air flow rate
5  generation rate (kg h), subscripts "r", "o", "s" refer to "room", "outside", and "supply" respectivel
6       streets, as in Table 1.  * The link numbers refer to Figure 7  Table 1?" The signals and the
7       level for each aggregate, which the Authors refer to as the 'ceiling strength", beyond which
8          for the remaining.  [For details, please refer to Chung & Poon, accepted for publication (b)]
9  a catchwater, 'minor', 'moderate' and 'serious' refer to the degree of blockage. (2) For landslides
10      the footway only. 'Moderate' and 'serious' refer to landslides affecting at least one lane of
```

Figure 38: Sample concordance lines of *refer to* in Transaction Proceedings (HKIE)

Regarding the colligation, there are noun phrases (19 times, 44.2%), adverbs (6 times, 14.0%), and modal verbs (6 times, 14.0%) at the N-1 position. In 20.9% of the instances (9 times), *refer to* is the first item of the clause. To the right, there is always a noun phrase following *refer to* (40 times, 93.0%), and so the most frequent configuration of *refer to* in this sub-corpus is:

noun phrase + refer to + noun phrase

Collocates to the left of *refer to* include "please" (6 times, 14.0%), "interested" (5 times, 11.6%), and "readers" (5 times, 11.6%). To the right, there are "figure" (7 times, 16.3%), "table" (4 times, 9.3%), and "agents" (2 times, 4.7%). For semantic preference, 25.6% of the instances (11 times) have collocates which are related to illustrative diagrams, i.e. "figure" and "table". However, apart from this mild preference, the words co-occurring with *refer to* in this sub-corpus, regardless of the part of speech, do not seem to share a more specific semantic feature.

These findings show that the colligation, collocation and semantic preference are optional for the phrasal verb in this sub-corpus. As for the semantic prosody, a sense of 'clarifying and defining' is observed in 37.2% of the instances (16 times). These instances

are characterised by the pattern in which the noun phrase to the right of *refer to* often repeats or features the noun phrase to the left. For example, in line 1 in Figure 38, "HK, UK, WG" occurs to the left of *refer to*, and they represent "Hong Kong, United Kingdom, West Germany" which occurs to the right of the phrasal verb. Another example is in line 9 in Figure 38, the noun phrase "the degree of blockage" to the right of *refer to* illustrates and clarifies the noun phrase to the left of the phrasal verb, i.e., '"minor", "moderate" and "serious"'. Thus, in Transaction Proceedings, there is a tendency to use *refer to* to provide a definition or clarification of some terms or phrases. The remaining lines express the two meanings illustrated earlier, they are 'recommending readers to read some source of texts' or 'directly asking readers to read some sources of texts'.

Table 20 shows the percentages of occurrence of different semantic prosodies of *refer to* across the sub-corpora. It can be seen that *refer to* in engineering genres is not specific to any one particular meaning, but generally it is used more often to express the second and third meanings described above, namely 'recommending readers to inter-text sources' and 'directly request readers to read some source of information'.

Table 20: Meaning distribution of *refer to*

	Sub-corpora	size	Freq	%	M1	M2	M3	M4
1	Letters to Editor	3,492	4	0.1145%	100%	0	0	0
2	Frequently Asked Questions	55,726	27	0.0485%	0	77.8%	22.2%	0
3	Consultation Papers	111,494	42	0.0377%	0	90.5%	0	4.8%
4	Transaction Discussions (HKIE)	7,149	2	0.0280%	100%	0	0	0
5	Handbooks	67,284	15	0.0223%	0	33.3%	66.7%	0
6	Manuals	296,299	64	0.0216%	0	53.1%	34.4%	7.8%
7	Guides	783,805	165	0.0211%	1.2%	47.3%	41.8%	5.5%
8	Media Releases	1,566,742	334	0.0213%	0.9%	60.8%	33.5%	4.8%
9	Code of Practice	997,228	180	0.0181%	0	7.2%	80%	11.1%
10	Position Documents	75,660	5	0.0066%	0	40%	60%	0
11	Reports	979,170	132	0.0135%	0.8%	48.5%	40.2%	9.8%
12	Notes	156,255	18	0.0115%	0	55.6%	44.4%	0
13	Publicity Material	599,407	77	0.0128%	0	39.0%	49.4%	11.7%
14	Standards	136,024	16	0.0118%	0	18.8%	68.8%	12.5%
15	Q & A	27,703	3	0.0108%	66.7%	33.3%	0	0
16	Product Descriptions	611,549	48	0.0078%	0	31.3%	60.4%	0
17	Technical Papers	65,731	5	0.0076%	0	40%	60%?	0
18	Conference Proceedings	196,498	9	0.0046%	0	22.2%	55.6%	22.2%
19	Transaction Proceedings (HKIE)	1,055,248	43	0.0041%	0	20.9%	27.9%	37.2%
20	About Us	647,013	24	0.0037%	0	15.4%	73.1%	11.5%
21	Ordinances	139,176	3	0.0022%	0	0%	100%	0
22	Review Papers	106,506	1	0.0009%	0	0	0	100%
23	Abstract	94,671	2	0.0021%	0	100%	0	0
24	Circular Letters	143,313	3	0.0021%	33.3%	66.7%	0	0
25	Agreements	127,895	1	0.0008%	0	0	0	100%

M1: making clear reference
M2: politely directing readers
M3: directly requesting readers
M4: defining

In some sub-corpora, non-contiguous instances are found. Table 21 shows the sub-corpora in which non-contiguous instances are found. The frequency of the non-contiguous instances is lower than 10 in all the sub-corpora, but the proportion of the non-

contiguous instances to the total instances of *refer to* varies across the sub-corpora, ranging from 0.8% to 66.7%.

Table 21: Sub-corpora in which non-contiguous *refer to* instances are found

	Sub-corpora	size	Freq. of the PV (Non-cont)	% of the total freq. of the PV
1	Handbooks	67,284	1	6.3%
2	Manuals	296,299	4	5.9%
3	Guides	783,805	8	4.6%
4	Media Releases	1,566,742	5	1.5%
5	Code of Practice	997,228	6	3.2%
6	Position Documents	75,660	8	61.5%
7	Reports	979,170	1	0.8%
8	Notes	156,255	3	14.3%
9	Publicity Material	599,407	2	2.5%
10	Standards	136,024	1	5.9%
11	Product Descriptions	611,549	3	5.9%
12	Transaction Proceedings (HKIE)	1,055,248	1	2.3%
13	About Us	647,013	2	7.7%
14	Ordinances	139,176	2	40.0%
15	Review Papers	106,506	2	66.7%

For the sake of illustration, Guides is selected as an example in this section as it has the highest frequency of non-contiguous instances (N=8, 4.6%). Figure 39 shows all the eight concordance lines of the non-contiguous instances of *refer to* in Guides.

```
1   be issued accordingly. If necessary, FEHD will refer the case to the Water Supplies Department or
2   Committee and all owners. You are suggested to refer the case to your property manager to admin
3 Secretary must appoint a disciplinary board and refer the matter to that board for inquiry.  The
4   the Secretary must appoint an appeal board and refer the appeal to that board for inquiry.
5 Secretary must appoint a disciplinary board and refer the matter to that board for inquiry.  The
6    of internal procedures,  for example, who to refer complaints to within the office or organisati
7  your building or not. If yes, you will have to refer the contractor to the Owners' Corporation for
8        deserving of censure,   the Director may refer the matter by notice in writing to the Secreta
```

Figure 39: Concordance lines of *refer...to* in Guides

In this sub-corpus, all the non-contiguous instances display constituency variation, i.e. *refer…to*. Syntactically, to the left of *refer…to*, there is mostly a modal verb or semi-modal verb (N=6, 75%), such as "will", "have to" and "may" in lines 1, 7 and 8 respectively. In lines 3 to 5, despite the fact that the modal verb "must" does not literally come immediately in front of *refer*, the line can be syntactically restructured as, for example, in line 4:

the Secretary must appoint an appeal board and must refer the appeal to that board for inquiry.

Thus, in six out of the eight lines, a modal or semi-modal verb precedes *refer…to*. In the middle of *refer* and *to*, there is a noun phrase in every line. To the right of *to*, there is also a noun phrase in almost all of the lines (7 times, 87.5%). To summarise the colligational pattern, the most frequent configuration of *refer…to* in Guides is:

modal verb + refer + noun phrase + to + noun phrase

This collocational pattern is not particularly strong. Collocates to the left include "secretary" (3 times, 37.5%), "must" (3 times, 37.5%), "appoint" (3 times, 37.5%), "board" (3 times, 37.5%), "will" (2 times, 25%), etc. Between *refer* and *to*, collocates are "matter" (3 times, 37.5%) and "case" (2 times, 25%). To the right of *to*, there is "that board for inquiry" (3 times, 37.5%).

A strong semantic preference of 'specific matter requiring formal and authoritative advice or action' is observed in 87.5% of the lines (7 times). This semantic feature is shared by the words or phrases such as "the case", "the matter", "complaints", "for inquiry" and "the appeal". This semantic preference, and the co-occurrence of other

words or phrases such as "must", "the Secretary", "a disciplinary board", "an appeal board", "if necessary", "suggested to", "manager", and "the Director", postulate a semantic prosody of 'serious' (7 times, 87.5%). Therefore, the constituency variation *refer...to* is used to express the meaning of passing a serious matter to a party for formal and authoritative advice or action.

In Position Documents, there are also eight instances of *refer...to*. The frequency of *refer...to* is even higher than that of *refer to* in this sub-corpus. Figure 40 shows the eight instances of *refer...to* in Position Documents. In this sub-corpus, the co-selections of *refer...to* are rather similar to those in Guides. Yet, a different attitudinal sense is observed.

```
1  who can respond to a request for information or refer the caller to a knowledgeable source. OSHA
2  who can respond to a request for information or refer the caller to a knowledgeable source. OSHA
3  questions related to these regulations and will refer specific problems to informed sources. Asbest
4  questions related to these regulations and will refer specific problems to informed sources. Asbest
5          in "nonstate-plan" states, they will refer a specific inquiry to other sources for
6          in "nonstate-plan" states, they will refer a specific inquiry to other sources for
7  requirements for worker protection and will also refer detailed inquiries on all aspects of asbestos
8  requirements for worker protection and will also refer detailed inquiries on all aspects of asbestos
```

Figure 40: Concordance lines of *refer...to* in Position Documents

First of all, the colligation pattern of *modal verb + refer + noun phrase + to + noun phrase* is observed in 75% of the lines (6 times). The collocation pattern tends to be slightly stronger than that in Guides, though the set of collocates is different. To the left of *refer*, there is "will" (6 times, 75%). In the middle of *refer* and *to*, there are "specific" (4 times, 50%), "inquiry" (2 times, 25%), "inquiries" (2 times, 25%), "the caller" (2 times, 25%), and "problems" (2 times, 25%). To the right, there are "sources" (6 times, 75%), "knowledgeable" (4 times, 50%), "source" (2 times, 25%), and "informed" (2 times, 25%).

The different set of collocates suggests a different semantic preference in this sub-

corpus, that is 'inquiries requiring informed source for solutions'. Such a semantic feature is observed in 75% of the lines (6 times) with words such as "problems", "inquiries", "informed source" and "knowledgeable source". The semantic prosody is a sense of 'intention' found in 75% if the lines (6 times). This is due to the use of "will" which has the modality of expressing intention (Biber et al., 1999). Due to the differences in the co-selections, there is a shift in the meaning of *refer…to* in Position Documents as compared to that in Guides. While in Guides, the constituency variation of the phrasal verb is used to express passing a serious matter to a party for formal and authoritative advice or procedures, in Position Documents, the meaning shifts to intending to direct inquirers or specific questions to informed sources for solutions.

In spite of the low frequencies, the extended units of meaning of *refer…to* postulated from the co-texts are distinct from each other with regard to semantic prosody. The functions of the sub-corpora play an important role in the distinction. Guides are regulative in nature. Irrespective of whether the guidelines and regulations stated in the Guides are legally-bound or not, the texts in Guides aim to govern and regulate people's behaviour and activities related to engineering. Thus, the cases and matters described in Guides may be more serious in nature and tend to require formal advice or procedures. As for Position Documents, they are texts presenting comments or views on a specific topic on behalf of government bodies or organisations. They are of a totally different nature from Guides. It is found that, all the *refer…to* instances in Position Documents come from the section 'Resources' in each piece of text. This 'Resources' section serves to remind the readers the sources from which information and assistance can be sought if there are questions and specific inquiries.

These extended units of meaning are not found in the contiguous form of the

phrasal verb *refer to*. This indicates that phraseological variation has an impact on the use
and the meaning of the phrasal verb. The contiguous *refer to* and the non-contiguous
refer…to have very different co-selections.

Table 22: Meaning distribution of non-contiguous *refer to*

	Sub-corpora	size	Freq (Non-cont)	%	M2	M3	M4	M5	M6
1	Position Documents	75,660	8	0.0106%	0	0	0	0	100%
2	Notes	156,255	3	0.0019%	0	33.3%	0	66.7%	0
3	Review Papers	106,506	2	0.0019%	0	0	0	100%	0
4	Handbooks	67,284	1	0.0015%	0	0	0	100%	0
5	Ordinances	139,176	2	0.0014%	0	0	0	100%	0
6	Manuals	296,299	4	0.0013%	0	0	0	100%	0
7	Guides	783,805	8	0.0010%	0	0	0	100%	0
8	Standards	136,024	1	0.0007%	0	100%	0	0	0
9	Code of Practice	997,228	6	0.0006%	0	33.3%	0	66.7%	0
10	Product Descriptions	611,549	3	0.0005%	0	0	0	100%	0
11	Publicity Material	599,407	2	0.0003%	0	0	50%	50%	0
12	Media Releases	1,566,742	5	0.0003%	0	0	0	100%	0
13	About Us	647,013	2	0.0003%	0	0	100%	0	0
14	Reports	979,170	1	0.0001%	0	0	0	0	100%
15	Transaction Proceedings (HKIE)	1,055,248	1	0.0001%	100%	0	0	0	0

M1: making clear reference
M2: politely directing readers
M3: directly requesting readers
M4: defining
M5: passing serious matter for formal procedures
M6: intending to direct questions to knowledgeable sources

However, not all *refer…to* instances have the extended units of meaning described
above. This can be seen from Table 22. In Codes of Practice, for example, there are six
instances of *refer…to*. They are shown in Figure 41.

```
1    and in vibration modes with higher frequencies.Refer also to clause 5.3.4. 13.2 GUIDANCE ON DESIGN
2    less than 1.5% of the characteristic dead load (refer also to clause 2.3.1); and accidental loads s
3   considerable public concern, the Authority shall refer the matter for advice to the Secretary. In
4   considerable public concern, the Authority shall refer the matter for advice to the Secretary. In
5   considerable public concern, the Authority shall refer the matter for advice to the Secretary. In
6   considerable public concern, the Authority shall refer the matter for advice to the Secretary. In
```

Figure 41: Concordance lines of *refer…to* in Codes of Practice

Four of the instances, lines 3 to 6 in Figure 41, have very similar co-selections to those in Guides. The colligation pattern *modal verb + refer + noun phrase + to + noun phrase* is the same. There are shared collocates such as "matter" and "the Secretary", and the collocates also share the semantic preference of 'specific matter requiring formal and authoritative procedures or advice'. The semantic prosody of 'serious' is also observed in these four lines as evident by the co-occurrence of the phrases "considerable public concern", "the Authority" and "shall". Yet, in lines 1 and 2, despite the intervening word, the co-selections of these two instances are the same as most of the contiguous instances in the sub-corpus. The intervening word in these two lines is "also". When the intervening word "also" is taken away, the extended unit of meaning is the same as other contiguous instances and it expresses the meaning of directly requesting readers to read some source of information. The adverb "also" adds the sense of 'repetition' to the instances, implying that the source of information, i.e. "clause 5.3.4" and "2.3.1" as in Lines 1 and 2, has been mentioned earlier in the text.

Interestingly, it is noted that the non-contiguous instance in Transaction Proceedings has a different form, i.e. *to…refer*. This is the only non-contiguous instance in the sub-corpus. The co-text of the instance is extracted below:

…used in this paper follow their usual meaning as defined in the Shore Protection Manual (1984) to which readers can refer for full details…

In this case, *to* is the fronted particle and is immediately followed by "which" which is a relative pronoun denoting the antecedent "the Shore Protection Manual". Thus, the instance expresses that readers can read "the Shore Protection Manual (1984)" for full details.

Therefore, unlike the *refer…to* instances, this positional variation *to…refer* does not change the meaning of the phrasal verb *refer to*. This instance demonstrates a preposed construction, i.e. fronting the particle of the phrasal verb and the object noun phrase (Fraser, 1976; Lindner, 1981; Hampe, 2002), as *refer to* is a Type II phrasal verb with a verb and a prepositional particle.

4.5.2 *Refers to*

Of all the inflectional forms of the lemma REFER TO, only *refer to* is found among the ten most frequently used phrasal verbs across the engineering genres. The other three forms, *refers to*, *referring to* and *referred to* are discussed one by one below.

Table 23 shows the 22 sub-corpora in which *refers to* is found. It shows a very different pattern of distribution of meanings from that of *refer to*. This phrasal verb *refers to* is predominantly associated with the fourth meaning associated with *refer to* as described above, i.e. providing a definition or clarification of some terms or phrases. This meaning is found in all the instances in 15 out of the 22 sub-corpora.

Table 23: Sub-corpora in which *refers to* occurs and meaning distribution

	Sub-corpora	Size of the sub-corpus	Freq	% in the sub-corpus	M3	M4	M7
1	Plans	4,173	1	0.0240%		100%	
2	Transaction Discussions (HKIE)	7,149	1	0.0140%		100%	
3	Handbooks	67,284	7	0.0104%		100%	
4	Standards	136,024	13	0.0096%		100%	
5	Code of Practice	997,228	66	0.0066%		93.9%	6.1%
6	Notes	156,255	8	0.0051%		100%	
7	Manuals	296,299	15	0.0051%		100%	
8	Frequently Asked Questions	55,726	2	0.0036%		100%	
9	Transaction Proceedings (HKIE)#	1,055,248	36	0.0034%		100%	
10	Media Releases	1,566,742	49	0.0031%		100%	
11	Reports	979,170	29	0.0030%		37.9%	58.6%
12	Guides	783,805	22	0.0028%		100%	
13	Product Descriptions #	611,549	17	0.0028%		100%	
14	Position Documents	75,660	2	0.0026%	50%	50%	
15	Agreements	127,895	3	0.0023%		100%	
16	Publicity Material	599,407	13	0.0022%		92.3%	7.7%
17	Ordinances	139,176	3	0.0022%		100%	
18	Conference Proceedings	196,498	4	0.0020%		100%	
19	Transaction Notes (HKIE)	79,058	1	0.0013%		100%	
20	Abstract	94,671	1	0.0011%		100%	
21	Review Papers	106,506	1	0.0009%		100%	
22	About Us	647,013	6	0.0009%		66.7%	

M3 (meaning 3): directly requesting readers
M4 (meaning 4): defining
M7 (meaning 7): applying to

Plans and Transaction Discussions (HKIE) are the sub-corpus with the highest relative percentage of the frequency of *refers to*. However, there is in fact only one time of occurrence in each of the sub-corpora. It is the small sizes of the two sub-corpora

which makes them the highest percentage of occurrence. The sub-corpus ranking the second highest is Handbooks. Figure 42 shows all the seven concordance lines of *refers to* in Handbooks.

```
1  the following four stages: a. Planning stage - refers to the preparation of contents, venue, date
2   arrangement of speakers. b. Publicity stage - refers to the promotion and information dissemination.
3   and information dissemination. c. Event stage refers to the actual staging of the third party event.
4  of the third party event. d. Reporting stage - refers to the information to be collected after the
5        1.1 The term "electrical safety" broadly refers to the safety of generation, transmission,
6  valve.8.3.2 Tee-branch Valve A tee-branch valve refers to an isolation valve at a branch pipe and wh
7     3.7, 7.10, 8.16 & 9.5) - A tee-branch valve refers to an isolation valve at a branch pipe and wh
```

Figure 42: Concordance lines of *refers to* in Handbooks

With regard to the colligation, a noun phrase occupies the positions on both sides of the phrasal verb, thus, forming the following configuration:

noun phrase + refers to + noun phrase

Collocates to the left of *refers to* include "stage" (4 times, 57.1%) and "a tree-branch valve" (2 times, 28.6%). To the right, the phrase "an isolation valve at a branch pipe" occurs twice (28.6%). There is no strong collocate, and the collocates do not share any particular semantic characteristics. However, a closer inspection of the concordance lines reveals that the words or phrases following the phrasal verb always repeat or characterise those to the left of the phrasal verb. For example, in line 2, the phrase to the right of *refers to* is "the promotion and information dissemination". This phrase explains the noun phrase "publicity stage" which occurs to the left of *refers to*. In lines 1 to 4, the phrases following the phrasal verb explain and characterise the noun phrases to the left of the phrasal verb. In the case of lines 5 to 7, some words to the right of the phrasal verb repeat those to the left. For example, in line 6, the phrase to the left is "a tee-branch valve". The words "valve" and "branch" also occur on the right side of *refers to*. The semantic prosody is, therefore, postulated to be "clarifying and defining". This prosody is observed

in all the instances in this sub-corpus.

Not only is the distribution of meanings of *refers to* different from that of *refer to*, there are also new meanings revealed which are not found in the *refer to* instances. In Codes of Practice, the frequency of *refers to* is 66. Figure 43 shows a sample of concordance lines of *refers to* in the sub-corpus.

```
1  Linear vs. non-linear analysis Linear analysis refers to any analysis assuming a linear relationship
2      strain and displacement. Nonlinear analysis refers to any analysis not fulfilling this linearity
3    hot rolled sections are uncommon, the clause refers to Section 11 of the Code for the effective wi
4   of a restrained beam or in-plane beam, which refers to a beam with lateral restraint against
5       and they are the same. Torsional buckling refers to twisting mode of column buckling under axia
6 in the bolts as for non-preloaded bolts, but it refers to the frictional slip resistance of the bolts.
7       bars inside concrete, and the second type refers to factory made type anchors bolts.    For the
8    calculation methods Clause 12.4.5 of the Code refers to the moment capacity method given in BS5950:
9   steel materials is self explanatory. The list refers to steel materials standards from Australia,
10     and forgings is self explanatory. The list refers to standards from Australia, China, Japan, Nor
```

Figure 43: Sample concordance lines of *refers to* in Codes of Practice.

93.9% of the instances (62 times) have the same extended unit of meaning described above which is providing clarification or definition. Except for line 8, the co-selections of *refers to* in all the lines in Figure 43 are very similar to those described above. For example, in line 9, the extract of the wider co-texts is given below:

…The list of standards for steel materials is self explanatory. The list refers to steel materials standards from Australia, China, Japan, North America and UK versions of European standards…

So "the list of standards for steel materials" is actually the referent of "the list" which precedes *refers to*. Thus, *refers to* in these lines is associated with the meaning of clarifying or defining.

In the remaining 6.1% of the instances (4 times), *refers to* has a different extended unit of meaning. In line 8 in Figure 43, *refers to* shares the same pattern of colligation, i.e.

noun phrase + refers to + noun phrase. However, unlike the instances described above, these four instances tend to have a semantic preference of 'codes and standards'. This is suggested by the collocates occurring to the left of *refers to*, they include "clause", "standard" and "section". These words and phrases represent some kind of codes or standards. In these instances, the noun phrases to the right of *refers to* do not characterise those to the left. They seem to express what the codes or standards apply to. For example, in line 8, the codes stated in "Clause 12.4.5" apply to "the moment capacity method".

For non-contiguous instances *refers…to*, only 3 sub-corpora have such occurrence and the frequency is very low. The non-contiguous instances are found in Transaction Proceedings (HKIE), Publicity Material and Ordinances. Unlike the case of *refer to* and *refer…to* in which meanings are influenced by phraseological variation, *refers…to* does not have different meanings from that of *refers to*. In Transaction Proceedings (HKIE), there are two non-contiguous instances and they display constituency variation. They are extracted below:

Blast furnace stag <u>refers</u> in particular <u>to</u> the slag produced from the…
…the term HEV in this paper <u>refers</u> only <u>to</u> the vehicle adopting the ICE and electric…

In spite of the intervening words, *refers…to* in these two instances has the same meaning as other *refers to* instances in this and other sub-corpora which is providing clarification or definition. The intervening words 'in particular' and 'only' in the two cases are adverbials, meaning specifically and solely respectively in the two lines. The addition of these intervening words does not shift the meaning of the phrasal verb but they serve to set a limit or emphasize the meaning of *refers to*.

There are two non-contiguous instances in both Publicity Material and Ordinances.

They all display positional variation, i.e. *to…refers*. It is found that the two instances in the two sub-corpora are identical, except for the numbering:

1.22.13 On each test certificate, sufficient information including the specification, contract number and equipment details, shall be given for the ready identification of the material or equipment <u>to</u> which the certificate <u>refers</u>.

1.21.9 On each test certificate, sufficient information including the specification, contract number and equipment details, shall be given for the ready identification of the material or equipment <u>to</u> which the certificate <u>refers</u>.

In these two instances, the phrasal verb displays a preposed structure. As discussed earlier, this is a feature of a Type II phrasal verb. In such a preposed construction, the relative pronoun "which" denotes the noun phrase immediately in front of it, i.e. "the material or equipment". These two instances tend to express that the test certificate applies to the particular material and equipment. Thus, in these cases, variation in the form of the phrasal verb does not affect the meaning.

4.5.3 *Referring to*

While *refers to* tends to be highly specific to a particular meaning across most of the sub-corpora, this is not the case for *referring to* and *referred to*. In total, *referring to* is found in 14 sub-corpora, but in one of the sub-corpora, only non-contiguous instances are found. The frequency of use of this phrasal verb is generally low. Even in the sub-corpus in which this phrasal verb has the highest percentage of occurrence, there are only 29 times of occurrence. Table 24 shows the sub-corpora in which *referring to* is found and the meaning distribution in different sub-corpora.

Table 24: Sub-corpora in which *referring to* is found and the meaning distribution

	Sub-corpora	Size of the sub-corpus	Freq	% in the sub-corpus	M2	M4	M8	M9	M10
1	Transaction Proceedings (HKIE)	1,055,248	30	0.0028%	6.7%	6.7%	72.4%	10%	0
2	Code of Practice	997,228	11	0.0011%	18.2%	18.2%	18.2%	45.5%	0
3	Reports	979,170	6	0.0006%	0	0	66.7%	33.3%	0
4	Product Descriptions	611,549	5	0.0008%	0	0	0	100%	0
5	About Us	647,013	5	0.0008%	0	60%	0	20%	0
6	Guides	783,805	4	0.0005%	0	50%	0	50%	0
7	Media Releases	1,566,742	4	0.0003%	0	0	25%	25%	50%
8	Manuals	296,299	3	0.0010%	0	0	0	100%	0
9	Technical Papers	65,731	2	0.0030%	0	0	50%	0	50%
10	Publicity Material	599,407	2	0.0003%	0	0	0	100%	0
11	Transaction Notes (HKIE)	79,058	1	0.0013%	0	0	0	100%	0
12	Conference Proceedings	196,498	1	0.0005%	0	0	100%	0	0
13	Abstract	94,671	1	0.0011%	0	100%	0	0	0

M2: politely directing readers
M4: defining
M8: illustrating information
M9: using as references
M10: addressing job titles

The majority of them tend to express the meaning of 'using some source of information as reference' and 'illustrating'. Transaction Proceedings, which is the sub-corpus with the highest percentage of occurrence of referring to, is used for illustration. Figure 44 shows a sample of concordance lines of *referring to* in Transaction Proceedings.

```
1    depth of the neutral axis, xbal is however referring to the theoretical balanced section and there
2  crack widths for immature concrete other than referring to BS5337 (old water retaining structures
3  Wave equation, velocity wave and force wave  Referring to Fig.A-1, according to Newton"s second law,
4    of Wave Reflection and Wave Transmission   Referring to Fig. 1, the wave reflection and transmissi
5   the outdoor air temperature. Occupancy Rate  Referring to the daily profiles in figure 8, the daily
6  so on for F and ? and E and ER4 respectively. Referring to Figure 5 microwaves travelled from the top
7  a clean LDPE CF (33%) sample used as control. Referring to Figure 4, it is found that the elements
8   the curves of Figures 6 and 7 respectively. Referring to Figure 7, it is found that there is only
9   atoms in C-H and C-C bonds in LDPE GF (33%). Referring to Figure 6, it is found that there are two
10  point Is 6000 x l0 (400 x 1200) = 12 5 MPa,  Referring to Table 2 the allowable nodal stress factor
```

Figure 44: Sample concordance lines of *referring to* in Transaction Proceedings

There is a strong colligation pattern in this sub-corpus. The colligation is characterised by the fact that *referring to* itself is always the first item in the clause (82.8%, 24 times), and to the right, there is always a noun phrase (100%). Thus the most frequent configuration can be expressed as:

referring to + noun phrase

This pattern of colligation looks the same as that of the imperative use of *refer to*, i.e. 'directly asking readers to read some source of information'. However, a closer inspection shows that there is always another clause following the clause with *referring to*, whereas in the case of *refer to*, there is none. For example, in line 4 in Figure 44, "Referring to Fig. 1" is the first clause in the sentence, it is followed by the clause "the wave reflection and transmission in a pile with a change in impedance are calculated and summarized in this appendix".

Since *referring to* is often the first item of the clause, collocates are mainly identified from the right side of the phrasal verb. A rather strong collocate is 'figure' or sometimes 'fig.' as shown in lines 3 and 4 in Figure 44. It is found in 69.0% of the lines (20 times). The semantic preference is thus 'figure'. Regarding the semantic prosody, it is observed that the clause following the clause with *referring to* is usually illustrating the information indicated in the figure. Hence, the semantic prosody is suggested to be 'illustrative' (75.9%, 22 times). To summarise, this phrasal verb *referring to* is used to express the meaning of 'illustrating the information indicated in figures'.

The sub-corpus with the second highest percentage of occurrence of *referring to* is Codes of Practice. There are 11 instances in the sub-corpus. Figure 45 shows the concordance of all these 11 instances.

```
1  is designated as case 3. 58 Case2 Unbraced case Referring to Table 1, the selected £ps is 0.00085 and
2    whether directly or indirectly as a result of referring to or using any information contained in
3   term 'steel pipelines' is  generally used when referring to welded steel pipelines operating at high
4      company' is used throughout this code when referring to the owner of a gas pipe.  For the purpo
5    whether directly or indirectly as a result of referring to or using any information contained in
6    whether directly or indirectly as a result of referring to or using any information contained on
7    whether directly or indirectly as a result of referring to or using any information contained in
8     relay contactors, or the electrical devices (referring to in Section K3) or (b) be a safety circui
9   a Banquet Room, a Seminar Room or a Ball Room. Referring to Form LG-2 of Appendix C3, the LPD
10 directly cutting the supply to the contactors (referring to Section K3) or their relay contactors,
11 to de-energized or current-balanced cables, by referring to the electricity supplier's cable plans.
```

Figure 45: Concordance lines of *referring to* in Codes of Practice

The colligation of *referring to* in this sub-corpus is not strong. To the left of the phrasal verb, there is a prepositional phrase in four of the lines (36.4%). In another four of the lines (36.4%), *referring to* is the first item in the clause. In the remaining three lines, there is an adverb (18.2%, 2 times) and a single instance of a preposition. To the right of the phrasal verb, there is a noun phrase (54.5%, 6 times), 'or' (36.4%, 4 times), and a single occurrence of a preposition.

In terms of collocation, collocates on the left side of the phrasal verb include "as a result of" (4 times, 36.4%) and "when" (2 times, 18.2%). Collocates to the right of the phrasal verb include "or" (4 times, 36.4%) and "section" (2 times, 18.2%). In fact, the four lines with "or" can be rewritten as the following:

…as a result of referring to any information or using any information contained…

Since "or" indicates alternative options, "using" and "referring to" are, therefore, alternatives. Thus, in other words, it can be said that *referring to* in these four lines is followed by a noun phase rather than the conjunction "or". No particular semantic preference is identified from these concordance lines, as there is no strong collocational pattern. As for semantic prosody, 'using as reference' is observed in 45.5% of the instances (5 times). This is suggested by the use of 'as a result of' and 'by' in lines 2, 5, 6, 7 and 11. In lines 2, 5, 6 and 7, the word 'using' occurs after the word 'or'. This indicates

that 'using' is an alternative of *referring to*. Thus *referring to* in these four lines are used to express 'using'. In line 11, there is 'by' immediately in front of *referring to*. This preposition has the meaning of indicating the means of achieving something.

The remaining instances are associated with other meanings. They are 'directly asking readers to read some source of information' as in lines 8 and 10, 'providing definition or clarification' as in lines 3 and 4, and 'illustrating information indicated in figures' as in lines 1 and 9.

In addition to the above extended units of meaning, some new meanings are revealed and they are not found in the use of *refer to* and *refers to*. In Media Releases, there are four instances of *referring to* in total. They are shown in Figure 46.

```
1  8  March 2007  The Regional Challenge Fund Referring to the December issue of the Newsletter, it has
2  required. Contractual roles Take care when referring to job titles to impose consistent and correct
3  orrect use of initial capital letters. When referring to an engineer in general, the title does not
4    - Check the status of the projector with referring to " INDICATORS AND PROJECTOR CONDITION" on pa
```

Figure 46: Concordance lines of *referring to* in Media Releases

As shown in Figure 46, *referring to* in all the four lines is followed by a noun phrase. However, in lines 2 and 3, as different from the other instances discussed above, the noun phrase in these two lines does not represent information of any sort. They are "job titles" and "an engineer". In line 3, the phrase "the title" represents "an engineer". Thus, these two lines have the semantic preference of 'job title'. It is argued that such words or phrases postulate a semantic prosody of 'addressing'. This semantic prosody is also identified in Technical Papers. Despite the limited frequencies of such meaning, the pattern of co-selection is distinctive. It is interesting to find a new meaning used within Engineering English, but further evidence will be needed to prove whether such meaning is specific to any engineering genres.

Agreements is the only sub-corpus in which non-contiguous instances are found. The raw frequency is very low. There are only two times of occurrence. Both cases display constituency variation. Their concordance lines are extracted below:

The need for normalization can be explained by referring again to the case of assessing the overall performance of students.

The function of the weighting step can be appreciated by referring once again to the case of student assessment.

The two instances of *referring…to* share the semantic prosody of 'using as references'. This is mainly due to the use of "by" in front of *referring*. Even though there are intervening words, which are "again" and "once again" respectively, they can be taken away without changing the meaning of the lines. The co-occurrence of these intervening words adds a sense of 'once more'. Thus, rather than changing the meaning of the phrasal verb, it indicates that the information, which in this case is "the case of assessing student performance", has been referred to earlier and can now be used one more time as reference.

4.5.4 *Referred to*

The last phrasal verb *referred to* comprising REFER TO is found in 26 sub-corpora. It has the second highest overall frequency among the four phrasal verbs. Table 25 shows the sub-corpora in which *referred to* occur and the distribution of meanings.

Table 25: Sub-corpora in which *referred to* occurs and meaning distribution

	Sub-corpora	Size of the sub-corpus	Freq	% in the sub-corpus	M1	M2	M3	M4	M5	M6	M11	M12
1	Letters to Editor	3,492	3	0.0859%	66.7%	0	0	0	0	0	0	0
2	Ordinances	139,176	114	0.0819%	0	0	0.9%	0	0	0	96.5%	2.6%
3	Circular Letters	143,313	34	0.0237%	0	0	14.7%	0	0	5.9%	58.8%	20.6%
4	Abstract	94,671	19	0.0201%	0	0	0	0	0	0	84.2%	10.5%
5	Manuals	296,299	56	0.0189%	0	0	16.4%	5.5%	10.9%	0	40.0%	27.3%
6	Agreements	127,895	23	0.0180%	0	0	0	0	0	0	0	100.0%
7	Code of Practice	997,228	157	0.0157%	0	3.2%	14.6%	0	1.3%	0	64.3%	15.3%
8	Guides	783,805	86	0.0110%	0	3.5%	15.1%	0	5.8%	0	47.7%	27.9%
9	Notes	156,255	16	0.0102%	0	0	0	6.3%	56.3%	0	6.3%	31.3%
10	Standards	136,024	12	0.0088%	0	8.3%	0	0	16.7%	8.3%	33.3%	25.0%
11	Publicity Material	599,407	50	0.0083%	0	10.0%	6.0%	2.0%	8.0%	0	58.0%	16.0%
12	Transaction Proceedings (HKIE)	1,055,248	83	0.0079%	0	12.2%	2.4%	4.9%	0	0	22.0%	57.3%
13	Technical Papers	65,731	5	0.0076%	0	20.0%	0	0	20.0%	0	0	60.0%
14	Transaction Notes (HKIE)	79,058	6	0.0076%	16.7%	0	16.70%	16.7%	0	0	16.7%	33.3%
15	About Us	647,013	40	0.0062%	2.5%	0	0	0	12.5%	0	55.0%	30.0%
16	Handbooks	67,284	4	0.0059%	0	0	0	0	25.0%	0	25.0%	50.0%
17	Position Documents	75,660	4	0.0053%	0	0	0	0	0	0	0	100.0%
18	Review Papers	106,506	5	0.0047%	0	0	0	0	20.0%	0	60.0%	20.0%
19	Q & A	27,703	1	0.0036%	0	0	0	0	0	0	100.0%	0
20	Frequently Asked Questions	55,726	2	0.0036%	0	0	0	0	0	0	0	100.0%
21	Project summaries	115,829	4	0.0035%	0	0	0	0	25%	0	0	75%
22	Media Releases	1,566,742	53	0.0034%	1.9%	0	0	0	9.4%	0	54.7%	34.0%
23	Reports	979,170	31	0.0032%	0	0	0	9.7%	19.4%	0	48.4%	22.6%
24	Conference Proceedings	196,498	6	0.0031%	0	0	0	16.7%	0	0	0	83.3%
25	Product Descriptions	611,549	14	0.0023%	0	0	0	0	7.1%	21.4%	14.3%	57.1%
26	Consultation Papers	111,494	2	0.0018%	0	0	0	0	50.0%	0	0	50.0%

M1: making clear reference
M2: politely directing readers
M3: directly requesting readers
M4: defining
M5: passing serious matter for formal procedures
M6: intending to direct inquiries
M11: locating information
M12: identifying with terms

As the same as the case of *refer to*, Letters to Editor is the sub-corpora in which *referred to* has the highest percentage of occurrence. Due to the small size of the sub-corpus, there are only three times of occurrence. The three concordance lines are extracted below:

The public opinion survey referred to by Mr Tam indicated that nearly 90% of the public

agreed that …

We also note that heavy-gauge plastic bin liners, which Mr Tam referred to, are generally

sold at a price, and should not be abused as in the case of free plastic shopping bags.

page A12 of SCMP on 3 December 2007. The letter referred to the Pak Tam Road street

lighting installation

There is no distinct pattern of colligation. In the first instance, *referred to* is a past participle. The prepositional phrase "by Mr Tam" serves to introduce that "Mr Tam" is the agent of the passive verbal action of *referred to*. In the second line, *referred to* is the verb in the relative clause "which Mr Tam referred to". In this case, *referred to* is used in the active structure. Thus, in both lines 1 and 2, "Mr Tam" is the agent of the action of the phrasal verb in spite of appearing on different sides of *referred to* respectively in the two instances. This suggests the semantic preference of 'people'. These two lines have the semantic prosody of 'making clear references'. This semantic prosody is also observed in *refer to* in this sub-corpus. Although the other categories of co-selections of *referred to* are different from those of *refer to* in this sub-corpus, these two phrasal verbs share the same semantic prosody. In these two cases, it is the addition of the reference about who mentioned "the public opinion survey" as in line 1 and "heavy-gauge plastic bin liners" as in line 2, which makes the information more precise and clear.

Although this is the major semantic prosody found in the sub-corpus with the highest percentage of occurrence of *referred to*, this semantic prosody is only found in three sub-corpora and it is not prominent in the other two sub-corpora. As shown in Table 7, most of the sub-corpora are particularly associated with the meaning of 'locating where the information is mentioned' and 'identifying with conventionally used terms'. These two meanings are illustrated using Ordinances and Agreements respectively.

In Ordinances, for example, which is the sub-corpus with the second highest percentage of occurrence of *referred to*, there are 114 instances of *referred to* and 96.5% of them expresses the meaning of 'locating where the information is mentioned'. Figure 47 shows a sample of concordance lines of *referred to* in the sub-corpus.

```
1  ESG14 (6 98) - 5 5.3 The colour video signal referred to in this Specification shall generally confo
2    (2) This Specification (3) Specifications referred to in Clause 3.1 (4) International Standards 4
3  (a) immediately thereafter remove the notice referred to in regulation 32(1)(ii)(B) which relates to
4    6 alleging the contravention of a provision referred to in that regulation that the person charged-
5  rson who is under the supervision of a person referred to in subparagraph (i). (Part IIA added 22 of
6    and (ii) after the expiration of the period referred to in paragraph (i), in the 5 years period
7  ehicle and licence granted in respect thereof referred to in subregulation (1) or (2) until-  (a) the
8  at specified in the prescribed qualifications referred to in subparagraph (i). (Enacted 1990)   (1) A
9       Goods Ordinance (Cap 295) of the licence referred to in that subregulation granted in respect of
10   be granted or refused. (2) For the purpose referred to in subregulation (1), the Authority may
```

Figure 47: Sample concordance lines of *referred to* in Ordinances

The overall pattern of co-selections is quite neat. Regarding the colligation, there is always a noun phrase immediately to the left of the phrasal verb (107 times, 93.9%). To the right, the phrasal verb is followed by a prepositional phrase (112 times, 98.2%), except only in two lines. Thus the most frequent configuration is:

noun phrase + referred to + prepositional phrase

Looking more closely, I found that the preposition leading the prepositional phrase is predominantly "in" (108 times, 94.7%). This makes it a very strong collocate of *referred to* in this sub-corpus. Other collocates include "paragraph" (36 times, 31.6%),

"subregulation" (33 times, 28.9%), "regulation" (19 times, 16.7%) on the right side, and "gas installation" (16 times, 14.0%) and "disciplinary action" (15 times, 13.2%) on the left side.

In 93.9% of the lines (107 times), words like "paragraph" and "subregulation" are found, and there are either a numeral or an English letter, or both, representing the particular section or paragraph, for example, "Clause 3.1" and "paragraph (i)". These phrases denote particular section or part embedded in a text. Thus, *referred to* has the semantic preference of 'inter-text references'. Due to the use of the preposition 'in' which is used to express the location of something, *referred to* has the semantic prosody of 'locating'. The noun phrases immediately to the left of *referred to* represent some information or item mentioned earlier, and the use of *referred to* in these 93.9% of the lines is to locate where the information appears earlier in the texts.

Regarding the meaning of 'identifying with conventionally used terms', Agreements is used as an example for illustration as all the *referred to* instances in this sub-corpus have the same co-selection and are associated with this meaning. Figure 48 shows a sample of concordance lines of *referred to* in the sub-corpus.

```
1        Environment Forum in Sweden, which is referred to as the Malm? Declaration (UNEP, 2003): "We
2  r different emissions in an impact category is referred to as characterization in LCA. Normalization
3  ndicators of a reference case. This process is referred to as normalization in LCA. Typical reference
4      in LCA. Typical reference impact values, referred to as normalization factors, that are common
5  or all the impact categories in a LCA study is referred to as weighting. The function of the weighti
6   rate r, the real interest rate, more commonly referred to as the real discount rate, d, is to be us
7  that have the effect of inflation adjusted are referred to as in 'constant dollars'. In the above
8  the building. The former type of energy use is referred to as the 'embodied energy' of the materials.
9  unit quantity of the material are collectively referred to as the impact profile of the material. The
10   calculation routine utilizes a set of data, referred to as a shading mask (Figure 4.8), to
```

Figure 48: Sample concordance lines of *referred to* in Agreements

There are 23 lines of *referred to* in total. The colligation is rather neat. The phrasal verb is always immediately followed by a preposition. In only one of the lines, an adverb

drops between *referred to* and the following preposition. And then a noun phrase always comes after the preposition. To the left of *referred to*, there is an auxiliary verb in about half of the lines (12 times, 52.2%), a noun phrase in eight lines (34.8%), and an adverb in three lines (13.0%). In fact, in the three lines where an adverb occurs immediately to the left of *referred to*, the adverb follows a noun phrase. Therefore, the two most frequent configurations of *referred to* in this sub-corpus are:

noun phrase + auxiliary verb + referred to + preposition + noun phrase

noun phrase + (adverb) + referred to + preposition + noun phrase

It is found that the preposition is always realised by "as". This makes it a strong collocate of *referred to*. Apart from "as", collocates to the right include "localisation" (5 times, 21.7%), "study" (2 times, 8.7%), "normalization" (2 times, 8.7%), etc. To the left, the auxiliary verb is realised by different forms of the verb *to be*, i.e. "is" (8 times, 34.8%) and "are" (4 times, 17.4%). Other collocates on the left side of *referred to* include "process" (5 times, 21.7%), "study" (2 times, 8.7%), "hereafter" (2 times, 8.7%), etc.

The semantic preference of *referred to* is suggested to be 'conventionally used terminology' (100%). This preference is identified in all the lines and is realized by the co-occurrence of a short noun phrase to the right of *referred to*. The noun phrase is a specific term or expression which precisely characterises the noun phrase to the left. Repetition of words is found in seven lines (30.4%). This feature is also observed in, for example, *refer to* in Transaction Proceedings in expressing 'clarifying or defining engineering-related terminology'. However, *referred to* in these cases are used in the passive voice. The terminology appears on the right side of the phrasal verb, rather than to the left. In other words, rather than defining the terminology, the terminology occurring to

the right of *referred to* is used to characterise or identify the concepts introduced on the left side of *referred to* or even earlier in the text. This pattern together with the co-occurrence of "as" contributes to a strong semantic prosody of 'identifying with a term' (23 lines, 100%). This is also supported by the fact that the strong collocate "as" carries a sense of introducing the character or quality of something.

As for non-contiguous *referred to* instances, the overall frequency is very low. There are only five instances and they are distributed in three sub-corpora, namely Transaction Proceedings, Handbooks and Standards. All of the non-contiguous instances display constituency variation, i.e. *referred…to*. Four of the five instances are associated with the meaning of 'passing a serious matter for procedure or advice'.

In Transaction Proceedings, three lines of *referred…to* are found. Their concordance lines are extracted below:

This requires that exceptional cases of constraint violation be referred back to the agent owner, the client, or his representative (depending on the contractual arrangements), for further instruction such as constraint relaxation by Increasing the budget threshold.

This requires that exceptional cases of constraint violation be referred back to the agent owner …

The exact value of the unaxial compressional strength of the loss at Gissar is not known, but in view of the cementation developed in the matrix structure, the strength may be referred roughly to be within the range of qu = 200-400 kPa.

In all the three lines, an auxiliary verb immediately precedes *referred*. In the middle of *referred* and *to*, there is an adverb in all the three lines. Following *to*, there are noun

phrase (2 times, 66.7%) and auxiliary verb (1 time, 33.3%). The first two instances have the semantic prosody of 'serious'. This is suggested by the collocating phrase "exceptional cases of constraint violation" and also the fact that the cases are required to be passed back to the "agent owner, the client, or his representative" for "further instruction", as described in the wider co-text. In these two lines, *referred...to* has similar co-selections to that of *refer...to* in Guides. The intervening word "back" adds the sense of returning the cases to "the agent owner, the client, or his representative". Thus, in spite of the intervening word, *referred...to* in these two lines is used to express 'passing serious matter for further action'.

In line 3, *referred...to* is used to provide a clarification of information. This line resembles those of *refer to* in Transaction Proceedings as described earlier. The phrase to the right of the particle *to* is "within the range of gu = 200-400" which characterise the phrase "the strength" which occurs to the left of *referred*. So, *referred...to* is used to provide a clarification of the information to the left hand side, i.e. "the strength" in this case. Even though there is an intervening word "roughly", it does not affect the meaning of *referred...to*. It only adds the sense of 'hedging' to the meaning of *referred...to*, implying that "the strength" is approximately within the described range.

In Handbooks and Standards, there is a single instance of *referred...to* and they are extracted below:

...inform the Industry Liaison Coordinator and the person who has referred the case to the Branch Officer of the development (Handbooks)

...the conflict shall be referred by DEP to SETW for him to resolve the matter with other Policy Secretaries (Standards)

229

In these two instances, *referred…to* has the meaning of 'passing a serious matter to an authority for advice or action', which is specific to *refer…to*. The first line shares the semantic preference of 'requiring formal and authoritative advice or action' which is suggested by the intervening words "the case". Expansion of the concordance line shows that "the case" refers to the condition of having an event organised by a third party. This matter needs to be passed to "the Branch Officer" for further action. In fact, this instance in Handbooks shares the co-selections of those of *refer…to*, as illustrated above in Position Documents, but this instance is used in the present perfect tense.

As for the line in Standards, *referred…to* also has the same meaning. The phrase "the conflict" occurring to the left of *referred* is a specific matter which requires further action, and it has a sense of 'serious'. In this case, the intervening words "by DEP" (Director of Environmental Protection) represent the person who passed "the conflict" to "SETW" (Secretary for Environment, Transport and Works). Therefore, the intervening word does not affect the meaning of the phrasal verb, but adds more specification.

4.5.5 Summary

Overall, *refer to* and *referred to* are the two most frequent forms of REFER TO in engineering genres. In academic English, these two phrasal verbs are also ranked as the most frequently used constructions (Liu, 2012). Interestingly, unlike the other phrasal verb lemmas described above, the four forms of REFER TO have rather different co-selections.

The most frequent one *refer to* is very commonly used as a bare infinitive form instead of simple present tense. It is mostly used to express the meanings of 'recommending readers to inter-text source of information' and 'directly ask readers to go

to some source of information'. Apparently, the two uses are to refer readers to the source of information for more details. They differ particularly in terms of the attitudinal meaning, with one having a recommending and more polite tone, and the other having a more directive attitude. The second most frequent form *referred to* is used in the passive voice as a past participle, rather than the simple past tense form. Used very differently, *referred to* mainly expresses 'locating where the information is found in the text' when it is followed by a prepositional phrase indicating a particular section or paragraph of a text, and 'identifying terminology' when it is followed by the preposition "as" and a noun phrase which represents a conventionally used terminology in the engineering setting. As for *refers to*, it has a very consistent co-selection across engineering genres. Predominantly, it has the colligation of *noun phrase + refers to + noun phrase*. The noun phrase to the right of *refers to* is often longer and more complex in structure, and it features and characterises the more simple noun phrase occurring to the left of *refers to*. And very often, the noun phrases to the left of *refers to* are some engineering terminology. A very strong semantic prosody of 'defining and clarifying' is observed for *refers to*. *Referring to* which has the lowest frequency among the forms comprising the lemma is used as a present participle instead of the progressive form. It tends to express the meanings of 'illustrating information in the figure' and 'using as references' in engineering genres.

It can be concluded that the four inflected forms comprising REFER TO are four different phrasal verbs due to the distinctive patterns of frequency and co-selection. Each of them is specific to a particular use and meaning.

As for the non-contiguous form of these phrasal verbs, some of them are simply the phraseological variation as they differ only in the phraseological form but share the

same co-selection. For example, *refers…to* shares the meaning of 'defining or clarifying' which is highly specific to *refers to*; *referring…to* has the same meaning as *referring to* in expressing 'using as reference'. However, *refer…to* and *referred…to* have a distinctive meaning which is not associated with any of their contiguous counterparts. *Refer…to* and *referred…to* have the semantic preference of 'specific matter requiring formal and authoritative advice or action' and the semantic prosody of 'serious'. In the case of *refer…to*, this semantic preference is evident by the words or phrases occurring in between *refer* and *to*. As for *referred…to* which is mainly used in the passive voice, the phrases occurring in front of *referred* tend to suggest such semantic preference. Despite the limited frequencies of *refer…to* and *referred…to*, they obviously demonstrate distinctive co-selection. Therefore, it is argued that *refer…to* and *referred…to* have a status different from other non-contiguous instances of BASE ON, CARRY OUT and COMPLY WITH which are simply phraseological variations, and thus *refer…to* and *referred…to* can be regarded as unique phrasal verbs.

In phrasal verb dictionaries, '*refer to'* is listed as an entry comprising of all the four inflected forms. Meanings of the entry are described with limited information regarding its usage. For example, in *Macmillan Phrasal Verb Plus* (Rundell & Fox, 2005), the entry *'refer to'* is printed in red colour with a 3-star rating. Based on the guidelines of use of the dictionary, phrasal verbs with such a rating label are the most essential and common English phrasal verbs. Four meanings are described for this entry in the dictionary, namely (1) to mention, (2) to describe, (3) to send someone to a person or place for advice or help, and (4) to consult a source of information. Findings in the current study show that some meanings are more specific to a particular form. The explanation provided in the dictionaries lacks details in this regard. The only grammatical information

provided for this entry is for the third meaning which is marked as "often passive". While the other three meanings are described with contiguous pattern, i.e. "refer to sb/sth", this third meaning is the only one with both contiguous and non-contiguous patterns provided, i.e. "be referred to sb/sth" and "refer sb to sb/sth".

In *Collins Cobuild Dictionary of Phrasal Verbs* (Sinclair et al., 1989), six meanings of *'refer to'* are provided, namely (1) to talk about or mention, (2) to describe something with a name or expression, (3) to describe, (4) to consult a source of information, (5) suggest somebody a source of information, and (6) to ask a specialist or an organisation to deal with a person or a problem. Grammar notes are provided in an extra column. Next to the explanation of the first four meanings, the grammatical information is "V+PREP: HAS PASSIVE". This indicates that this phrasal verb construction is formed by an intranstive verb followed by a prepositional particle, and that the construction "can have a passive, with the object of the preposition becoming the subject of the passive verb" (Sinclair et al., 1989: xvi). Findings of the current study show that *refer to* and the passive form *referred to* do not always share the same co-selections and meaning. For the second meaning, "USUALLY + as" is also stated in addition to the above grammar note, indicating that "as" is a frequent collocate of the phrasal verb in expressing this particular meaning. Our findings conform to this description, specifically for *referred to.* For the last two meanings, the grammar notes are "V+PRON+PREP, V+N+PREP" and "V+N+PREP, V+PRON+PREP" respectively. Such grammatical information suggests that the two elements of the phrasal verb are used non-contiguously in expressing the two particular meanings. These two meanings are found to be specific to *refer…to* and *referred…to* in the current study.

4.6 Use of phrasal verbs: SET OUT

This section analyses and discusses the fifth most frequent phrasal verb in the HKEC and its inflectional counterparts. There are only three inflectional forms of SET OUT, namely *set out, sets out* and *setting out*.

4.6.1 *Set out*

The fifth most frequent phrasal verb found across the engineering sub-corpora is *set out*. This phrasal verb is an irregular verb. It has the same form for its base form, the past tense and past participle. As different from the other sets of phrasal verbs discussed above, the analysis of *set out* requires even more attention in order to determine how many instances are used in the infinitive form, the present tense form, the past tense form or the past participle form. There are 1,288 instances of *set out* in the whole corpus. They are spread across 25 sub-corpora. Table 26 shows the sub-corpora in which *set out* is found and the respective frequency and percentage of occurrence in each sub-corpus. There are four instances of *set out* in the sub-corpus of Plans. Due to the small size of the sub-corpus, it has the highest percentage of occurrence (0.0959%) of *set out* even with only four instances.

Table 26: Sub-corpora in which *set out* is found

	Sub-corpora	size	Freq	%
1	Plans	4,173	4	0.0959%
2	Circular Letters	143,313	121	0.0844%
3	Consultation Papers	111,494	49	0.0439%
4	Standards	136,024	46	0.0338%
5	Handbooks	67,284	21	0.0312%

6	Review Papers	106,506	31	0.0291%
7	Guides	783,805	208	0.0265%
8	Abstract	94,671	25	0.0264%
9	About Us	647,013	155	0.0240%
10	Ordinances	139,176	27	0.0194%
11	Frequently Asked Questions	55,726	9	0.0162%
12	Publicity Material	599,407	83	0.0138%
13	Code of Practice	997,228	136	0.0136%
14	Notes	156,255	21	0.0134%
15	Reports	979,170	122	0.0125%
16	Media Releases	1,566,742	166	0.0106%
17	Manuals	296,299	15	0.0051%
18	Conference Proceedings	196,498	9	0.0046%
19	Product Descriptions	611,549	26	0.0043%
20	Position Documents	75,660	3	0.0040%
21	Q & A	27,703	1	0.0036%
22	Technical Papers	65,731	2	0.0030%
23	Transaction Notes (HKIE)	79,058	2	0.0025%
24	Agreements	127,895	1	0.0008%
25	Transaction Proceedings (HKIE)	1,055,248	5	0.0005%

Figure 49 shows the four concordance lines of *set out* in Plans. Regarding the word class of the phrasal verb, *set out* in lines 1, 3 and 4 are the base form used as infinitive, whereas the one in line 2 is a past participle form. To the left of *set out*, the colligation of a verb followed by an infinitive marker is observed in lines 1 and 3. In line 2, the preceding item is a conjunction and in line 4 a modal verb. On the right side, there are two instances of a prepositional phrase, and two of a noun phrase. For collocation, there are "to" (3 times), "the" (3 times), "propose" (2 times) and "we" (2 times) on the left side of *set out*. On the other side, there are "of" (2 times) and "the" (2 times).

```
1      allowed for the power sector, we propose to set out by means of a Technical Memorandum (TM) und
2   to amend the Air Pollution Control Ordinance as set out in paragraphs 8 to 19 of this paper.
3  waste management legislation, we now propose to set out the essential elements of individual PRSs in
4    that the tentative body of the PER Bill would set out the purpose of the legislation; the types of
```

Figure 49: Concordance lines of *set out* in Plans

The semantic preference of *set out* in this sub-corpus is 'legal information related to pollution control' (100%). This semantic preference is evident by "emission allowances for the power sector" in line 1, "the Air Pollution Control Ordinance" in line 2, "the essential elements of individual PRSs", in which "PRS" refers to "Producer Responsibility Scheme", in line 3, and "the purpose of the legislation" in line 4. These phrases represent legally-related information on the control and restriction of pollution. Regarding the semantic prosody, the words "propose" in lines 1 and 3, and "would" in line 4 carry a sense of intention. Thus, there is the semantic prosody of 'intending to provide details' (75%). The word "propose" conveys a meaning of putting forward a plan to do something, and the word "would" is a modal verb marking volition (Biber et al., 1999). They indicate that the information has not been elaborated in details and they now offer to do so. Therefore, the phrasal verb *set out* is used to express the meaning of intending to provide details on legal information related to pollution control. The fact that the high percentage of occurrence of this semantic prosody (75%) is contributed by only three instances may hedge the generalization of the findings. Nonetheless, the findings are indeed consistent with the nature of the texts which are plans and proposals.

The phrasal verb in line 2, which is the only instance in this sub-corpus used in past participle, does not to share the same meaning as the other three instances. This instance also shares the semantic preference of 'legal information related to pollution control' which is realized by "the Air Pollution Control Ordinance". But for the semantic prosody, the instance has a different one. The phrasal verb is followed by a prepositional

phrase "in paragraphs 8 to 19 of this paper" which seems to provide a specification of the information in terms of where the information is detailed. To the contrary to the other three instances which denote a sense of intending to provide information that has not been stated, this instance identify the specific location of the information denoting that the information has been provided already. This single instance may not provide enough data for any description of patterns, but it is observed that this instance resembles the majority of the instances in other sub-corpora. More detailed analysis of their co-selections will be discussed in the following paragraphs using other sub-corpora.

The sub-corpus with the second highest percentage of occurrence of *set out* is Circular Letters. There are 121 instances (0.0844%) in the sub-corpus. Figure 50 shows a sample of the concordance lines. In this sub-corpus, no instance is used in the base form. One instance is used in the past tense active voice, and all the other instances are the past participle as they are used in the passive voice. In Figure 50, *set out* is used in passive voice in all the ten lines.

```
1       which commercial terms have been agreed and set out in the Agreement. Pursuant to the  Operation
2       the following expressions have the meanings set out below  unless the context otherwise requires.
3      and, if thought fit, passing the resolutions set out in the Notice convening the Meeting and at
4       of Directors of Yau Lee Holdings Limited is set out on pages 1 to 7 of this circular. A notice
5 Kowloon Bay, Kowloon, Hong Kong at 10:00 a.m. is set out on pages 8 to 11 of this circular. If you
6             details of the retiring Directors are set out below. Mr. Sun Chun Wai, aged 47, earned a
7   OF THE SHARE ISSUE MANDATE Resolution No. 5(1) set out in the notice of AGM (the "AGM Notice") da
8      ("Shares"). Subject to certain exceptions as set out in the AGM Notice, the additional shares
9        Period (as defined in Resolution No. 5(1) set out in the AGM Notice). EXERCISE OF THE
10REPURCHASE MANDATE Resolutions No. 5(2) and 5(3) set out in the AGM Notice will, if passed, give a
```

Figure 50: Sample concordance lines of *set out* in Circular Letters

The phrasal verb often follows a noun phrase (53 times, 43.8%) or an auxiliary verb (47 times, 38.8%). The auxiliary verb is mainly realized by "is" (28 times) and "are" (18 times). In some cases, the pattern preceding *set out* is a noun followed by a conjunction (16 times, 13.2%). The conjunction is always "as". An example is line 8. To the right of the phrasal verb, a prepositional phrase occurs in the majority of the lines (99

times, 81.8%). In the remaining lines, the phrasal verb is followed by an adverb (22 times, 18.2%). Thus, the most frequent configurations of *set out* in this sub-corpus are:

noun phrase + set out + prepositional phrase
auxiliary verb + set out + prepositional phrase

In both configurations, *set out* is used in the past participle form in the passive voice. This means that in the first configuration, *set out* plus its following prepositional phrase functions as an adjective phrase of the noun phrase. In other words, the whole sequence *noun phrase + set out + prepositional phrase* forms a larger noun phrase.

The top collocates to the left of *set out* include "the" (71 times), "of" (32 times), "is" (31 times), "as" (19 times), "are" (18 times), "in" (16 times), "information" (16 times), "resolution" (15 times), "mandate" (10 times), "period" (10 times), "resolutions" (9 times), "financial" (8 times), etc. On the right side, there are "in" (85 times), "the" (85 times), "of" (41 times), "to" (40 times), "this" (39 times), "circular" (36 times), "notice" (35 times), "AGM" (27 times), "below" (24 times), "on" (24 times), "pages" (23 times), "note" (14 times), "appendix" (13 times), "convening" (9 times), "meeting"(9 times), etc. In 117 instances (96.7%), *set out* co-occurs with words or phrases denoting information related to the operation, policies, decisions and financial performance of a company, for example, "resolutions" as in line 3, "details of the retiring Directors" as in line 6, "the AGM notice", "a letter from the Board", "the financial information", "the accounting policy for contract revenue", "their principal activities", and "proposed amendments to the Company's Bye-laws". Such kinds of information generally reflects the performance and condition of a company, and thus, is either required by the government to be disclosed to the public or is of interest to the shareholders or potential investors of the

companies. These co-occurring phrases create a semantic preference of 'company disclosures'.

A very strong semantic prosody is observed across the instances of *set out* in this sub-corpus. The prepositional phrases or the adverb following *set out* in all the instances denote a location or direction as to where the details of company disclosures can be found. The heads of the prepositional phrases are "in" (74 times, 61.2%), "on" (23 times, 19.0%) and "under" (a single occurrence). In these cases, the preposition "in" is used to show something within a particular space or area, or at a specific part of something; "On" denotes where something has been written; and "under" is used to express something can be found in a particular section or below a particular title. For example, in line 3, "the resolutions" itself may be too generic. The prepositional phrase "in the Notice convening the Meeting" functions to direct the location of the details of "the resolutions". In a sense, this co-selection of *set out* with the prepositional phrase specifies what particular resolutions are being referred to, i.e. those described and detailed in the Notice. In addition, the adverbs used with *set out* in the sub-corpus are "below" (21 times, 17.4%) and "above" (a single occurrence). These two adverbs are respectively used to show something in an earlier part or an upper position, and in a later section or a lower position. For example, in line 6, the adverb "below" is used immediately to the right of *set out*. It indicates that the company disclosures, i.e. "Brief biographical details of the retiring directors" in this instance, are provided in a later section of the document. Therefore, *set out* has a semantic prosody of 'directing the location' (100%) in Circular Letters. It is noticed that the only instance in Plans which does not have the semantic prosody of 'intending to provide details', i.e. line 2 in Figure 49, share this semantic prosody of

'directing the location'. The phrasal verb *set out* in that instance is followed by a prepositional phrase headed by "in" and the phrase "paragraphs 8-19 of this paper".

In fact, this semantic prosody is found in all the sub-corpora in which *set out* occurs. Table 27 shows the percentage of occurrence of this semantic prosody in each of the sub-corpora. Out of the 25 sub-corpora, 21 of them (84%) have a percentage of over 50% of the semantic prosody, and 16 of which have even got a percentage of over 90%. Those instances which do not share this semantic prosody tend to have other meanings.

Table 27: Percentage of occurrence of the semantic prosody of 'directing the location'

	Sub-corpora	size	Freq. of set out	%	Freq. of the semantic prosody 'Directing'	%
1	Plans	4,173	4	0.0959%	1	25%
2	Circular Letters	143,313	121	0.0844%	121	100%
3	Consultation Papers	111,494	49	0.0439%	45	91.8%
4	Standards	136,024	46	0.0338%	44	95.7%
5	Handbooks	67,284	21	0.0312%	21	100%
6	Review Papers	106,506	31	0.0291%	29	93.5%
7	Guides	783,805	208	0.0265%	201	96.6%
8	Abstract	94,671	25	0.0264%	25	100%
9	About Us	647,013	155	0.0240%	149	96.1%
10	Ordinances	139,176	27	0.0194%	22	81.5%
11	Frequently Asked Questions	55,726	9	0.0162%	5	55.6%
12	Publicity Material	599,407	83	0.0138%	81	97.6%
13	Code of Practice	997,228	136	0.0136%	129	94.9%
14	Notes	156,255	21	0.0134%	19	90.5%
15	Reports	979,170	122	0.0125%	112	91.8%
16	Media Releases	1,566,742	166	0.0106%	142	85.5%
17	Manuals	296,299	15	0.0051%	13	86.7%
18	Conference Proceedings	196,498	9	0.0046%	7	77.8%
19	Product Descriptions	611,549	26	0.0043%	24	92.3%
20	Position Documents	75,660	3	0.0040%	1	33.3%
21	Q & A	27,703	1	0.0036%	1	100%

22	Technical Papers	65,731	2	0.0030%	1	50.0%
23	Transaction Notes (HKIE)	79,058	2	0.0025%	2	100%
24	Agreements	127,895	1	0.0008%	1	100%
25	Transaction Proceedings (HKIE)	1,055,248	5	0.0005%	1	20.0%

In Consultation Papers, there are 49 instances of *set out* (0.0439%). Figure 51 shows a sample of concordance lines.

```
1     PROPOSAL The implications of the proposal are set out at Annex B.  PUBLIC CONSULTATION We will
2        account of all the relevant considerations set out in the preceding sections, we propose that -
3       penalty ticket unless any of the exemptions set out in paragraph 6.2 applies; and - the ban be
4   views on the proposed implementation framework set out in sections 6.1-6.2 above. To cater for the
5         A list of the key consultation points is set out at Annex D. Please send in your comments to
6   views on the proposed implementation framework set out in sections 6.1-6.2 above. [Sections 6.1-6.2]
7  on the proposed ban. Details of the proposal are set out in the Legislative Council Brief issued to
8         districts; and - Other than the exemptions set out in the consultation document, no further
9                    caps per vehicle, which are set out at Annex. These caps have been set with
10     Projects ~Project proponents are required to set out targets deliverables in their funding
```

Figure 51: Sample concordance lines of *set out* in Consultation Papers

In this sub-corpus, five of the instances are used as the base form in the active voice (10.2%). One instance is used in the past tense active voice (2.0%). The others are used as the past participle in the passive voice (87.8%). Regarding the colligation, the phrasal verb is preceded by either an auxiliary verb (20 times, 40.8%) or a noun phrase (19 times, 38.8%) in most of the lines. The auxiliary verb is mostly "are" (12 times) and "is" (6 times). In some cases, there is a conjunction "as" (4 times, 8.2%) or an infinitive marker "to" (4 times, 8.2%). To the right of the phrasal verb, there is mostly a prepositional phrase (39 times, 79.6%) which is headed by "in" (26 times) and "at" (13 times). In five instances, *set out* is followed by the adverb "below" (10.2%). In the remaining five instances which are used in the active voice, the item following *set out* is a noun phrase (10.2%). Thus, the two most frequent configurations in Consultation Papers are:

auxiliary verb + set out + prepositional phrase

noun phrase + set out + prepositional phrase

They are the same as those in Circular Letters. Again, these configurations are conformed to only when *set out* is used in the past participle form in the passive voice. When used in the base form in the active voice, *set out* has the following configuration:

infinitive marker + set out + noun phrase

The top collocates to the left of *set out* include "the" (26 times), "of" (16 times), "are" (14 times), "and" (8 times), "to" (7 times), "is" (6 times), "buildings" (5 times), "on" (5 times), "as" (4 times), "those" (4 times), "a" (3 times), "areas" (3 times), "for" (3 times), "framework" (3 times), "implementation", etc. To the right, top collocates include "in" (29 times), "the" (26 times), "at" (14 times), "annex" (13 times), "to" (10 times), "a" (7 times), "below" (6 times), "and" (4 times), "paper" (4 times), "above" (3 times), "building" (3 times), "of" (3 times), "sections" (3 times), etc. Based on the words or phrases co-occurring with *set out*, the phrasal verb tends to have three semantic preferences in this sub-corpus, namely 'proposal-related matters' (26 times, 53.1%), 'legislation and regulations' (15 times, 30.6%) and 'data and findings' (8 times, 16.3%). The semantic preference of 'proposal-related matters' is evident from phrases such as "The implications of the proposal" in line 1 and "the proposed implementation framework" in lines 4 and 6. In some instances, the phrases may not be explicitly denoting a proposal or proposal-related matters, but the extended co-texts provide that they are. For example, "the exemptions" in line 3 refers to the items or conditions which are free from or not regulated under the proposed ban on idling vehicles with running engines. Also in line 10, "targets/deliverables" (checking the original text data shows that there is a slash between the two words) are the required information project proponents need to provide in their

application for funding. Since the projects are subject to approval before funding is granted for the carrying out of the projects, they can be considered as a kind of proposal.

The other semantic preference 'legislation and regulations' is suggested by the co-occurrence of phrases such as "the guiding principles and criteria", "the vetting criteria", "technical requirements", "SC" which stands for sewage charge and "the regulatory regime under existing legislation". There are also some instances indirectly relating to regulations or legislation, but their referents become clear when longer co-texts are considered. For example, in one of the concordance lines "These are set out below" the pronoun "these" actually refers to "further measures to combat climate change" and the measures more specifically include "Mandatory Implementation of the Building Energy Codes" and "Mandatory Energy Efficiency Labelling Scheme". In four instances, there are the phrases "the developers of those new buildings" or "those buildings" as in "The developers of those new buildings set out at Annex C are required to submit a self-declaration…" or "…in respect of the new floor areas of those buildings set out at Annex C". The co-texts show that "the developers of those new buildings" are regulated entities specified in the stipulated standards for building energy efficiency. The last semantic preference 'data and findings' denotes more general information or data. Examples are "the latest information of these schemes", "the key initiatives undertaken since 2004", "The historical trend of GHG emissions in Hong Kong from 1990 to 2005", and "the key findings".

There seems to be three distinct semantic preferences of *set out* in this sub-corpus, but the three preferences are indeed closely related. This sub-corpus comprises of texts of consultation papers. They are to detail the proposed change or new items and to seek support and opinion from the public or relevant committee members about the proposal

before it can be revised and approved. Thus there is a strong preference for *set out* to co-occur with words and phrases that describe proposals or any related matters. Some of the proposed matters involve legislative or legal issues, for example, the proposed ban on idling vehicles with running engines, the proposed tax concession to promote the use of environmental friendly vehicles, and the proposal to amend the Air Pollution Control Regulations regarding fuel restriction. In order to provide detailed background and solid arguments to support the proposals, researches, pilot studies or tests were conducted to investigate current conditions and possible impacts (for example, economical implications and environmental impacts) of the proposed change. Therefore, the three semantic preferences are all related to the main functions of the sub-corpus of Consultation Papers.

The semantic prosody of 'directing the location' is observed in 46 instances (93.9%). Lines 1 to 9 in Figure 51 have this semantic prosody as evident by the co-occurrence of the prepositional phrase following *set out*. As mentioned earlier, *set out* co-selecting a prepositional phrase typically headed by "in" and followed by a noun phrase denoting a text document or part of a text document expresses that some information are discussed within a particular space or area, or at a specific part of a document. For example, in line 2, readers are directed to "the preceding sections" for details of "all the relevant considerations". In line 3, more specific location of "the exemptions" is indicated by the prepositional phrase, i.e. "in paragraph 6.2". While in Circular Letters the prepositional phrases are mostly headed by "in" and "on", the latter preposition is not found in this sub-corpus. The other preposition used is "at" which also functions to indicate a particular place or location. For example, in line 1, the prepositional phrase "at Annex B" shows where you can find the information "the implications of the proposal". It is noticed that "at" co-selects very frequently with "Annex" as shown in lines 1, 5 and 9.

In general, *set out* and the prepositional phrases, whether it is headed by "in" or "at", constitute the semantic prosody and share the same meaning of directing the location of information. The use of the adverb "below" right after *set out* also creates the same semantic prosody of 'directing the location'.

Three instances seem to have a different semantic prosody. Line 10 in Figure 51 is one of them. This instance has different colligation. It is used in the active voice. The preceding phrase "required to" seems to suggest a sense of 'requiring' indicating that it is a requirement to provide the information "targets/deliverables". It also implies that, unlike those instances used in the passive voice and with a prepositional phrase which suggest that the information is already somewhere within the same document or at another piece of text, the information in this instance is not provided initially and now it is required to be stated out.

Another instance not sharing the semantic prosody is considered to have the semantic prosody of 'intending to provide details'. The instance is extracted as follows:

To implement the fee increment proposal, we propose to set out the SC rates in the coming ten years in a single item of subsidiary legislation ...

The co-occurring phrase "propose to" also suggests that the information "the SC rates in the coming ten years" has not been provided initially, and they are now putting forward a plan or a proposal to detail this information. Although there is only this single instance in this sub-corpus, the co-selections of this instance resemble those in the sub-corpus of Plans as described above.

The last instance seems unclassifiable. The extract of the co-text of the instance is given below:

Environmental Protection Department

...

- Set out technical requirements for the environmental impact assessment (EIA) processes at the planning stage, to avoid, minimize and control potential adverse impacts to the local environment of designated projects

The texts are from a table showing the roles and responsibilities of relevant bureaux and departments in environmental and human health protection. Under each of the bureaux or departments, there are bullet points listing the roles and responsibilities of the departments, and the legislative instruments they adopt to perform their duties. Thus, *set out* in this instance is used in the base form in the active voice. The whole clause with *set out* above is describing one of the responsibilities of the Environmental Protection Department. While this instance does not share the same meaning with all the other instances in the sub-corpus, this single occurrence provides insufficient data for a comprehensive description of it.

In Standards, there are 46 instances of *set out* (0.0338%). Three of which is used in the base form (6.5%), one in the past tense form (2.2%) and all the others are the past participle used in the passive voice (91.3%). Figure 52 shows a sample of concordance lines of the phrasal verb in this sub-corpus.

```
1  since been reviewed further and the details are set out in this Circular. 7. This Circular covers
2   criteria, and the submission requirements are set out in Appendix B. DEP may waive the sediment
3 Assessment The purpose of these guidelines is to set out the requirements for assessing, sampling,
4  or discrepancy between this Circular and those set out above , this Circular  shall prevail. Policy
5     activities comply with the requirements as set out in the Particular Specification. (e)
6 to reinstate the unauthorized disposal ground as set out  in Clause 13 above, the Employer may in
7  under PFSES, the standard terms and clauses as set out in paragraph 15 below shall be incorporated
8 above, certain minimum environmental measures as set out in Appendix F shall still be required.
9 the PS other relevant requirements conditions as set out in the Environmental Permit); and ETWB TCW
10 As mentioned in para. 15(d), the MoM and BQ SoR set out in Appendices D and E respectively cover
```

Figure 52: Sample concordance lines of *set out* in Standards

For the patterns of colligation on the left side, there are mostly noun phrase (22 times, 47.8%), conjunction (14 times, 30.4%) which is always the word "as", and an auxiliary verb (6 times, 13.0%). To the right of *set out*, a prepositional phrase is found in most of the instances (37 times, 80.4%). All of these prepositional phrases are headed by "in". There are five instances of adverb (10.9%), and three instances of a noun phrase (6.5%). Thus the two most frequent configurations of *set out* in Standards are:

noun phrase + set out + prepositional phrase

noun phrase + conjunction + set out + prepositional phrase

The configuration *noun phrase + set out + prepositional phrase* is commonly used in this sub-corpus and in the other sub-corpora discussed above, with *set out* and its following prepositional phrase functioning as the adjective phrase modifying the preceding noun phrase. The other configuration *noun phrase + conjunction + set out + prepositional phrase* which is not very commonly found in the previously discussed sub-corpora but is fairly frequent in Standards. The configuration of *infinitive marker + set out + noun phrase* is also observed in two of the instances in which *set out* is the base form.

Top collocates to the left of *set out* include "the" (32 times), "as" (14 times), "and" (12 times), "requirements" (9 times), "of" (8 times), "environmental" (6 times), "procedures" (6 times), "to" (6 times), "are" (4 times), "guidelines" (4 times), "measures" (4 times), "is" (3 times), "mitigation" (3 times), etc. On the other side, there are "in" (42

247

times), "the" (26 times), "of" (13 times), "and" (11 times), "to" (9 times), "for" (8 times), "appendix" (7 times), "a" (5 times), "as" (5 times), "below" (5 times), "circular" (5 times), "EIA" (5 times), "above" (4 times), "environmental" (4 times), etc. Most of the phrases co-occurring with *set out* in Standards are related to guidelines or requirements, for example, "the submission requirements" in line 2, "the requirements for assessing, sampling, testing and categorizing the sediment" in line 3, "the standard terms and clauses" in line 7, "certain minimum environmental measures" in line 8, "the guidelines and procedures", "the specified grade strength", etc. Thus, *set out* has a semantic preference of 'guidelines and requirements' (39 times, 84.8%) in Standards.

The semantic prosody of 'directing the location' is observed in 44 instances (95.7%). Most of these instances are realized by the co-occurrence of a prepositional phrase or an adverb following the phrasal verb, such as "in this Circular" as in line 1 and "above" as in line 4. All the prepositional phrase are headed by "in" (37 times), and the choice of adverb includes "below" (3 times), "above" (1 time) and "herein" (1 time). The adverb "herein" is often used in more official and legal documents expressing something in or within the document. Thus, *set out* used with "herein" also has the semantic prosody, directing the readers where the information, i.e. guidelines and requirements in this case, is located. Similar to the cases of "above" and "below", "herein" only suggests a general location, which is within the document. The use of these adverbs does not specify locations in a way as specific as in the case of certain prepositional phrases, for example, "in paragraph 15".

There are three instances which have the same semantic prosody but they do not share the same pattern. Line 3 in Figure 52 is one of them. The phrasal verb in this

instance is used in the active voice in the base form. The line beginning with "The purpose of these guidelines" suggests that "set out the requirements for assessing…" is the purpose. This line may appear irrelevant to the semantic prosody of 'directing the location'. However, it actually means that "the requirements for assessing…" are provided in details in "these guidelines". In other words, this instance can be considered as a paraphrase, and it conforms to the most frequent configuration with some turbulence. And it shares the same semantic prosody as it is to direct to the location of the information, i.e. "the requirements" are detailed in "these guidelines. The other two instances are:

WBTC No. 18/98 set out the procedures for EIA of development projects and proposals…

The purpose of these guidelines is to set out the requirements for assessing, sampling, testing and categorising the sediment.

For the first instance, it is used in the active voice and expresses that "the procedures…" are described in details in "WBTC No. 18/98". In other words, it has the typical semantic prosody because it functions to direct to the location of where the information is described. That means "WBTC No. 18/98" is the text where readers can find the details of "the procedures for EIA…". Thus, this instance adheres to the semantic prosody of 'directing the location'. The second instance has the same structure and condition as line 3 in Figure 52.

The only two instances that are not classified as having this semantic prosody are:

… [EI] section should set out in concise fashion whether or not the proposal has environmental implications …

FIFA has set out to ensure that the balls used in top matches meet the most exacting standards.

Reading the larger co-texts reveals that "EI" in the first instance refers to "Environmental Implications". The phrasal verb *set out* in this instance seems to have a different semantic preference. The co-occurring phrase "whether or not the proposal has environmental implications" is not related to guidelines and requirements. Rather, it seems to be about the effects of particular actions on the environment. The modal verb "should" suggests that such information about the environmental implications is not provided initially, and it is now required to be stated "in concise fashion". This instance tends to carry a sense of requirement which is similar to one of the instances described above in Consultation Papers. For the final instance, a non-finite clause is used immediately to the right of *set out*. This is unlike in all other instances in which *set out* is followed by a prepositional phrase, an adverb or a noun phrase. This instance has a different semantic preference because its co-occurring words are not about guidelines or requirements, or any kind of information. Based on the co-texts, "FIFA" refers to Federation Internationale de Football Association which is a governing body for the organisation and related matters of major international football tournaments. The non-finite clause seems to be introducing a goal or an aim which FIFA wants to achieve. This single instance does not provide enough data for a concrete description of its pattern and meaning, but the disparate co-selections of this instance suggest a meaning distinct from those in other instances.

There are 21 instances of *set out* in Handbooks (0.0312%). Table 25 shows that the percentage of occurrence of the semantic prosody of 'directing the location' is 100%. That means all the instances in Handbooks conform to the canonical form and meaning of

set out described above. In Review Papers, where there are 31 instances of *set out* (0.0291%), 93.5% of the instances (29 times) have the typical semantic prosody. The only two instances that do not share the same meaning seem to have the sense of intending to provide details and requiring respectively.

In Guides, there are 208 instances (0.0265%) of the phrasal verb. 21 of them are used in the base form (10.1%), the remaining ones are used as the past participle in the passive voice (89.9%). Figure 53 shows a sample of concordance lines in the sub-corpus.

```
1     and considerations suggested in this section set out principles for maintenance and renovation
2  not less than 6% v v oxygen at a temperature as set out in 4.2.1 (c) (ii) for a residence time of
3   requirements of the Regulation. The Notes also set out lists of standards that are deemed to sati
4       associated with their operation. These are set out below for each type of laser. (a)* Carbon
5   Class 4 lasers and the control - 20 - measures set out in paragraph 2.4.4 above should be applied.
6   of medical devices and the scope of control as set out in the consultation document.  Following the
7  & Mechanical Services Department (EMSD) aims to set out the minimum design requirements on energy
8  and Mechanical Services Department (EMSD), is to set out the minimum energy-efficient design standard
9   & Mechanical Services Department (EMSD) aims to set out the minimum design requirements on energy
10 nd Mechanical Services Department (EMSD), is to set out the minimum energy-efficient design standard
```
Figure 53: Sample concordance lines of *set out* in Guides

The co-selections of the majority of the instances resemble the typical co-selections as described above. The most commonly found grammatical class of items on the left side of *set out* is a noun phrase (136 times, 65.4%), a conjunction (26 times, 12.5%), an auxiliary verb (25 times, 12.0%) and an infinitive marker (11 times, 5.3%). The conjunction includes "as" (22 times) and "and" (4 times). To the right of the phrasal verb, there are a prepositional phrase (165 times, 79.3%), a noun phrase (24 times, 11.5%), an adverb (15 times, 7.2%), an infinitive marker followed by a verb (2 times, 1.0%), and a conjunction (2 times, 1.0%). Thus, the most frequent configurations of *set out* in Guides are:

noun phrase + set out + prepositional phrase

conjunction + set out + prepositional phrase

auxiliary verb + set out + prepositional phrase

infinitive marker + set out + prepositional phrase

In the first configuration, the phrasal verb is used in the past participle functioning as an adjective phrase modifying the preceding noun phrase. In the second sequence, there is a noun phrase preceding the conjunction when it is "as". In the other cases when the conjunction is "and", the clause with *set out* is connected to the previous clause. The phrasal verb is used in the base form when it follows the infinitive marker "to", and it takes a noun phrase as its object.

Regarding the collocation, top collocates to the left of *set out* are "the" (155 times), "to" (85 times), "requirements" (46 times), "emission" (42 times), "limits" (36 times), "as" (32 times), "and" (27 times), "a" (25 times), "in" (25 times), "meet" (25 times), "of" (25 times), "are" (20 times), "be" (17 times), "limit" (16 times), "satisfy" (16 times), etc. On the other side of *set out*, top collocates include "in" (169 times), "of" (67 times), "the" (65 times), "this" (64 times), "note" (55 times), "section" (47 times), "for" (29 times), "and" (27 times), "a" (21 times), "annex" (20 times), "paragraph" (20 times), "to" (20 times), "above" (18 times), etc. 185 instances (88.9%) co-occurs with phrases which are related to guidelines or requirements, for example, "list of standards" in line 3, "the minimum design requirements" in line 7, "the emission limits", "the concentration limits", "the specifications", "the tightest IAQ objectives", "noise standards", "the criteria", "the guidelines", etc. They have the semantic preference of 'guidelines and requirements'.

200 instances (96.2%) of *set out* in Guides conform to the semantic prosody of 'directing the location'. They are characterised by the use of a prepositional phrase or an adverb which function to specify the location of the mentioned guidelines or requirements.

The remaining eight instances which do not share this semantic prosody seem to have other meanings. For example, consider the following two instances:

a written notice of appeal which should be substantially in accordance with Form 1 in the Schedule (reproduced in the Appendix) and should set out the grounds of the appeal

Item 10: Waste types
Set out the major types of chemical waste generated or expected to be generated at the establishment(s).

The phrasal verb in both of the instances is used in the base form. The modal verb "should" in the first instance indicates that "the grounds of the appeal" is required information to be provided in details in the "written notice of appeal". In the second instance, the wider co-texts suggest that this instance is about information to be provided by chemical waste producer for registration before engaging in any activity that generates chemical waste. The phrasal verb *set out* in this case is used as an imperative instructing the readers about what information is required to be detailed under the heading of "Item 10: Waste types". Thus, there is a sense of requirement in these two instances.

Another meaning is observed in the other three instances. They have the identical concordance lines:

They can be discarded if the respective methods mentioned above do not set out any procedures for determination.

The phrase "any procedures for determination" denotes a set of steps to be followed for making certain decision, and in the context the decision is to measure the volume of water and exempt compounds of paints. Thus, it also shares the semantic features of guidelines

and conforms to the typical semantic preference of 'guidelines and requirements'. The word "if" introduces a condition and the phrase "do not" used right in front of *set out* indicates not doing. And the clause "they can be discarded" is the consequence in the presence of the condition. Therefore, *set out* in these three instances seems to have the meaning of describing the consequence of not giving details of the guidelines and requirements.

Another unclassifiable instance is the following:

Key Environmental Outcomes or Influences

...

2. Set out a Strategic Environmental Monitoring and Audit framework to check the environmental performance and ensure follow-up actions are taken.

The phrasal verb is in the past tense form and is used in one of the bullet points under the heading "Key Environmental Outcomes or Influences". It takes a noun phrase object "a Strategic Environmental Monitoring and Audit framework" which is followed by a non-finite clause. The noun phrase denotes the essential structure and guidance for monitoring and auditing the environmental influences. It also shares the semantic characteristics of 'guidelines and requirements'. The non-finite clause following the noun phrase seems to introduce the aim of setting out the framework.

The remaining two instances tend to have different semantic preference and semantic prosody. The instances are extracted as follows:

On arrival at site the drilling rig should be set-up up-wind of the borehole location, 'No smoking' signs set out and the working area should be roped or coned-off.

… except that the appropriate road hazard warning lanterns must be set out at the same time as the signs and cones …

The co-occurring words "signs" and "lanterns" respectively in the two instances tend to suggest a different semantic preference. They are notices on public display that gives information, instructions or signals in a written or symbolic form, unlike in the majority instances in which *set out* co-occurs with information of any kind on any documents, such as company disclosures and guidelines and requirements. Therefore, *set out* in these two instances possibly has a meaning other than directing the location of the details of information or intending to provide details of information. *Set out* in these instances expresses the meaning of fixing or placing signs in a particular place or position. And in these two particular cases, the signs have "No smoking" on them, and the lanterns have illuminating or flashing effect. They are to warn people not to smoke in the area and to be cautious about the road conditions. This seems to extend the meaning of *set out* to warning people by fixing or placing signs in a particular place or position.

In Publicity Materials where there are 83 instances of *set out* (0.0138%), all the instances share the most common co-selections as described above, i.e. the meaning of directing readers to where the details of the information can be found. Initially, two instances seem not to share the semantic prosody. Consider the following instance:

The architect shall be responsible for ensuring that the substation complies with the latest requirements set out by Fire Services Department, Hong Kong SAR Government.

It shares the semantic preference of 'guidelines and requirements' as suggested by the co-occurring phrase "the latest requirements". While this instance also has the most frequent configuration *noun phrase + set out + prepositional phrase*, the prepositional phrase in

this instance is headed by the preposition "by" which seems to suggest a different semantic prosody, rather than 'directing the location' as described earlier when the preposition is "in", "on" or "at". The noun phrase in this prepositional phrase headed by "by" is regarded as the agent (Biber et al., 1999), i.e. "Fire Services Department, Hong Kong SAR Government". Thus it is this authoritative body which introduced and put forward the requirements. Readers could go to or contact the authority for the details of "the latest requirements". Therefore, this instance also adheres to the typical semantic prosody of *set out* as it functions to direct to the location of where to find the details.

The following instances from Codes of Practice, Reports, Media Releases and Products Descriptions also have the same co-selections:

The owners of petrol-cum-LPG filling stations and dedicated LPG filling stations shall comply with the fire service requirements as set out by the Director of Fire Services and/or the Gas Authority.

… many European facilities are built to meet very stringent air emission requirements set out by the proposed European Commission (EC) Directives for Incineration.

All four of the objectives set out by The Government of the Hong Kong Special Administration Region in the Contract as follows are covered in the present review…

The new SOC incorporates three key principles set out by the Government…

developed countries have the responsibility to provide leadership under the core principle of common but differentiated responsibilities set out by the United Nations

The quality of the effluent well exceeds the water reuse standards set out by HKSAR, WHO & USEPA…

These instances share the colligational pattern of *noun phrase + set out + prepositional phrase*. They all have the preposition "by" which is followed by a noun phrase denoting an authority, such

as "the Government" and "the United Nations". They are all considered as sharing the semantic prosody of 'directing the location' but only that these instances do not indicate the exact location or general direction of the information in a text or document, but the official source or the issuer of the information.

The following instances, from About Us, Notes and Media Releases, have a different configuration but they may be regarded as sharing the same semantic prosody just discussed:

HEC has set out some high level key principles which must be met.

CEDD has set out the basic standard and the time frame for mandatory site safety-related training ...

In fact the Group has set out targets for cost savings ...

They have the pattern of *noun phrase + auxiliary verb + set out + noun phrase*. The phrasal verb has different colligation. They are used in the active voice as the past participle in the present perfect tense. Following *set out*, the noun phrases denote authoritative bodies. "CEDD" in the second line refers to Civil Engineering and Development Department which is a government department. In the other instances, "HEC" and "the Group" refer to companies and are not authoritative body, but in terms of setting out "some high level principles" and "targets for cost saving", they have the power to make decisions and formulate company policies. It is expressed that "HEC", "CEDD" and "the Group" are the sources which provide details of the information, i.e. "some high level principles", "the basic standard and the time frame for mandatory site safety-related training" and "targets for cost saving" respectively. Therefore, it is suggested that these three instances have the same semantic prosody of 'directing the location' as the above discussed instances.

In Media Releases, there is yet another meaning expressed by *set out*. Four instances are extracted as follows:

Fifteen Hong Kong secondary school students, who are winners of the Hong Kong Student Science Project Competition 2001, today (July 24) set out for the United Kingdom to attend the London International Youth Science Forum …

Ten Hong Kong secondary school students, who are winners of the Hong Kong Student Science Project Competition 2003, today (July 22) set out for the United Kingdom to attend the London International Youth Science Forum.

Twenty-seven Hong Kong secondary school students, who are winners of this year's Hong Kong Student Science Project Competition, today (July 25) set out for Singapore to attend the Second Asia-Pacific Economic Cooperation (APEC) Youth Science Festival.

Thirteen Hong Kong secondary school students, who are winners of the Hong Kong Student Science Project Competition 2002, today (July 23) set out for the United Kingdom to attend the London International Youth Science Forum.

The configuration of *set out* in these four instances is:

noun phrase + relative clause + adverb + set out + prepositional phrase

The noun phrase denotes people. The relative clause functions to provide additional information to modify the noun phrase. The adverb is always realized by "today" and followed by an exact date in a pair of parenthesis. The prepositional phrase following *set out* is headed by "for" followed by a place name. Due to the co-occurrence of the phrase "Hong Kong secondary school students", the semantic preference of *set out* in these instances is 'people'. In terms of the semantic prosody, the co-selection of the place names suggests that *set out* is related to a journey. And the use of the word "today" seems to suggest that the journey just started on the present day. Therefore, *set out* in these

instances have a distinct semantic preference and semantic prosody. They express the meaning of 'people beginning a journey'.

To summarise, *set out* is predominantly used in the past participle form in a passive voice structure in the engineering sub-corpora. Only a few instances are used as the active voice in the base form and even less for the past tense form. This is different from Sinclair's (1991) short analysis of *set out* in the central corpus of the Birmingham Collection of English Texts. The frequency of *set out* in the Birmingham corpus was not reported but it is said that half of the occurrences are in the passive. The Birmingham corpus comprises of approximately 7.3 million words of a representative set of texts in contemporary written and spoken English (Renouf, 1984). A preponderance for *set out* to be used in passive is, therefore, expected for engineering-specific texts as opposed to more general English texts.

The phrasal verb is frequently preceded by a noun phrase or an auxiliary verb, and followed by a prepositional phrase or an adverb. In the majority of the case, the prepositional phrase following *set out* is headed by "in". The semantic preference of the phrasal verb tends to vary from sub-corpus to sub-corpus, but generally speaking, it co-selects with words or phrases which denote some kind of information, such as guidelines and requirements, company disclosures, or legal information related to certain environmental issues. These are all the things which are set out in documents or texts. As compared to the items identified in Sinclair's (1991) analysis, including "agenda, item, criteria, themes, hopes, lists, stages, theory, argument, programme" (p. 77), some of them shared the nature of being some kind of written, issued or published information. However, the semantic preference of *set out* in the engineering genres refers to more

technical communication. This reflects the nature of the engineering genres. Regarding the semantic prosody, *set out* has a number of meanings within and across the sub-corpora. Overall, it has a very strong association to the sense of 'directing the location', meaning directing to the location of the information. This semantic prosody is mainly characterised by the co-occurrence of a prepositional phrase, typically with "in", "on" and "at", and an adverb of place showing the direction of "above" or "below". In some cases, it is characterised by the occurrence of a official body with authoritative power. They demonstrate a varying degree of specification, from very specifically locating the particular section or paragraph in which the details are provided, to the particular piece of text, to the official source or body who issues or provides the information.

Almost all of the sub-corpora have a very high percentage of the semantic prosody of 'directing the location'. Therefore, it can be considered as the canonical meaning of *set out* in engineering genres. This meaning indeed corresponds to one of the meanings Sinclair (1991) identified for *set out* in the Birmingham corpus, i.e. "a meaning close to 'expound'" (p. 77). The example cited by Sinclair, *The report set out the alternatives*, conforms to the canonical pattern of *set out* instances in the engineering genres with some turbulence. This example can be restructured to form a passive construction which would then conform to the canonical configuration with the noun phrase "the alternatives" becoming the subject of the passive sentence, and "the report" would come after the phrasal verb *set out* and be probably headed by the preposition "in". Instances such as this one is regarded as having the semantic prosody of 'directing the location' due to the focus on the prepositional phrase coming after *set out* in passive constructions. And as pointed out above, the *set out* instances in the engineering genres are overwhelmingly used in passive, thus the phrases with the locative prepositions stand out in the current analysis

contributing to the semantic prosody. Whereas in Sinclair's analysis only half of the occurrences are in passive, the focus is thus not so much on the locative function but rather more directly expounding the information. The distinction may seem subtle but it reveals the inherent difference of the style and use between engineering English and more general English.

The sequence *be set out in* is described as a phrasal-prepositional verb in *Longman Grammar of Spoken and Written English* with the function of identifying the source of information, and has an occurrence of over 20 times per million words in the Academic Prose sub-corpus of the Longman Spoken and Written Corpus (LSWC) (Biber et al., 1999). Consider the two examples provided in *Longman Grammar of Spoken and Written English*:

These project objectives were set out in the first project report.

There are three forms of statutory demand set out in Schedule 4.

These two examples have highly comparable co-selections as those of *set out* as described in the current study. While the findings in this study show a high tendency for *set out* to co-select with "in", this is not always the case. There are other choices of preposition, such as "on" and "at", or other grammatical categories following *set out*, such as adverbs "above" and "below". And since the co-selection with these different items contribute to the same semantic prosody, i.e. 'directing the location', which is similar to the function of *set out in* as identified by Biber et al. (1999), it is argued that describing *be set out in* as a phrasal-prepositional verb may be misleading. Rather, it would be more seemly to treat *set out* as a phrasal verb which is predominantly used in the passive construction and has

a strong tendency to be followed by a prepositional phrase indicating a location of text or source of the mentioned information.

Quantitatively, the fact that *set out* is one of the most frequently used phrasal verbs in the HKEC corresponds to Biber et al.'s (1999) findings. Considering that the texts collected in the HKEC are from the daily engineering workplaces, and the academic prose sub-corpus in LSWC consists of book extracts mostly targeting audience with technical background knowledge across disciplines and published research articles, the resemblance borne by them is the technical and professional nature and style of the texts. Therefore, it can be argued that *set out* is a phrasal verb more typically used in professional and formal settings. In a study investigating the use of phrasal verbs by French-speaking EFL learners, *set out* is one of the distinctive phrasal verbs in native speakers of English as compared to French speaking learners with no occurrences at all in the learner corpora (Gilquin, 2015). It is suspected whether the nature of *set out* being used more frequently in professional and formal contexts has an influence to such findings about the underuse of this phrasal verb by English language learners since their exposure to English is probably more restricted to the general contexts.

Only in a few sub-corpora with very few instances there may be another dominant semantic prosody, namely 'intending to provide details', 'warning people by fixing signs in a particular place' and 'people beginning a journey'. The analysis of *set out* across the sub-corpora demonstrates how different co-selections create different meanings. While the semantic prosody of 'directing the location' is typically created by the use of a prepositional phrase or an adverb with a locative function, the prosody of 'intending to provide details' is attributed to the use of mainly two words in the data, namely "propose"

and "would". For 'warning people by fixing signs in a particular place', *set out* co-occurs with words denoting signs or notices on public display with a sense of warning. The instances with the last semantic prosody 'people beginning a journey' are typically followed by "for" and a place name, and collocate with words identifying people, places and time. Sinclair (1991) also identified instances of *set out* which are regularly followed by "on" or "for", referring to the start of a journey.

4.6.2 *Sets out*

The other inflectional form *sets out* is found in 17 sub-corpora (see Table 28). Its frequency is much lower than that of *set out*. Overall, there are 196 instances in the whole corpus.

Table 28: Sub-corpora in which *sets out* occurs

	Sub-corpora	size	Freq	%
1	Standards	136,024	13	0.0096%
2	Consultation Papers	111,494	9	0.0081%
3	Notes	156,255	12	0.0077%
4	Guides	783,805	54	0.0069%
5	Frequently Asked Questions	55,726	3	0.0054%
6	Review Papers	106,506	4	0.0038%
7	About Us	647,013	24	0.0037%
8	Manuals	296,299	10	0.0034%
9	Handbooks	67,284	2	0.0030%
10	Position Documents	75,660	2	0.0026%
11	Code of Practice	997,228	26	0.0026%
12	Conference Proceedings	196,498	4	0.0020%
13	Reports	979,170	15	0.0015%
14	Publicity Material	599,407	7	0.0012%
15	Project Summaries	115,829	1	0.0009%
16	Media Releases	1,566,742	8	0.0005%

17	Product Descriptions	611,549	2	0.0003%

Standards is the sub-corpus with the highest percentage of occurrence of *sets out*, but the actual frequency is only 13 (0.0096%). Figure 54 shows a sample of concordance lines of *sets out* in Standards.

```
1     Dredged Excavated Sediment Scope This Circular sets out the procedure for seeking approval to dred
2              including dredged material. The protocol sets out generic guidelines for considering waste
3           Specification (PS) from Appendix A, which sets out the requirements for the Contractor to dra
4   Works Contracts Introduction This interim note sets out further guidance on the application of ETW
5            on Construction Sites Scope This Circular sets out the policy and procedures requiring
6      Projects and Proposals  Scope  This Circular sets out the guidelines and procedures for
7  has agreed to a class assessment document which sets out the mitigation measures necessary for this
8    FROM NEW ROADS ON PLANNED USES    This section sets out the requirements for relevant departments
9     same or a nearby building. This practice note sets out guidelines regarding design measures to be
10 aggregates and put them into good use. This PNAP sets out the technical guidelines for using recycle
```

Figure 54: Sample concordance lines of *sets out* in Standards

The patterns of colligation are rather neat. On the left side of *sets out*, there is mostly a noun phrase (11 times, 84.6%). In the other two instances, there is a relative pronoun followed by a noun phrase (15.4%). To the right, there is always a noun phrase (100%). Thus, the most frequent configuration of *sets out* in Standards is:

noun phrase + sets out + noun phrase

Since this inflectional form with the affix -*s* is to mark the third person singular in the present tense (Biber et al. 1999), the noun phrase occurring to the left of *sets out* is always a singular, such as "This Circular" as in line 1 and "This section" as in line 8.

The collocates to the left of the phrasal verb include "this" (7 times), "the" (4 times), "circular" (3 times), "note" (3 times), and "scope" (3 times). On the other side of *sets out*, there are "the" (15 times), "for" (8 times), "to" (6 times), "guidelines" (4 times), "of" (4 times), "and" (3 times), "requirements" (3 times), "environmental" (2 times),

"measures" (2 times), "procedures" (2 times), and "waste" (2 times). Based on the collocates, particularly those occurring to the right, *sets out* in Standards has the semantic preference of 'guidelines and requirements' (11 times, 84.6%), for example, "the procedures for seeking approval" as in line 1, "generic guidelines" as in line 2, "the requirements" in line 3, "further guidance" in line 4, etc.

While at first glance, *sets out* may have its distinct co-selections, they actually have the canonical semantic prosody of *set out*, i.e. 'directing the location'. The noun phrases to the left of *sets out* denote the source of information, or the location where the details can be found. For example, in line 1, "the procedures for seeking approval" are provided in "This Circular". Or in line 8, the information about "the requirements for relevant departments" is detailed in "This section". And in fact, the instances can be re-constructed to become passive constructions. In such case, they would share the most frequent configurations of *set out*. Thus, all these *sets out* instances in Standards can be considered paraphrases of the canonical *set out* instances and share the semantic prosody of 'directing the location'.

In Consultation Papers, there are nine instances of *sets out* (0.0081%). The concordance lines are shown in Figure 55.

```
1  during bad air pollution days. The table below sets out the comparison of the emissions from an
2            Impact Assessment Ordinance (EIAO-TM) sets out the technical requirements for the EIA
3       air pollution. The consultation document sets out the Administration's proposals on the
4   HKEERSB. The consultation document (at Annex) sets out the Administration's proposals on the
5      requirements stipulated therein. BEC (PB) sets out a means to evaluate and assess the energy
6     Introduction 1. 1 The 2007-08 Policy Agenda sets out the Government's new and on-going
7  account of the initiative The Policy Framework sets out a comprehensive waste management strategy
8  Activities on Private Land  Purpose This paper sets out a possible option for addressing the issue
9  for introducing legislative changes. This paper sets out the response of the Administration after
```

Figure 55: Concordance lines of *sets out* in Consultation Papers

All of the instances have the colligational pattern of *noun phrase + sets out + noun phrase*. The collocates to the left are "the" (6 times), "consultation" (2 times), "document" (2 times), "paper" (2 times), "policy" (2 times), "this" (2 times). To the right of *sets out*, there are "the" (14 times), "of" (5 times), "a" (3 times), "and" (3 times), "for" (3 times), "on", "administration's" (2 times), "implementation" (2 times), "mandatory" (2 times), and "proposals" (2 times).

Different semantic preferences seem to be observed in this sub-corpus. Some of the instances are related to proposals and related matters, for example, "proposals" as in lines 3 and 4, "a possible option" in line 8. In line 1, "the comparison of the emissions" is data in support of the proposal for "banning idling vehicles with running engines". And "the response of the Administration" in line 9 were made after some examinations had been done by government departments in order "to come up with a proposal for resolving the problem of disposal of waste materials on private land". Thus, these are regarded as sharing the semantic feature of 'proposals and related matters'. The instances in lines 2 and 5 seem to have the semantic preference of 'guidelines and requirements'. In line 2, "the technical requirements for the EIA" is explicit. And in line 5, "a means to evaluate and assesses the energy efficiency performance of a building" serves as the recommended method or guideline for building developers. As for lines 6 and 7, the semantic preference of *sets out* is 'the government's policy' constituted by "the Government's new and on-going initiatives" and "a comprehensive waste management strategy". They are actions and policy introduced by the government to solve particular problems. In fact, whether they are proposals, requirements or government policies, they are all related to addressing environmental concerns.

As the same as the instances in Standards, *sets out* in Consultation Papers all adhere to the semantic prosody of 'directing the location'. The noun phrases to the right of *sets out* denote information of different kinds, represented by the different semantic preferences, and they are detailed in the noun phrase to the left of *sets out*. For example in line 1, "the comparison of the emissions …" is detailed in "The table below", or in line 4, the details about "the Administration's proposals" are provided in "The consultation document (at Annex)". Thus, they express the meaning of 'directing to the location of the details'. And they can be formed the passive voice instances and would then have the canonical configurations of *set out*.

This semantic prosody of *sets out* is prevalent across the sub-corpora. Table 29 shows the percentage of occurrence of this semantic prosody in each of the sub-corpora in which *sets out* occurs. In 13 out of the 17 sub-corpora, all of the instances have this semantic prosody. In the remaining four sub-corpora, the semantic prosody is observed in the majority of the instances (66.7% to 94.4%).

Table 29: Percentage of occurrence of the semantic prosody across the sub-corpora

	Sub-corpora	size	Freq. of *sets out*	%	Freq. of the semantic prosody 'Directing'	%
1	Standards	136,024	13	0.0096%	13	100%
2	Consultation Papers	111,494	9	0.0081%	9	100%
3	Notes	156,255	12	0.0077%	12	100%
4	Guides	783,805	54	0.0069%	51	94.4%
5	Frequently Asked Questions	55,726	3	0.0054%	2	66.7%
6	Review Papers	106,506	4	0.0038%	4	100%
7	About Us	647,013	24	0.0037%	22	91.7%
8	Manuals	296,299	10	0.0034%	10	100%
9	Handbooks	67,284	2	0.0030%	2	100%

10	Position Documents	75,660	2	0.0026%	2	100%
11	Code of Practice	997,228	26	0.0026%	26	100%
12	Conference Proceedings	196,498	4	0.0020%	3	75%
13	Reports	979,170	15	0.0015%	15	100%
14	Publicity Material	599,407	7	0.0012%	7	100%
15	Project Summaries	115,829	1	0.0009%	1	100%
16	Media Releases	1,566,742	8	0.0005%	8	100%
17	Product Descriptions	611,549	2	0.0003%	2	100%

As illustrated above, these instances has the prevailing pattern of colligation *noun phrase + sets out + noun phrase*. It is observed that the noun phrase preceding *sets out* is generally a shorter noun phrase representing a more generic term, while the noun phrase following the phrasal verb tends to be a complex and larger chunk of noun phrase with more specific details. For example, in Consultation Papers (see Figure 55 above), the noun phrases preceding *sets out* are "The table", "The consultation document", "The 2007-08 Policy Agenda" and "This paper". These noun phrases represent some types of documents or parts of a document, and they may give little or no information of what they are concerned with or what kinds of texts and information are contained in them. "The consultation document" is essentially related to the consultation of a public issue, and "The 2007-08 Policy Agenda" is related to the government policies in year 2007-08. "The table" and "This paper" are more abstract and generic in terms of what are included in the section or the document. Whereas on the other side of the phrasal verb, the longer and more complex noun phrases provide information with higher specificity. They represent more specific details of what the noun phrase to the left of *sets out* contain or are about. This pattern of having a more generic noun phrase to the left of *sets out* and a long complex noun phrase with specific details to the right of it constitutes evidence of the

most frequent semantic prosody 'directing the location'. Also, further to such evidence is that the typical *sets out* instances are paraphrases of the passive sentences which are the canonical form of *set out* with the canonical semantic prosody. A minor distinction of the semantic prosody of *set out* and *sets out* is on the specificity of the location or source of the information. The indication of the location of the information in the use of *set out* tends to be more specific as there is a frequent use of a prepositional phrase directing to the particular page, section or paragraph, e.g. "in sections 5 to 8", "in Clause 13 above", "in paragraph 1(b)(iv)", "on pages 139 to 141", whereas for *sets out*, there is such specific indication in some cases, but in most of the instances it tends to be referring to a whole piece of text or document, e.g. "This Circular", "The Guidance Note", "This code".

The few instances which do not conform to the canonical meaning may have their distinct use and meaning. In Guides, there are 54 instances of *sets out*. 94.4% (51 times) of them have the typical co-selections described above. Three instances do not fit in the typical pattern. Extracts of the three instances are given below:

Overall objectives are goals that an organization sets out in order to live up to its vision, values and policies ...

... environmental targets and milestones represent the detailed performance requirements that an organization sets out to achieve.

... environmental targets and milestones represent the detailed performance requirements that the bureau or department sets out to achieve.

The sequence coming in front of *sets out* is a noun phrase followed by a relative pronoun and then another noun phrase. To the right of *sets out*, unlike the majority of the instances in which there is also a noun phrase, there is an infinitive marker followed by a verb. In short, these three instances have the following configuration:

269

noun phrase + relative pronoun + noun phrase + sets out + infinitive marker + verb

The noun phrase coming immediately in front of *sets out* represents an organisation or office, i.e. "an organization" in the first and second line, and "the bureau or department" in the third line. These instances of *sets out* can be considered as having the semantic preference of 'guidelines and requirements' as suggested by the phrases "overall objectives and goals" and "the detailed performance requirements". They are targets which aim to guide people for certain actions and behaviour. It is observed that *sets out* carries a sense of 'fulfilling' in these three instances. This semantic prosody is postulated by the infinitive phrase following the phrasal verb, i.e. "in order to live up to its vision, values and policies" in the first line and "to achieve" in the second and the third lines. These phrases express a purpose or an intention of fulfilling or attaining certain goals and targets. Therefore, *sets out* in these instances tend to have the meaning of intending to fulfill the guidelines or requirements.

There are three instances in Frequently Asked Questions and they are extracted as follows:

Table 1 sets out the 2-level IAQ objectives.

Table 2 also sets out the recommended objectives for individual volatile organic compounds …

EPD sets out the wastewater discharge standards for individual operations in accordance with the Technical Memorandum …

They have the colligation of *noun phrase + sets out + noun phrase*. The phrases "IAQ objectives", "the recommended objectives" and "the wastewater discharge standards" constitute to the semantic preference of 'guidelines and requirements'. All these three

instances conform to the canonical semantic prosody of 'directing the location' because they locate where the details about the guidelines and requirements can be found. In line 1, "the 2-level IAQ objectives" are detailed in "Table 1", and in line 2, "the recommended objectives for individual volatile organic compounds" are provided in "Table 2". As for the third line, although "EPD" is not a piece of text or document, it stands for Environmental Protection Department and is the source authoritative body who issues the information regarding "the wastewater discharge standards…". Thus, this instance also has the canonical semantic prosody of 'directing to the location'.

In About Us, there are 24 instances. 95.8% (23 times) of them adhere to the canonical form and meaning of *sets out*, with some turbulence. There are the configurations *noun phrase + sets out + noun phrase* (17 times, 70.8%) and *noun phrase + relative pronoun + sets out + noun phrase* (6 times, 25%). They are more typical forms of *sets out*. The noun phrases to the left of *sets out* are more generic and the noun phrases to the right are more specific and informative. And they can be re-structured to form passive sentences which share the canonical meaning of *set out*. There is only a single instance which seems to express a different meaning. The instance is extracted below:

… the Hub sets out to attract enterprises and IT professionals to Dalian from around the world.

The noun phrase preceding *sets out* does not refer to some documents or parts of documents, nor authoritative body. Based on the wider co-texts, "the Hub" refers to "DALIAN TIANDI.software hub" which is a well-equipped centre for IT and software services and development. The phrasal verb is not followed by a noun phrase in this case, but by a non-finite clause "to attract enterprises and IT professionals". It seems to suggest what the Hub aims for. This seems a distinct meaning from other instances but there is

271

only this single instance in the sub-corpus. An instance in Conference Proceedings also has a very similar meaning to this one:

HKSTP therefore sets Out to help Hong Kong assert itself as a city of creativity and execution.

The noun phrase preceding *sets out* is an acronym standing for "Hong Kong Science & Technology Parks Corporation". Although it is not exactly of the same nature as "the Hub", they do share some feature that is to operate for some purposes. In this case, the phrase following *sets out* shares a very similar sense to that about "the Hub". It seems to identify the aim or the purpose of "HKSTP". A closer inspection of the relationship of the phrases on both sides of *sets out* suggests that, the phrases to the right of *sets out* tend to be pinpointing the core nature of the noun phrases to the left of *sets out*. In other words, attracting enterprises and IT professionals to Dalian from around the world is the main objective of "DALIAN TIANDI.software hub", i.e. "the Hub" in the above instance from About Us, and in the instance from Conference Proceedings, helping Hong Kong assert itself as a city of creativity and execution is the core function of HKSTP. It is argued that these two instances share the semantic prosody as those three instances in Guides expressing 'intending to fulfill the goals'.

To summarise the above discussion, *sets out* is used much less frequently in engineering sub-corpora as compared to *set out*. As an inflected verb form for marking the third person singular in the present tense, *sets out* has a typical configuration of *noun phrase + sets out + noun phrase*, or sometimes *noun phrase + relative pronoun + sets out + noun phrase*. At the first glance, *sets out* has its distinct patterns, but it indeed shares the canonical form of *set out* by means of paraphrasing the *sets out* instances to

passive sentences. And *sets out* has a typical pattern of having a more generic noun phrase to the left of it, and a more complex noun phrase to the right of it denoting more specific information. *Sets out* conforms to the canonical meaning of *set out*, i.e. 'directing the location', but with some turbulence. It is used to direct readers to the location or source of information, but with generally lower specificity when comparing to the use of *set out*.

Across the sub-corpora in which *sets out* occurs, there are only a few instances (3.6%) which do not have the canonical meaning. They may be expressing fulfilling the guidelines and requirements, or introducing the particular aim of organisations. Their occurrences are too limited and therefore it is doubtful whether these other meanings can be regarded as specific to engineering genres. This meaning is also identified as signifying intention (Sinclair, 1991). They share the same co-selection of having a to-infinitive clause following the phrasal verb, as in the example cited from the Birmingham corpus, *Babbage set out to build a full scale working version*. The intriguing distinction is that this meaning which is found to be associated with nearly half of all the occurrences of *set out* in Sinclair's (1991) analysis is only observed in 9 *set out* instances (0.7% of all *set out* occurrences) in the whole HKEC distributing in six sub-corpora, and 4 *sets out* instances (2.0% of all *sets out* occurrences) in the HKEC. This suggests that the meaning of intending to achieve something or signifying intention expressed by *set out* may be common in general English, but rare in engineering English.

4.6.3 *Setting out*

The last inflectional form *setting out* has very low frequency in the corpus. There are 84 times of occurrence in total, and they are spread across 18 sub-corpora. Table 30 shows the sub-corpora in which the phrasal verb occurs and the respective frequency in each sub-corpus. It can be seen that the frequency of *setting out* in each sub-corpus is very low. There are only two sub-corpora in which more than ten instances are found.

Table 30: Sub-corpora in which *setting out* is found

	Sub-corpora	size	Freq	%
1	Handbooks	67,284	5	0.0074%
2	Consultation Papers	111,494	5	0.0045%
3	Fact Sheets	26,059	1	0.0038%
4	Q & A	27,703	1	0.0036%
5	Notes	156,255	5	0.0032%
6	Standards	136,024	3	0.0022%
7	Ordinances	139,176	3	0.0022%
8	Review Papers	106,506	2	0.0019%
9	Frequently Asked Questions	55,726	1	0.0018%
10	Code of Practice	997,228	16	0.0016%
11	Reports	979,170	11	0.0011%
12	Manuals	296,299	3	0.0010%
13	Publicity Material	599,407	5	0.0008%
14	Guides	783,805	6	0.0008%
15	About Us	647,013	4	0.0006%
16	Media Releases	1,566,742	9	0.0006%
17	Transaction Proceedings (HKIE)	1,055,248	3	0.0003%
18	Product Descriptions	611,549	1	0.0002%

Handbooks, the sub-corpus with the highest percentage of occurrence of *setting out*, has five times of occurrence of the phrasal verb (0.0074%). Figure 56 shows the concordance lines of the five instances.

```
1      highway geometry  design computation, its setting out on ground and subsequent monitoring.(H)
2    purpose.(DR)  check and verify contractors' setting out in accordance with drawings and
3  check as built works.(DR)   check contractor's setting out works.(TR)  assist in producing earthwork
4  data for planning and design, quality control. setting out and complies  with all survey requests of
5  execution of their role.(H)  check contractor' setting-out work and to verify positions of construct
```

Figure 56: Concordance lines of *setting out* in Handbooks

On the left side of *setting out*, there is a possessive pronoun (4 times, 80%). One of the instances is in the form of "its" and the others are a noun phrase ending with -*'s*. Having this colligation indicates that *setting out* is used as a gerund instead of a verb. Line 4 is the only case in which a possessive pronoun is not found to the left. In this instance, *setting out* and its following words is one of the phrases listed as a bullet point under a job title. To the right of *setting out*, there are a noun phrase (3 times, 60%) and a prepositional phrase (2 times, 40%).

Collocates to the left of *setting out* include "check" (3 times) and "contractor" (2 times). There is no outstanding collocates on the other side of *setting out*. There is a semantic preference of 'building and construction works' (100%) as suggested by the co-occurring words or phrases "highway geometry design computation" in line 1, "contractors" in line 2, "contractor" and "works" in line 3, "survey requests of the engineer in relation to the design and construction of the project" in line 4, and "contractor" and "work" in line 5. In this context, *setting out* has the meaning of positioning bolts or pins in the ground or marking concrete to mark out the locations of walls and structures for the proposed construction to be performed (Scott, 1991; Webster, 1997; Blockley, 2005). Regarding the semantic prosody, the words or phrases

"monitoring", "check and verify", and "check" found in lines 1, 2, 3 and 5 constitute to a semantic prosody of 'checking' (4 times, 80%). Therefore, *setting out* in Handbooks is used to express the meaning of 'checking the *setting out* works for building and construction'.

Consultation Papers is the sub-corpus with the second highest percentage of occurrence of *setting out*. There are also five instances (0.0045%). Figure 57 shows the concordance lines the five instances.

```
1  the contractor to submit for approval a plan setting out the waste management measures, which will
2  the contractor to submit for approval a plan setting out the waste management measures, which will
3  to provide a supplementary information paper setting out the extent of land resumption necessitated
4  the contractor to submit for approval a plan setting out the waste management measures, which will
5          to the overall interest of Hong Kong.  Setting out the future fees in one piece of legislation
```

Figure 57: Concordance lines of *setting out* in Consultation Papers

Inspection of the concordance lines shows that the co-selections of *setting out* in this sub-corpus differ from those in Handbooks. Colligationally, there is mostly a noun phrase (4 times, 80%) immediately occurring to the left of *setting out*. To the right, there is always a noun phrase (100%). Following a noun phrase as in lines 1 to 4, *setting out* is a present participle functioning as an adjective phrase modifying its preceding noun phrase.

Collocates to the left are "a" (4 times), "approval" (3 times), "for" (3 times), "plan" (3 times), and "submit" (3 times). To the right, there are "the" (5 times), "appropriate" (3 times), "include" (3 times), "management" (3 times), "measures" (3 times), and "waste" (3 times). In fact, the co-texts in lines 1, 2 and 4 are very much the same. This has an influence on the collocational pattern. There does not seem to be a particular semantic preference of *setting out* in this sub-corpus. The co-occurring phrases "the waste

276

management measures", "the extent of land resumption", and "the future fees" tend to denote rather different things. They are information of different kind, but there is not much in common about them. As for the semantic prosody, it is argued that they all share the canonical semantic prosody of *set out* and *sets out*, i.e. 'directing the location'. *Setting out* together with its following noun phrase is modifying the noun phrase to the left of *setting out*. This means that, for example in line 1, what "a plan" is about is "setting out the waste management measures". In other words, the plan is where readers can find the details of the waste management measures. Or in line 3, "a supplementary information paper" is where the details of "the extent of land resumption". As for line 5, it is expressed that "the future fees" are detailed and described in "one piece of legislation". Therefore, these instances in Consultation Papers are considered as having the semantic prosody of 'directing the location' with some turbulence.

This semantic prosody is common across most of the sub-corpora. It is observed in the majority of the instances in 11 out of 18 sub-corpora. In the other sub-corpora, some instances share this semantic prosody, others may express other meanings. In the sub-corpus of Notes, there are five total instances of *setting out* (0.0036%). They are shown in Figure 58.

```
1  control of air pollution in vehicle tunnels by setting out :- (i) the air pollution concentration
2     design opportunities  and facilitates easy setting out. Current productions of paving units by
3          Set of screed boards,  Brooms.  4.6.4 Setting Out  Accurate setting out of the paving patt
4 determine the exact location of base lines  for setting out on site  4.6.5 Sub-base and Road base
5  boards,  Brooms.  4.6.4 Setting Out  Accurate setting out of the paving pattern on site eliminates
```

Figure 58: Concordance lines of *setting out* in Notes

In line 1, s*etting out* conforms to the typical semantic prosody. Consider its wider co-text:

This Practice Note provides guidelines on control of air pollution in vehicle tunnels by setting out :- (i) the air pollution concentration limits required for protection of public health; and ...

This means that "the air pollution concentration limits require for protection of public health" is one of the "guidelines on control of air pollution in vehicle tunnels". It can be seen that the message expressed in this instance is to show that some information, i.e. the guidelines, is detailed in the Practice Note. Thus, this instance shares the semantic preference of 'guidelines and requirements' and has the semantic prosody of 'directing the location'.

The instance in line 5 seems to have similar co-selections as those instances in Handbooks. The noun phrase "the paving pattern" refers to one of the construction processes involving the paving of concrete, rocks or other materials. Thus, *setting out* in this case shares the semantic preference of 'building and construction works'. In terms of the semantic prosody, the word "accurate" and the phrase "eliminates unnecessary cutting of paving units along edges and ensures a well finished job" imply that the "setting out" works, i.e. positioning of the bolts, should be checked and monitored for quality assurance. Therefore, it is suggested that this instance has the semantic prosody of 'checking'. For the instances in lines 2 and 4, the semantic preference is observed to be the same as in that of line 5, i.e. 'building and construction works'. This is suggested by the phrases occurring in the co-texts "precast concrete paving units" and "paving works". These two instances are observed to have a distinct semantic prosody. The phrases

"facilitates easy" in line 2 and "shall be prepared to determine the exact location of base lines" suggest that something is done prior to the "setting out" works to facilitate the smooth performing of the work. Thus, the semantic prosody for these two instances is 'facilitating'. In line 3, *setting out* is used independently with a number preceding it. This form of occurrence indicates that it is a numbered sub-heading of a section.

In Codes of Practice, there are 16 instances of *setting out* (0.0016%). They seem to express a variety of meanings in this sub-corpus. Figure 59 shows a sample of concordance lines.

```
1   simply added together such that the geometrical setting out of the overall structure exceeds reasona
2              It should be noted that problems of setting out  and tolerance often occur at the interf
3        of a column to 50mm from its theoretical setting out.  This  overall tolerance is clearly of
4    criteria for the test site trial are met. N2 Setting out of soil nails Check whether the position
5   No. Stage Description F1 (a) Driven Piles (i) Setting out of piles Check that the locations of pil
6     Bored Piles, Barrette Piles and the like (i) Setting out of piles Check that the locations of pil
7   Mini-piles, Socketted H-piles and the like (i) Setting out of piles Check that the locations of pil
8   drilling. F22 (d) Rafts and Spread Footings (i) Setting out of rafts and spread footings Check that
9    Authority. "method statement" means a document setting out working procedures and sequences includ
10    2(1) of the Buildings Ordinance. It is a plan setting out the plan of safety management of build
```

Figure 59: Sample concordance lines of *setting out* in Codes of Practice

In terms of colligation, there is a preposition (4 times, 25%), a determiner at the L1 or L2 position (3 times, 18.8%) and a noun phrase (2 times, 12.5%). In five instances, *setting out* is preceded by a number. Following *setting out*, there are mostly a noun phrase (7 times, 43.8%) and a prepositional phrase (6 times, 37.5%). *Setting out* is used as a gerund functioning as a noun phrase in 13 instances (81.3%), such as lines 1 and 2. For the other instances, *setting out* is a present participle which functions to modify its preceding noun phrase, such as lines 9 and 10. In lines 4 to 8, the configuration of *setting out* is *number + setting out + prepositional phrase*. They denote numbered sub-headings.

There does not seem to be any specific collocates to the left of *setting out*. To the right, there are "of" (11 times), "the" (7 times), "piles" (3 times), "lines" (2 times), "bolts"

(2 times), etc. In 11 instances (68.8%), *setting out* occurs in the contexts of 'building and construction works', such as "manufacture and fabrication" and "overall structure" in line 1, "the interface between steel and concrete components" in line 2, "multi-storey column plumb" and "1/600 of storey height" in line 3, "soil nails" in line 4, "piles" in lines 5 to 7, "rafts and spread footings" in line 8, "the bolts" and "the centroidal axis" in lines 11 and 12, and "erection" in line 13. These are various processes and procedures involved in building and construction works. In some cases, there is the semantic prosody of 'checking' (5 times, 31.3%). This is observed in lines 4 to 8, and another instance in the sub-corpus. In lines 4 to 8, although *setting out* is used as the heading of a sub-section which is simply a short phrase, its following sentences elaborate what the sub-section is about. The sentences all begin with the word "check", for example, the following two lines are extracted from lines 4 and 5:

Check whether the positions of the soil nails agree with the approved plans.

Check that the locations of piles agree with approved plans.

They describe the checking of the accuracy and quality of the works.

Setting out in lines 1 to 3 are mentioned in relation to the concept of "tolerance limits", "tolerance", "deviations" and "deviation". They suggest that the specifications of the actual "setting out" works may deviate from the designed structure or the model, and that such difference can be tolerated to a certain extent or within a certain limits. Therefore, these three instances have a semantic prosody of 'accepting departure from a standard'.

Another semantic preference 'guidelines and requirements' is observed in four instances (25%), through the co-occurrence of the phrases "working procedures and

sequences" in line 9, "the plan of safety management" in line 10, "safety precautions" in line 14 and "the selection criteria" in line 15. These phrases denote recommended or required procedures or principles to be followed. The instances in lines 9 and 10 instances tend to have a semantic prosody of 'defining' as suggested by the words "means" in line 9, and "defined" in the wider co-texts in line 10:

A supervision plan is defined under section 2(1) of the Buildings Ordinance. It is a plan setting out the plan of safety management of building works or street works.

Also, the noun phrase preceding *setting out* used with the indefinite determiner "a", as in "a document" and "a plan", suggests that they are not referring to a specific piece of text, but rather, any document providing the details of "working procedures and sequences" and "the plan of safety management of building works or street words" would be referred to as "method statement" and "a supervision plan" respectively. Line 15 can be regarded as sharing the semantic prosody of 'directing the location'. It has a more generic noun phrase "the following guideline" in which the information "the selection criteria of the key operating staff" is detailed. Thus, this is an instance conforming to the typical meaning but with some turbulence.

Overall, the form *setting out* is used the least frequently in engineering genres. Instead of being used in the present tense progressive form, it is mostly used as a gerund or present participle as an adjective phrase. As the same as *set out* and *sets out*, *setting out* collocates with words which denote some kind of information, such as guidelines and requirements. The majority of the instances have the semantic prosody of 'directing the location', expressing that some sorts of information is detailed in particular documents.

More interestingly, a distinct meaning is observed in some instances of *setting out*. They typically collocate with words sharing the feature of building and construction works, rather than documents or information of any kinds. This seems to be a technical use and is defined as marking out the location for construction of walls or structures by marking out pegs or concrete. For such use of *setting out*, they have the semantic prosody of 'checking' and, in a few cases, 'facilitating' and 'accepting departure from a standard'. This means that *setting out* is used to express the meaning of 'checking the marking works for constructions and building' and 'facilitating the marking works for constructions and building' and 'accepting the marking works depart from a standard'. Although the frequency of *setting out* having these semantic preference and prosodies is limited, these semantic preference and semantic prosodies are found in more than one sub-corpus and are not observed in the use of *set out* or *sets out*. Thus, these co-selections are specific to the form *setting out* in engineering genres.

4.6.4 Summary

The findings show that *set out* is used predominantly as a past participle in the passive voice. In the majority of the instances, set out, sets out and setting out share similar co-selections in expressing the meaning of directing readers to a location in a document where readers can find details of some information. In addition to this canonical extended unit of meaning, other meanings are found to be specific to different inflectional forms, despite the low frequency. *Set out* can be used to express 'intending to provide details', 'warning people by fixing signs in a particular place' and 'people beginning a journey'. *Sets out* can be used to express 'fulfilling goals in relation to some guidelines and

requirements'. The form *setting out* is not used as the present progressive tense and can express 'checking' or 'facilitating' 'the marking works for building and construction, or 'accepting some departure from a standard'.

Sinclair et al. (1989) list four meanings for SET OUT, namely (1) to start a journey, (2) to take action to achieve some goals, (3) to explain facts, ideas and opinions in a clear way, and (4) to put something in an organised way to show them. Rundell and Fox (2005) also list the same four meanings for SET OUT. These four meanings are found in the use of SET OUT in the current study. However, it is argued that some meanings are more specific to a particular inflectional form.

The non-contiguous instances for the previous sets of phrasal verbs are infrequent, there is even none for SET OUT. The syntactic flexibility of the phrasal verbs, i.e. whether it is influenced by their idiomaticity, may have an influence. Those phrasal verbs with a literal spatial meaning have a higher tendency to display the discontinuous order, i.e. verb + particle + object, whereas idiomatic phrasal verbs predominantly display the continuous order (Gries & Stefanowitsch, 2004). As found in the current study, almost all the meanings expressed by SET OUT are idiomatic. The only non-idiomatic meaning found is specific to *set out* in expressing 'warning people by fixing signs in a particular place or position'.

In another study, SET OUT was identified as a distinctive phrasal verb for native English as compared to French learner of English (Gilquin, 2015). It is believed that French learners of English may demonstrate a preference for phrasal verbs which are more transparent in meaning. In addition to variation between native and non-native English learners, usage difference between American English and British English is also

observed for the use of SET OUT. British academic English shows a much higher preference to SET OUT than American academic English (Liu, 2012). However, rather than concluding this as a usage difference between the two varieties of English, Liu (2012) suggests that a difference in the topics and areas of the data or a difference in diction in academic writing between the two varieties could be a contributing factor. These studies show how SET OUT may have its distinctive usage in some situations, but details in terms of the particular use and meanings were not investigated.

Chapter 5   Conclusions

The final chapter of this thesis concludes the present study by summarising the major results of the analysis and discussing the implications of the findings in terms of the methodology, research and pedagogy. The limitations of the study are given in the last section.

5.1 Major findings and conclusions

The study explores the use of phrasal verbs in the English used in the engineering profession in Hong Kong. The most frequent phrasal verbs, their variant forms and inflectional forms are examined for frequency analysis and textual analysis. The guiding research questions of the study are as follows:

1) How do phrasal verbs within and across the engineering genres investigated compare in frequencies? What are engineering genre-specific phrasal verbs?

2) To what extent can a phrasal verb and its variant forms be specific to engineering genres in terms of its frequency and co-selection?

3) How are the inflectional forms of a phrasal verb lemma similar to and/or different from each other in terms of their frequencies and patterns of co-selection?

The following part summarises the major findings in response to these research questions.

Regarding the first research question, the comparison and analysis of the 31 lists of the top 10 most frequent phrasal verbs show that only eight phrasal verbs are shared in more than 10 sub-corpora. These eight phrasal verbs are *based on, carried out, comply with, refer to, set out, associated with, carry out* and *set up*. Among these top phrasal

verbs, only *based on* and *carried out* are shared in more than 20 sub-corpora. Not only are these two phrasal verbs frequent in terms of their overall frequencies, but they are also generally very frequent within the sub-corpora. For example, *based on* is among the top three most frequent phrasal verbs in 22 sub-corpora. These most frequent phrasal verbs can be considered as specific to the engineering profession.

It should be pointed out again that this study examines and compares the use of phrasal verbs only within one specialised corpus, the Hong Kong Engineering Corpus. The most frequent phrasal verbs argued as specific to the engineering profession are based on the fact that they have the highest overall frequencies in the whole corpus, they are very commonly used across different genres (i.e. found in more than one-third of the sub-corpora), and they are among the top three most frequent phrasal verbs in more than half of the sub-corpora. This may or may not render them register-specific if a general English corpus is taken into account, which is outside the scope of the current study and could be further investigated in future research.

In exception of *based on* and *carried out*, not many other phrasal verbs are shared in the majority of the sub-corpora. This indicates that, in terms of genre level, the engineering genres having different nature and functions do not share the same set of the most frequent phrasal verbs.

The second and the third questions were researched with the concordance analysis of the most frequent phrasal verbs in addition to the frequency analysis. With regard to the extent to which a phrasal verb can be specific to engineering genres in terms of their co-selections, the results of the present study reveal that different phrasal verbs have different behaviour. While some phrasal verbs such as *based on* are not specific to a

particular genre but have very similar co-selection irrespective of which genre or situation they are used in, others such as *refer to* have varying degree of specificity to a particular meaning depending on the genres in which they occur.

The most frequent phrasal verb *based on* has the same co-selection across the genres. It is predominantly used as a past participle in the passive structure. It collocates with words or phrases sharing semantic features of evaluation and assessment, and overall there is a strong semantic prosody of 'rigorous'. This extended unit of meaning is found in the vast majority of the instances across the sub-corpora. There is minimal occurrence of other extended units of meaning of this phrasal verb and it exhibits little turbulence.

The phrasal verbs *carried out* and *set out* tend to have a major extended unit of meaning across the sub-corpora. For *carried out*, it is mostly used as a past participle in passive structure. Its collocates denote various kinds of engineering-related activities, such as construction works, building and installation, testing and inspection, and research. They share the semantic preference of 'engineering-related works'. It has a strong semantic prosody of 'requirement' across the sub-corpora. This co-selection dominates in most of the sub-corpora in which the phrasal verb is found. In only a few sub-corpora, there are other extended units of meaning. One of such units of meaning is the use of *carried out* with words denoting names of authors and years of publication. This different semantic prosody observed is 'acknowledging'. As for *set out*, it is used either as a past participle in the passive construction or an infinitive. Overall, the phrasal verb has various semantic preferences, namely 'guidelines and requirements', 'proposal-related matters', 'legal information related to pollution control', 'company disclosure', 'legislation and regulation' and 'data and findings'. Despite a wide range of semantic preferences, a strong semantic prosody of 'directing the location' is observed for the use of *set out*. In

addition to this major extended unit of meaning, *set out* has different co-selections in some sub-corpora expressing other minor meanings, including the co-occurrence with words or phrases describing consequences of not providing details of particular information, denoting signs to warn people, and signalling the start of a journey.

The results of the other two most frequent phrasal verbs examined show that they have a wider range of extended units of meaning. For *comply with*, although there is a tendency for this phrasal verb to have a major extended unit of meaning, it also has several others across the sub-corpora. The most frequent co-selection involves the pattern of a modal verb preceding *comply with*. The collocates denote different forms of requirements, thus the phrasal verb has a semantic preference of 'requirement'. Here the phrasal verb has the semantic prosody of 'obligatory'. In addition to this extended unit of meaning, a number of other meanings are observed for the use of *comply with*. They share the same semantic preference of 'requirement', but have different semantic prosodies, namely 'evaluating', 'describing consequences', 'intending' and 'making certain'. The results suggest that *comply with* has a very strong semantic association with words denoting requirements regardless of various semantic prosodies. In a number of sub-corpora, there is a rather even distribution of the percentage of occurrence of different units of meaning. This result indicates that a phrasal verb can have different functions and co-selections within a genre.

The other phrasal verb *refer to* is associated with five main units of meaning. Overall, the phrasal verb co-occurs with words denoting different kinds of information. There is a tendency for *refer to* to have the semantic prosodies of 'directives' and 'recommendation'. The other two extended units of meaning are created with the pattern of having a pronoun preceding *refer to*, with the semantic prosody of 'precise'; and the

pattern of having a noun phrase preceding the phrasal verb, with the semantic prosody of 'defining and clarifying'. Despite an obvious tendency for *comply with* and *refer to* to associate with a particular co-selection, they have different co-selections in some sub-corpora. They are more specific to some genres in terms of a particular extended unit of meaning. For example, *refer to* can be regarded as to a large extent specific to the genre of Letters to Editor in terms of expressing the meaning of 'making precise reference'. The uniqueness of this meaning in this genre is both due to the high frequency of it in the genre and the rarity of it in other genres.

To summarise, the five most frequent phrasal verbs examined in the study exhibit varying extent of specificity to engineering genres.

In terms of the other phraseological forms a phrasal verb exhibits, except for SET OUT, BASE ON, CARRY OUT, COMPLY WITH and REFER TO display non-contiguous instances in the form of constituency variation, such as *base…on*, *carried…out*, *complying…with* and *refer…to*. Positional variation is only observed for *on…based*, *on…base* and *with…comply*. Regardless of whether it is constituency variation or positional variation, the number of frequency is minimal as compared to contiguous form.

Generally, the constituency variation does not shift the meaning of the phrasal verbs. The type of words or phrases that drops in between the verb and the particle of the phrasal verbs depends on the type of phrasal verbs. In most cases, the words or phrases dropping in between the verb and the particle are adverbials, for example, "again", "only", "solely", and "in particular". These intervening adverbials do not change the meaning of the phrasal verb but function to vary the intensity of the meaning. CARRY OUT is a transitive Type 1 phrasal verb with a verb and an adverbial particle. Particle movement is

possible when the direct object is a pronoun (Sroka, 1962, 1972; Live, 1965; Bolinger, 1971; Lindner, 1981; Celce-Murcia & Larsen-Freeman, 1999; Jackendoff, 2002; Downing & Locke, 2006; Thim, 2012). The words intervening in *carried…out*, *carry…out* and *carrying…out* are pronouns, such as "them" and "it". In such cases, the phrasal verb is regarded as exhibiting the discontinuous order (Dehé, 2001, 2002). These phrasal verbs in the discontinuous order have the same co-selection of the contiguous form. BASE ON and REFER TO can be a doubly transitive Type II phrasal verb. They may take two noun objects. Thus the words dropping in between the verb and the prepositional particle can be a noun phrase, as in for example *base + noun phrase + on + noun phrase* and *refer + noun phrase + to + noun phrase*. Only in the instances of refer…to and referred…to, a shift in the meaning of the contiguous form is observed. Therefore, *refer to* and *refer…to* are regarded as different extended units of meaning, as well as *referred to* and *referred…to*.

Positional variation is exhibited in only a few phrasal verbs, namely *on…based*, *on…base* and *with…comply*. All these instances of positional variation are demonstrating the preposed structure of Type II phrasal verbs. The prepositional particle together with a following relative pronoun referring to the object noun phrase is fronted as in, for example, *on + relative pronoun (object) + noun phrase (subject) + base*. Despite the difference in form, these positional variants share the same meaning of their contiguous counterparts.

The final research question concerns the uniformity and variance of different inflectional forms of a lemma. Based on the hypothesis that different inflectional forms do not necessarily share the frequency of use and the same co-selection, each of them is

treated as distinct phrasal verbs for analysis. Five sets of phrasal verbs comprising five lemmas are examined. The results show a remarkably uneven distribution of frequencies across the phrasal verbs despite that they comprise the same lemma.

Regarding the co-selection, on the one hand, some phrasal verbs comprising the same lemma show a high resemblance of co-selections. In other words, regardless of the form, they are used with the same units of meaning with little turbulence. On the other hand, some phrasal verbs demonstrate a noticeable difference in terms of their dominating pattern of co-selections. This indicates that the hypothesis can be confirmed and inflection forms of a lemma do not necessarily share the same meaning.

Results obtained from the present study underpin some implications in terms of methodology and future research, and pedagogy. They are discussed in the following sections.

5.2 Pedagogical implications

The findings of the current study have implications for the teaching and learning of English for specific purposes. The results of the present study could be utilized for the pedagogy of teaching English for engineering students in terms of the selection of phrasal verbs for teaching, the way phrasal verbs are taught and learnt, and the presentation of the items in grammar references and dictionaries.

With the use of naturally occurring data from the professional engineering contexts, the authentic profile of high-frequency phrasal verbs are provided. This profile provides teachers and learners with a clear starting point as there is a higher chance that students will encounter the high-frequency phrasal verbs in the professional workplace, as

the case with other lexical items (Nation, 2001; Leech, 2011; Schmitt & Redwood, 2011). These most frequent phrasal verbs are recommended to be selected in the teaching of engineering English as they have a particular role in the register and genres which may not be replaced by their single-word counterparts.

In addition to the selection of phrasal verbs, the way in which the phrasal verbs are taught and introduced is also crucial in the pedagogy. Even with the right selection of items for teaching, in other words, key vocabulary, students are not familiar with the appropriate textual environment in which the item occurs (Flowerdew, 2001). One of the possible consequences is "collocational mismatches" (ibid, p. 371) or other stylistic problems. The analysis of the phrasal verbs based on Sinclair's (1991) five categories of co-selection provides a more accurate and complete picture in respect to the use of phrasal verbs in authentic contexts. Enhancing students' knowledge of the appropriate grammatical patterns, collocation, semantic preference and semantic prosody of the phrasal verbs can help students achieve the unmarked use of the phrasal verbs in the professional settings. As it has been proved that phrasal verbs can be register- or genre-specific, teachers are recommended to make use of corpora representing different genres and disciplines and corpus concordances of phrasal verbs in teaching to demonstrate the distinctive patterns of use of phrasal verbs in different contexts.

Another important implication for improving teaching and learning of phrasal verbs addresses the treatment of lemma and inflectional forms in textbooks, grammar reference books and dictionaries. Despite the increasing use of corpus as the basis, many practical aspects of information regarding the use of phrasal verbs are found insufficient or even lacking in most published materials and dictionaries. The importance of these aspects of information, such as frequency, co-selections, phraseological variation and

variation among different inflectional forms, highlights the needs to incorporate the information in the pedagogical and reference materials. Particularly, findings from the present study reveal that inflectional forms comprising the same lemma can behave considerably in different ways in terms of the use and frequency. Therefore, the traditional approach in lemmatising inflectional forms and presenting each lemma as a single entry or distinct linguistic item may be problematic and ineffective for users. Textbook writers, grammar reference developers and lexicographers are recommended to take into account the issues concerned and provide more comprehensive and detailed accounts of phrasal verbs in the materials for the enhancement of students' awareness and knowledge of phrasal verbs.

One possible direction, therefore, is the use of data-driven learning (Johns, 1991) to English for specific purposes pedagogy. The theoretical framework and findings in this study can be applied to the classroom with engineering students. Teachers can provide them with corpus data of authentic engineering text and support to use basic corpus tools. Possible language learning activities include students extracting the most frequent phrasal verbs from the corpus using the corpus search function and identifying patterns of typical use of the phrasal verbs using the concordancer. More specifically, teachers can provide two concordances of the same phrasal verb from different genres in which the phrasal verb has different use (e.g. *comply with* in Ordinances and Reports), and two concordances of a phrasal verb and its variant form (e.g. *refer to* and *refer…to*) which show different use and meaning, and also concordances of two inflectional forms of a phrasal verb lemma (e.g. *refer to* and *refers to*). By examining and comparing the concordances, and identifying regularities and patternings, students' learning is driven by linguistic data and they are able to discover how different genres may change the way a

phrasal verb is used, the extent to which a phrasal verb may display phraseological variation and some variant forms may have a shift in meaning, as well as how they should be aware of the use of different inflectional forms.

This data-driven approach to language learning is believed to be much more beneficial to language learners. Teachers take the role of a facilitator instead of an instructor deciding what phrasal verbs to teach and how to teach them; students are provided with access to real linguistic evidence so that they can act as researchers themselves and answer the linguistic queries they have by identifying regularities and patternings in the language data (Johns, 1991).

5.3 Methodological and research implications

The present study points to three main aspects of implications on methodology and research. The first methodological implication is concerned with the dual approach used in the present study with both quantitative and qualitative analysis. The second implication addresses the use of phrasal verbs as an indicator of genre distinction. The third implication points to a future direction of research.

The quantitative results of phrasal verbs provide a general picture about the different set of phrasal verbs used across the sub-corpora, the frequency of use of a phrasal verb across the sub-corpora, and the distinction between the uses of the inflectional forms. In terms of the comparison of the set of phrasal verbs used across the sub-corpora, common and unique phrasal verbs are identified. While there is an overlap of a small set of phrasal verbs in most of the sub-corpora, a considerable difference in the most frequent phrasal verbs used in each sub-corpus is noticed, and there are even unique

phrasal verbs which are only found in a particular sub-corpus. The frequencies of the most frequent phrasal verbs are tracked across the sub-corpora and a noticeable uneven distribution is observed. The qualitative analysis then shows some phrasal verbs are used in the same way across different sub-corpora despite the frequency variation, whereas others have different co-selections and express different meanings in different sub-corpora.

Regarding the inflectional forms comprising a lemma, the quantitative results show a striking variation between the inflection forms in all the five sets of phrasal verbs examined in the present study. Apparently they can be considered distinct phrasal verbs for teaching and learning based on the quantitative results. However, the qualitative analysis reveals that some sets of inflectional forms bear a close resemblance in their co-selections despite variation in form and frequency, whereas others have different uses. It is worth noting and pointing out that the two approaches complement each other in revealing a more complete and accurate view of the use of phrasal verbs in engineering English.

The second research implication addresses the issue of phrasal verbs being an indicator to distinguish genres. Previous efforts have been put on the variation within registers. Most of the effort has been on the distinction between spoken and written registers (for example, Biber, 1988; Biber et al., 2002; Louwerse et al., 2004), or between formal and informal settings (for example, Darwin & Gray, 1999). However, phrasal verbs did not have an essential role in the issue. Later, an attempt was made to explore particularly the contribution of phrasal verbs as an indicator of spoken versus written distinction and informal versus formal distinction, and it was suggested that phrasal verbs

are capable of making significant distinction on the two dimensions (Dempsey et al., 2007). While these above mentioned studies provide insight into the use of phrasal verbs as a text indicator, the emphasis has been restricted to the two dimensions on formalness and mode of communication. The analysis of the use of phrasal verbs within a genre and across genres in the present study has shown that certain phrasal verbs are specific to a particular genre in terms of its frequency of use and its specific co-selections. Results of the study have implications for using phrasal verbs for more specific genre distinction, for example, between regulative genres and reporting genres, or more specifically between specific genres such as Ordinances and Publicity Materials.

In addition to the results obtained, the methodology used in the current study can be applied to the study of phrasal verbs in corpora representing other registers and genres. Future studies can examine the use of phrasal verbs in other corpora to examine whether there are distinctive sets of phrasal verbs that are frequent in different registers, and across the genres within other registers, and whether there are particular uses of phrasal verbs which are specific to the registers and genres. Particularly, it is intriguing to investigate the use of phrasal verbs in more formal and specialised registers and genres. Another possible direction of future studies is to compare phrasal verbs in engineering English with a general corpus of English.

The methodological implication of the current study is not restricted to the examination of phrasal verbs. It can be extended to studying other linguistic items, such as reporting verbs, modal verbs, evaluative adjectives, degree adverbs, phraseologies, or any possible items. Particularly, analysis using the model of five categories of co-selection is useful in revealing and providing a more comprehensive and complete description of meanings of the items investigated. This study also contributes to further

understanding of Sinclair's (1996, 2004) five categories of co-selections by underpinning the concept that semantic prosody is often more specific than simply a positive-negative analysis. It is the specific semantic prosody which describes and distinguish meanings.

5.4 Limitations of the study

The researcher is fully aware of the limitation with regard to the sizes of the sub-corpora. The present study utilizes the 9.2 million word Hong Kong Engineering Corpus which consists of 31 genre-based sub-corpora. This profession-specific corpus undoubtedly offers invaluable authentic language data for researchers in conducting analysis, irrespective of whether to examine lexical items and phraseologies across the genre-based sub-corpora within the corpus or to compare the analysis in this corpus with a corpus representing another register. It is rather well-represented in terms of the collection of a wide variety of genres which engineers encounter in their daily workplaces. Yet, the sizes of certain genre-based sub-corpora of the corpus are small. On the one hand, the imbalance of the sub-corpora, to some extent, represents the proportion of the genres engineers produce and encounter. This imbalance of the sub-corpora may affect the comparability of the results of the phrasal verb analysis across the genres, resulting in low raw frequencies of some phrasal verbs in certain sub-corpora. Thus, some of the conclusions derived from the findings might not be generalisable.

Appendix

Lists of the top ten most frequent phrasal verbs in the 31 sub-corpora

R	Abstract	Freq.	R	Agreements	Freq.	R	About Us	Freq.
1	based on	84	1	based on	143	1	based on	196
2	carried out	38	2	lead to	51	2	set out	155
3	compared with	35	3	accounted for	37	3	committed to	131
4	dispose of	31	4	associated with	31	4	invest in	107
5	set out	25	5	account for	28	5	associated with	102
6	carry on	22	6	referred to	23	6	set up	95
6	entitled to	20	7	compared to	22	7	relating to	89
8	referred to	19	7	related to	22	8	listed on	83
9	deal with	15	7	regarded as	22	9	lead to	87
9	related to	15	10	contribute to	18	9	contribute to	76

R	Code of Practice	Freq.	R	Circular Letters	Freq.	R	Conference Proceedings	Freq.
1	carried out	807	1	set out	121	1	based on	94
2	comply with	435	2	based on	64	2	carried out	75
3	based on	347	3	recognised as	41	3	consists of	55
4	provided with	244	4	arising from	40	4	connected to	43
5	connected to	232	5	entered into	35	5	equipped with	31
6	carrying out	204	6	referred to	34	6	compared with	28
7	refer to	186	7	engaged in	27	6	lead to	28
8	carry out	166	8	addressed to	21	6	related to	28
9	provided for	160	9	charged to	22	9	associated with	26
10	referred to	157	10	entitled to	20	10	comply with	21

R	Consultation Papers	Freq.	R	Frequently Asked Questions	Freq.	R	Fact Sheets	Freq.
1	based on	54	1	refer to	27	1	set up	17
2	set out	49	2	based on	25	2	opened to	14
3	refer to	42	3	comply with	24	3	carried out	10
4	comply with	34	4	apply for	23	4	cope with	7
5	set up	27	5	dispose of	17	5	setting up	6
6	carried out	25	6	carried out	16	5	provided for	6
7	taking into	25	7	disposed of	14	7	built up	4
7	arising from	22	8	compare with	11	7	cater for	4
9	associated with	21	8	set up	11	7	connected to	4
10	switch off	19	8	carrying out	11	7	constructed of	4
						7	ranging from	4

R	Guides	Freq.	R	Handbooks	Freq.	R	Letters to Editor	Freq.
1	carried out	458	1	carried out	32	1	thank for	8
2	comply with	307	1	assist in	32	2	signed up	5
3	based on	273	3	carry out	31	3	refer to	4
4	carry out	273	4	deal with	30	4	point out	3
5	set out	208	5	based on	27	4	referred to	3
6	refer to	172	6	liaise with	26	4	set up	3
7	carrying out	154	7	comply with	22	7	compared with	2
8	submitted to	124	8	set out	21	7	drop off	2
9	fitted with	111	8	relating to	21	7	extended to	2
10	result in	106	10	refer to	16	7	look forward to	2
			10	fitted with	16	7	obtained from	2
			10	provided with	16	7	participate in	2

R	Manuals	Freq.	R	Media Releases	Freq.	R	Notes	Freq.
1	carried out	162	1	refer to	339	1	carried out	122
2	based on	125	2	participate in	336	2	based on	73
3	comply with	101	3	set up	289	3	carry out	57
4	provided for	73	4	carried out	227	4	set up	28
5	refer to	68	5	based on	210	5	related to	26
6	carry out	60	6	take of	191	6	comply with	23
7	referred to	56	7	set out	166	7	arising from	22
8	associated with	39	8	listed on	163	8	set out	21
9	provided with	38	9	relating to	158	8	refer to	21
10	depends on	36	10	compared with	132	10	taken into	20
						10	obtained from	20

R	Ordinances	Freq.	R	Position Documents	Freq.	R	Plans	Freq.
1	carried out	129	1	associated with	36	1	provide for	4
2	referred to	114	2	based on	35	1	set out	4
3	comply with	103	3	related to	17	1	allow for	4
4	carry out	90	3	lead to	17	4	cater for	2
5	carrying out	66	5	result in	16	4	committed to	2
6	provided with	61	6	committed to	15	4	consults on	2
7	apply to	53	7	refer to	13	4	give out	2
8	served on	47	8	contribute to	12	4	lead to	2
9	provided for	36	9	carried out	11	4	taking into	2
10	provide for	36	10	resulted in	9	4	apply to	1
			10	set up	9	4	consulted on	2
						4	made of	2

R	Publicity Material	Freq.	R	Product Descriptions	Freq.	R	Project Summaries	Freq.
1	based on	313	1	based on	169	1	aims at	53
2	comply with	287	2	carried out	150	2	comprises of	52
3	carried out	269	3	equipped with	102	3	ranging from	29
4	carry out	265	4	comply with	84	4	consists of	28
5	participate in	179	5	connected to	79	5	carried out	21
6	submitted to	163	6	carry out	77	5	comprise of	21
7	participating in	124	7	ranging from	64	7	associated with	18
8	provided with	113	8	provided with	61	7	cater for	18
9	carrying out	104	9	made of	59	7	cope with	18
10	connected to	77	10	provided for	53	10	based on	17

R	Q & A	Freq.	R	Reports	Freq.	R	Review Papers	Freq.
1	based on	15	1	carried out	646	1	based on	44
2	thank for	13	2	based on	334	2	set up	37
3	comply with	7	3	follow up	320	3	set out	31
4	carried out	5	4	arising from	152	4	comply with	29
4	compared with	5	5	carry out	149	4	taking into	29
4	engaged in	5	6	refer to	133	6	pointed out	25
4	relating to	5	7	set out	122	7	carried out	24
4	switch off	5	8	comply with	117	8	carry out	20
4	apply to	5	9	clean up	104	8	tie in	20
10	arising from	4	10	associated with	98	10	complying with	19
10	equipped with	4				10	disposed of	19
10	focus on	4						
10	provide for	4						
10	related to	4						
10	set up	4						
10	switching off	4						
10	taking into	4						

R	Standards	Freq.	R	Speeches	Freq.	R	Tender Notices	Freq.
1	comply with	52	1	carried out	2	1	based on	6
2	provided for	46	1	stand up	2	2	addressed to	4
2	set out	46	1	come from	2	2	obtained from	4
4	carried out	44				4	complied with	2
5	apply for	43				4	submitted to	2
6	based on	42						
7	exempted from	27						
7	disposed of	27						
9	provided with	25						
9	carrying out	25						

R	Technical Papers	Freq.	R	Transaction Discussions (HKIE)	Freq.	R	Transaction Notes (HKIE)	Freq.
1	start up	36	1	based on	8	1	based on	63
2	based on	33	2	carried out	4	2	carried out	29
3	shut down	23	3	depend on	3	3	compared with	22
4	consists of	15	3	related to	3	4	obtained from	19
5	set up	11	5	aim at	2	5	consists of	11
6	associated with	10	5	applied to	2	5	applied to	11
6	depends on	10	5	associated with	2	7	comply with	10
8	result in	9	5	built upon	2	7	limited to	10
9	carry out	8	5	correlated with	2	9	connected to	9
9	compared with	8	5	inferred from	2	10	associated with	8
9	combined with	8	5	refer to	2	10	lead to	8
9	contribute to	8						

R	Transaction Proceedings (HKIE)	Freq.
1	based on	908
2	carried out	422
3	compared with	250
4	obtained from	210
5	associated with	192
6	related to	181
7	applied to	149
8	depends on	146
8	lead to	146
10	subjected to	138

References

Aarts, B. (2011). *Oxford modern English grammar*. Oxford: Oxford University Press.

Akbari, O. (2009). *A corpus-based study on Malaysian ESL learners' use of phrasal verbs in narrative compositions* (Unpublished doctoral dissertation). Universiti Putra Malaysia.

Akbari, Z. (2017). Reading scientific texts: Some challenges faced and strategies used by EFL readers. *Journal of Applied Linguistics and Language Research, 4*(1), 14-28.

Alejo-González, R. (2010). Making sense of phrasal verbs: A cognitive linguistic account of L2 learning. *AILA Review, 23*(1), 50-71.

Alejo-González, R. (2012). Variability in L2 acquisition of formulaic sequences: a study of linguistic factors in the use and learning of phrasal verbs by non-native speakers. *Language Value, 4*(1), 33-62.

Barekat, B., & Baniasady, B. (2014). The impact of phrasal verb avoidance on the writing ability of the university EFL learners. *Procedia - Social and Behavioral Sciences, 98*, 343-352.

Bastow, T. (2003). Friends and allies: Binomials in a corpus of US defence speeches. Paper given at ASLA Symposium, Orebro University, 6 November.

Bhatia, V. K. (2013). *Analysing genre: Language use in professional settings*. London/New York: Routledge.

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

Biber, D. (1992). The multi-dimensional approach to linguistic analyses of genre variation: an overview of methodology and findings. *Computers and the Humanities, 26*(5/6), 331-345.

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at…: Lexical bundles in university teaching and textbooks. *Applied Linguistics, 25*(3), 371-405.

Biber, D., Conrad, S., Reppen, R., Byrd, P., & Helt, M. (2002). Speaking and writing in the university: A multi-dimensional comparison. *TESOL Quarterly, 36*, 9-48.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. London: Longman.

Blockley, D. (2005). *The new Penguin dictionary of civil engineering*. London: Penguin.

BNC. (2007) http://www.natcorp.ox.ac.uk/docs/URG/BNCdes.html

Bolinger, D. (1971). *The phrasal verb in English*. Cambridge: Harvard University Press.

Breeze, R. (2012). Verb-particle constructions in the language of business and finance. *Language Value, 4*(1), 84-96.

Brinton, L. J. (1985). Verb particles in English: Aspect or aktionsart? *Studia Linguistica, 39*, 157-168,

Campoy Cubillo, M. C. (2002). Phrasal and prepositional verbs in specialized texts: a creative device. *IBÉRICA, 4*, 95-111.

Carter, R., & McCarthy, M. (2006). *Cambridge grammar of English: A comprehensive guide*. Cambridge: Cambridge University Press.

Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL/EFL teacher's course*. Boston: Heinle & Heinle Publishers.

Census and Statistics Department (2017) The four key industries and other related industries in the Hong Kong economy. Hong Kong Monthly Digest of Statistics, May 2017 Issue, FA1-FA12.

Chen, M. (2013). Overuse or underuse: A corpus study of English phrasal verb use by Chinese, British and American university students. *International Journal of Corpus Linguistics, 18*(3), 418-442.

Chen, M. (2016). A corpus-based comparative study on the use of prepositional verbs in writing by Chinese, American and British students. In F. S. Topor (Ed.), *Handbook of research on individualism and identity in the globalized digital age* (pp. 478-499). Hershey PA: IGI Global.

Chen, M. (2017). Beyond frequencies: Investigating the semantic and the stylistic features of phrasal verbs in a three-year longitudinal corpus by Chinese university students. In V. Brezina, & L. Flowerdew (Eds.), *Learner corpus research: New perspectives and applications* (pp 134-154). Bloomsbury.

Chen, P. (1986). Discourse and particle movement in English. *Studies in Language, 10*, 79-95.

Cheng, L., Cheng, W., & Li, J. (2016). Defamation case law in Hong Kong: A corpus-based study. *Journal of the International Association for Semiotic Studies, 208*, 203-222.

Cheng, W. (2006). Describing the extended meanings of lexical cohesion in a corpus of SARS spoken discourse. *International Journal of Corpus Linguistics, 11*(3), 325-344.

Cheng, W. (2010). Hong Kong Engineering Corpus: Empowering professionals-in-training to learn the language of their profession. In M. C. Campoy-Cubillo, B. Bellés-Fortuño, & M. L. Gea-Valo (Eds.). *Corpus-based Approaches to English Language Teaching* (pp. 67-78). London and New York: Continuum.

Cheng, W. (2012). *Exploring corpus linguistics: language in action*. London: Routledge.

Cheng, W., Greaves, C., Sinclair, J. McH., & Warren, M. (2009). Uncovering the extent of the phraseological tendency: towards a systematic analysis of concgrams. *Applied Linguistics, 30*(2), 236-252.

Cheng, W., Greaves, C., & Warren, M. (2006). From n-gram to skipgram to concgram. *International Journal of Corpus Linguistics, 11*(4), 411-433.

Cheng, W., & Leung, M. (2012). Exploring phraseological variations by concgramming:

The realization of complete patterns of variations. *Linguistic Research, 29*(3), 617-638. Institute for the Study of Language and Information, Kyung Hee University, South Korea

Clear, J. (1993). From Firth principles: Computational tools for the study of collocation. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair*. Philadelphia: John Benjamins.

Conrad, S. (2000). Will corpus linguistics revolutionize grammar teaching in the 21st century? *TESOL Quarterly, 34*(3), 548-560.

Cornell, A. (1985). Realistic goals in teaching and learning phrasal verbs. *IRAL-International Review of Applied Linguistics in Language Teaching, 23*(1-4), 269-280.

Courtney, R. (1983). *Longman dictionary of phrasal verbs*. England: Longman.

Dagut, M., & Laufer, B. (1985). Avoidance of phrasal verbs - A case for contrastive analysis. *Studies in Second Language Acquisition, 7*(1), 73-79.

Dalle, T. S. (1983). *Phrasal verbs: The effectiveness of current grammatical theory in applied ESL and some pedagogical implications.* University of Arizona.

Darwin, M., & Gray, L. (1999). Going after the phrasal verb: an alternative approach to classification. *TESOL Quarterly, 33*(1), 65-83.

Dehé, N. (2000). On particle verbs in English: More evidence from information structure. In N. M. Antrim, G. Goodall, M. Schulte-Nafeh & V. Samiian (Eds.). *Proceedings of the 28th Western Conference on Linguistics, 11 (WECOL 1999)* (pp. 92-105). California State University, Fresno.

Dehé, N. (2001). Transitive particle verbs in English: The neutral order. Evidence from speech production. In N, Dehé & A. Wanner (Eds.). *Structural aspects of semantically complex verbs* (pp. 165-189). Berlin/Frankfurt/New York: Peter Lang.

Dehé, N. (2002). *Particle verbs in English: Syntax, information structure and intonation.*

Amsterdam/Philadelphia: John Benjamins.

Dehé, N., Jackendoff, R., McIntyre, A., & Urban, S. (2002). *Verb-particle explorations*. Berlin/New York: Mouton de Gruyter.

Dempsey, K. B., McCarthy, P. M., & McNamara, D. S. (2007). Using phrasal verbs as an index to distinguish text genres. In D. Wilson & G. Sutcliffe (Eds.), *Proceedings of the Twentieth International Florida Artificial Intelligence Research Society Conference* (pp. 217-222). Menlo Park CA: AAAI Press.

Dixon, R. M. W. (1982). The grammar of English phrasal verbs. *Australian Journal of Linguistics, 2*, 1-42.

Downing, A., & Locke, P. (2006). *English grammar: A university course*. (Second edition). London/New York: Routledge.

Elenbass, M. (2003). Particle verbs in early Middle English: The case of *up*. In L. M.E.A. Cornips & P. Fikkert (Eds.), *Linguistics in the Netherlands* (pp. 45-57), Amsterdam: John Benjamins.

Fairclough, N. L. (1965). Some English phrasal types: Studies in the collocations of lexical items with prepositions and adverbs in a corpus of spoken and written present day English. MA thesis, London: University College.

Firth, J.R. (1935). The technique of semantics. *Transactions of the philological society, 34*(1), 36-73.

Firth, J. R. (1957). Modes of Meaning. In F. R. Palmer (Ed.), *Papers in Linguistics* (pp.190-215). London: Oxford University Press.

Firth, J. R. (1968). A synopsis of linguistic theory, 1930-55. In F. R. Palmer (Ed.) *Selected papers of J.R. Firth* (pp. 168-205). London: Longmans.

Fletcher, B. (2005). Register and phrasal verbs. *MED Magazine, 33*.

Flowerdew, L. (2004). The argument for using English specialized corpora to understand academic and professional language. In U. Connor & T. A. Upton (Eds.), *Discourse*

in the Professions: Perspectives from corpus linguistics (pp. 11-36). Amsterdam/Philadelphia: John Benjamins.

Fraser, B. (1965). An examination of the verb-particle construction in English. Cambridge, MA: M.I.T PhD dissertation.

Fraser, Bruce (1976). *The verb-particle combination in English*. New York: Academic Press.

Gardner, D., & Davies, M. (2007). Pointing out frequent phrasal verbs: a corpus-based analysis. *TESOL Quarterly, 41*(2), 339-359.

Garnier, M., & Schmitt, N. (2015). The PHaVE list: A pedagogical list of phrasal verbs and their most frequent meaning senses. *Language Teaching Research, 19*(6), 645-666.

Garside, R., & Smith, N. (1997). A hybrid grammatical tagger: CLAWS4. In R. Garside, G. Leech & A. McEnery (Eds.). *Corpus Annotation: Linguistic Information from Computer Text Corpora* (pp. 102-121). London: Longman.

Gilquin, G. (2009). The (non-)use of phrasal verbs in L2 varieties of English. Paper presented at 30th International Computer Archive of Modern and Medieval English Conference (ICAME 30), Lancaster University, United Kingdom.

Gilquin, G. (2015). The use of phrasal verbs by French-speaking EFL learners. A constructional and collostructional corpus-based approach. *Corpus Linguistics and Linguistic Theory, 11*(1), 51-88.

Goźdź-Roszkowski, S. (2011). *Patterns of linguistic variation in American legal English: a corpus-based study*. Frankfurt am Main: Peter Lang.

Greaves, C. (2009). *ConcGram 1.0: A phraseological search engine*. Amsterdam: John Benjamins.

Greaves, C., & Warren, M. (2007). Concgramming: A computer-driven approach to learning the phraseology of English. *ReCALL Journal, 19*(3), 287-306.

Greenbaum, S. (1996). *The Oxford English grammar*. New York: Oxford University Press.

Gries, S. Th., & Stefanowitsch, A. (2004). Extending collostructional analysis: A corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics, 9*(1), 97-129.

Haddican, B., & Johnson, D. E. (2012). Effects on the particle verb alternation across English dialects. *University of Pennsylvania Working Papers in Linguistics, 18*(2), 31-40.

Halliday, M. A. K. (2004). *An introduction to functional grammar*. (Third edition). London: Arnold.

Hampe, B. (2002). *Superlative verbs: A corpus-based study of semantic redundancy in English verb-particle constructions*. Tubingen: Gunter Narr Verlag Tubingen.

Hampe, B. (2012). Transitive phrasal verbs in acquisition and use: A view from construction grammar. *Language Value, 4*(1), 1-32.

Ho, D. C. P., Ahmed, S. M., Kwan, J. C., & Ming, F. Y. W. (2000). Site safety management in Hong Kong. *Journal of Management in Engineering, 16*(6), 34-42.

Hoey, M. (1991). *Patterns of Lexis in Text*. Oxford: Oxford University Press.

Hoey, M. (2004). Lexical priming and the properties of text. In A. Partington, J. Morley, & L. Haarman (eds.), *Corpora and discourse* (pp. 385-412). Bern: Peter Lang.

Hoey, M. (2005). *Lexical priming: A new theory of words and language*. London/New York: Routledge.

Huddleston, R., & Pullum, G. K. (2002). *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Hunter, P. J., & Prideaux, G. D. (1983). Empirical constraints on the verb-particle construction in English. *Journal of the Atlantic Provinces Linguistics Association, 5*, 3-15.

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

Hunston, S. (2003). Lexis, wordform and complementation pattern: A corpus study.
    Functions of Language, 10(1), 31-60.

Hunston, S. (2007). Semantic prosody revisited. *International Journal of Corpus
    Linguistics, 12*(2), 249-268.

Hunston, S., & Francis, G. (2000). *Pattern grammar: A corpus-driven approach to the
    lexical grammar of English.* Amsterdam/Philadelphia: John Benjamins.

Hunston, S., & Thompson, G. (2000). Evaluation: An introduction. In S. Hunston, & G.
    Thompson (Eds.), *Evaluation in text: authorial stance and the construction of
    discourse* (pp. 1-27). New York: Oxford UP.

Ishii, Y. (2009). Making a list of essential phrasal verbs based on large corpora and
    phrasal verb dictionaries. In Y. Kawaguchi, & M. Minegishi, & J. Durand (Eds.),
    Corpus analysis and variation in linguistics (pp. 121-140). Amsterdam/Philadelphia:
    John Benjamins.

Ishikawa, K. (1999). English verb-particle constructions and a V0-internal structure.
    English Linguistics, 16(2), 329-352.

Jackendoff, R. (2002). English particle constructions, the lexicon, and the autonomy of
    syntax. In N. Dehé, R. Jackendoff, A. McIntyre & S. Urban (Eds.). *Verb-particle
    explorations* (pp. 67-94). Berlin/New York: Mouton de Gruyter.

Jackson, H. (2007). *Key terms in linguistics*. London: Continuum.

Kamarudin, R. (2013). A study on the use of phrasal verbs by Malaysian learners of
    English. PhD dissertation. University of Birmingham.

Kennedy, G. (2002). Variation in the distribution of modal verbs in the British National
    Corpus. In R. Reppen, S. M. Fitzmamurice & D. Biber (Eds.) *Using corpora to
    explore linguistic variation* (pp. 73-90). Amsterdam/Philadelphia: John Benjamins.

Khir, A. N. (2012). A semantic and pragmatic approach to verb particle constructions used in cartoons and puns. *Language Value, 4*(1), 97-117.

Knowles, G., & Don, Z. M. (2004). The notion of a "lemma": Headwords, roots and lexical sets. *International Journal of Corpus Linguistics, 9*(1), 69-81.

Kovács, É. (2007). *Reflections on English phrasal verbs*. Miskolci Egyetemi Kiadó. Universitatis Miskolcinensis.

Leech, G. (2011). Frequency, corpora and language learning. In F. Meunier, S. De Cock, G. Gilquin & M. Paquot (Eds). *A taste for corpora: In honour of Sylviane Granger* (pp. 7-32). Amsterdam / Philadelphia, PA: John Benjamins.

Legum, St. E. (1968). The verb-particle constructions in English: Basic or derived? In papers from the Fourth Regional Meeting of the Chicago Linguistic Society (pp. 50-62). Chicago: Chicago Linguistic Society.

Liao, Y., & Fukuya, Y. J. (2004). Avoidance of phrasal verbs: The case of Chinese learners of English. *Language Learning, 54*(2), 193-226.

Lin, Y. Y., & Chung, S. F. (2016). A corpus-based study on the semantic prosody of challenge. *Taiwan Journal of TESOL, 13*(2), 99-146.

Lindner, S. J. (1981). A lexico-semantic analysis of English verb particle constructions with *out* and *up*. PhD dissertation. University of California.

Lipka, L. (1972). *Semantic structure and word-formation: Verb-particle constructions in contemporary English*. Munich: Wilhelm-Fink-Verlag.

Liu, D. (2011). The most frequently used English phrasal verbs in American and British English: a multicorpus examination. *TESOL Quarterly, 45*(4), 661-688.

Live, A. H. (1965). The discontinuous verb in English. *Word, 21*, 428-451.

Lohse, B., Hawkins, J. A., & Wasow, T. (2004). Domain minimization in English verb-particle constructions. *Language, 80*(2), 238-261.

Louw, B. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair* (pp. 157-76). Amsterdam: John Benjamins.

Louw, B. (2000). Contextual prosodic theory: Bringing semantic prosodies to life. In C. Heffer, H. Sauntson, & G. Fox (Eds.), *Words in context: A tribute to John Sinclair on his retirement* (pp. 48-94). Birmingham: University of Birmingham.

Louwerse, M. M., McCarthy, P. M., McNamara, D. S., & Graesser, A. C. (2004). Variation in language and cohesion across written and spoken registers. *Proceedings of the Annual Meeting of the Cognitive Science Society, 26*, 843-848.

Mahpeykar, N., & Tyler, A. (2015). A principled cognitive linguistics account of English phrasal verbs with *up* and *out*. *Language and Cognition, 7*(1), 1-35.

McArthur, T. (Ed.). (1992). *The Oxford companion to the English language*. Oxford, UK: University Press.

McCarthy, M. & O'Dell, F. (2007). *English Phrasal Verbs in Use: Advanced*. Cambridge: Cambridge University Press.

McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge: Cambridge University Press.

McIntosh, C. (Ed.). (2006). *Oxford phrasal verbs dictionary for learners of English*. Oxford: Oxford University Press.

Mendis, D. (2010). Formality in academic writing: The use/non-use of phrasal verbs in two varieties of English. In M. F. Ruiz-Garrido, J. C. Palmer, & I. Fortanet-Gomez (Eds.), *English for professional and academic purposes* (pp. 11-24). Amsterdam/New York: Rodopi.

Meyer, G. (1975). *The two-word verb: A dictionary of the verb preposition phrases in American English*. The Hague: Mouton.

Milizia, D. (2012). *Phraseology in political discourse: a corpus linguistics approach in the classroom*. Milano: LED.

Milizia, D. (2013). Phrasal verbs and phrasal units: political corpora within the walls of the classroom. In C. Desoutter, D. Heller, & M. Sala (Eds.), *Corpora in specialized communication* (pp. 135-164). Bergamo, CELSB.

Moon, R. (2010). What can a corpus tell us about lexis? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 197-211). Abingdon: Routledge.

Morgan, P. S. (1997). Figuring out *figure out*: Metaphor and the semantics of the English verb-particle construction. *Cognitive Linguistics, 8*(4), 327-357.

Morley, J., & Partington, A. (2009). A few Frequently Asked Questions about semantic - or evaluative - prosody. *International Journal of Corpus Linguistics, 14*(2), 139-158.

Nakano, T., & Koyama, Y. (2005). e-Learning materials development based on abstract analysis using web tools. In R. Khosla, R. J. Howlett & L. C. Jain (Eds.), *Knowledge-Based Intelligent Information and Engineering Systems. KES 2005. Lecture Notes in Computer Science, vol 3681* (pp. 794-800). Berlin: Springer.

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

Nelson, G. (2001). *English: An essential grammar. London*, New York: Routledge.

Nelson, G., & Ren, H. (2012). Particle verbs in African Englishes. In M. Hundt, & U. Gut (Eds.), *Mapping unity and diversity world-wide: Corpus-based studies of new Englishes* (pp. 197-214). Amsterdam/Philadelphia: John Benjamins.

Newman, J. A. (2016). A Corpus-Based Comparison of the Academic Word List and the Academic Vocabulary List. *All Theses and Dissertations*. 6080.

O'Dowd, E. M. (1998). *Prepositions and particles in English: A discourse-functional*

account. New York, Oxford: Oxford University Press.

Oprit-Maftei, C. (2014). Challenges of phrasal verbs to Ecology and environmental protection students. *Cultural Intertexts, 1-2*, 346-352.

Owen, C. (1993) Corpus-based grammar and the Heineken effect: Lexico-grammatical description for language learners. *Applied Linguistics, 14*(2), 167-187.

Palmer, F. R. (1965). *A linguistic study of the English verb*. London: Longmans. Green & Co.

Parrott, M. (2010). *Grammar for English language teachers*. Cambridge: Cambridge University Press.

Partington, A. (2003). *The linguistics of political argument*. London: Routledge.

Partington, A. (2004). "Utterly content in each other's company": Semantic prosody and semantic preference. *International Journal of Corpus Linguistics, 9*(1), 131-156.

Partington, A. (2007). Irony and reversal of evaluation. *Journal of Pragmatics, 39*, 1547-1549.

Pelli, M. G. (1976). *Verb-particle constructions in American English: A study based on American plays from the end of the 18th century to the present*. Bern: Francke.

Qu, L. (2010). The Characteristics of Phrasal Verbs in Marine Engineering English. *Studies in Literature and Language, 1*(1), 50-56.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. London: Longman.

Rayson, P. (2009) Wmatrix: a web-based corpus processing environment, Computing Department, Lancaster University. http://ucrel.lancs.ac.uk/wmatrix/

Renouf, A. (1984). Corpus development at Birmingham University. In J. Aarts, & W. Meijs (Eds.), *Corpus linguistics: Recent developments in the use of computer corpora in English language research* (pp. 3-40). Amsterdam: Rodopi.

Requejo, M.D.P., & Díaz, C.P. (2008). A cognitive approach to some phrasal verbs in English for Specific Purposes. *Ibérica: Revista de la Asociación Europea de Lenguas para Fines Específicos (AELFE), 16*, 109-128.

Rosca, A., & Baicchi, A. (2016). Digging up the frequency of phrasal verbs in English for the police: The case of *up. Círculo de Lingüística Aplicada a la Comunicación (clac), 37*, 273-296.

Rudzka-Ostyn , B . (2003). *Word power: phrasal verbs and compounds (a cognitive approach*). Berlin / New York : Mouton de Gruyter .

Rundell, M., & Fox, G. (2005). *Macmillan phrasal verbs plus*. Oxford: Macmillan Education.

Sansome, R. (2000). Applying lexical research to the teaching of phrasal verbs. *International Review of Applied Linguistics in Language Teaching, 38*(1), 59-69.

Schmitt, N., & Carter, R. (2004). Formulaic sequences in action: An introduction. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use* (pp. 1-22). Amsterdam/Philadelphia, PA: John Benjamins.

Schmitt, N., & Redwood, S. (2011) Learner knowledge of phrasal verbs: a corpus-informed study. In F. Meunier, S. De Cock, G. Gilquin, & M. Paquot (Eds.), *A taste for corpora: In honour of Sylviane Granger* (pp. 173-208). Amsterdam/Philadelphia, PA: John Benjamins.

Schneider, E. W. (2004). How to trace structural nativization: particle verbs in world Englishes. *World Englishes, 23*(2), 227-249.

Schneider, G., & Zipp, L. (2013). Discovering new verb-preposition combinations in new Englishes. Studies in variation, contacts and change in English 13. In J. Mukherjee, & M. Huber (Eds.), *Studies in Variation, Contacts and Change in English*. Helsinki, Finland: Varieng.

Scott, J. S. (1991). *Dictionary of civil engineering*. Third Edition. London: Granada.

Scott, M. (2012). *WordSmith Tools* version 6, Liverpool: Lexical Analysis Software.

Side, R. (1990). Phrasal verbs: Sorting them out. *ELT Journal, 44*(2), 144-152.

Sinclair, J. McH. (1987). *Looking up: An account of the COBUILD project in lexical computing*. London: Collins.

Sinclair, J. McH. (1991). *Corpus, Concordance, Collocation*. Oxford: OUP.

Sinclair, J. McH. (1996). The search for units of meaning. *Textus, 9*, 75-106.

Sinclair, J. McH. (1998). The lexical item. *Amsterdam Studies in the Theory and History of Linguistic Science Series 4*, 1-24.

Sinclair, J. McH. (2000). Lexical grammar. *Naujoji Metodologija, 24*, 191-203.

Sinclair, J. McH. (2002). *Collins COBUILD phrasal verbs dictionary*. London: HarperCollins.

Sinclair, J. McH. (2004). *Trust the text: language, corpus and discourse*. London: Routledge.

Sinclair, J. McH. (2007). Collocation reviewed. Ms, Tuscan Word Centre, Italy.

Sinclair, J. McH. (2010). Defining the Definiendum. In G. De Schryver, & P. Hanks (Eds.), *A way with words: Recent advances in lexical theory and analysis. A festschrift for Patrick Hanks* (pp. 37-47). Kampala: Menha Publishers.

Sinclair, J. McH. et al. (Eds.). (1989). *Collins COBUILD Dictionary of Phrasal Verbs*. London: Collins.

Sinclair, J. McH. et al. (Eds.). (1990). *Collins COBUILD English grammar*. London: HarperCollins.

Sinclair, J. McH. et al. (Eds.). (2005). *Collins COBUILD English Grammar*. London: Collins.

Siyanova, A. & Schmitt, N. (2007). Native and nonnative use of multi-word vs. one-word

verbs. *International Review of Applied Linguistics, 45*, 119-139.

Spears, R. A. (2005). *McGraw-Hill's dictionary of American idioms and phrasal verbs*. New York: McGraw-Hill.

Sroka, K. A. (1962). Critique of the traditional syntactic approach to adverb/preposition words in modern English. *Bulletin de le Societe Polonais de Linguistique, 21*, 127-140.

Sroka, K. A. (1972). *The syntax of English phrasal verbs*. The Hague: Mouton.

Stubbs, M. (1995). Collocations and semantic profiles: On the cause of the trouble with quantitative studies. *Functions of Language, 2*(1), 23-55.

Stubbs, M. (2001). *Words and phrases: corpus studies of lexical semantics*. Oxford: Blackwell.

Stubbs, M. (2002). Two quantitative methods of studying phraseology in English. *International Journal of Corpus Linguistics, 7*(2), 215-244.

Stubbs, M. (2009). Technology and phraseology: with notes on the history of corpus linguistics. In U. Römer, & R. Schulze (Eds), *Exploring the lexis-grammar interface* (pp. 15-32). Amsterdam / Philadelphia, PA: John Benjamins.

Tekin, B. C. (2008). The construction of Turkey's possible EU membership in French political discourse. *Discourse and Society, 19*(6), 727-763.

Teubert, W. (2004). Units of meaning, parallel corpora, and their implications for language teaching. In U. Connor, & T. A. Upton (Eds.), *Applied corpus linguistics: A multidimensional perspective* (pp. 171-189). Amsterdam/New York: Rodopi.

Teubert, W., & Krishnamurthy, R. (2007) General introduction. In W. Teubert, & R. Krishnamurthy (Eds.), *Corpus linguistics: Critical concepts in linguistics* (pp. 1-37). London/New York: Routledge.

Therapeutic Goods Administration (2015) Department of Health, Australian Government. Retrieved from https://www.tga.gov.au/tga-basics

Thim, S. (2012). *Phrasal verbs: The English verb-particle construction and its history*. Berlin/Boston: Mouton de Gruyter.

Tognini-Bonelli, E. (2001) *Corpus linguistics at work*. Amsterdam/Philadelphia, PA: John Benjamins.

Trebits, A. (2009). The most frequent phrasal verbs in English language EU documents – A corpus-based analysis and its implications. *System, 37*, 470-481.

Tsui, W. (2016) Engineering industry in Hong Kong. HKTDC. Retrieved 11 September 2017 http://hong-kong-economy-research.hktdc.com/business-news/article/Hong-Kong-Industry-Profiles/Engineering-Industry-in-Hong-Kong/hkip/en/1/1X47J8WG/1X003URR.htm

Van Rooy, B., & Schäfer, L. (2003). Automatic POS tagging of a learner corpus: The influence of learner errors on tagger accuracy. In D. Archer, P. Rayson, A. Wilson, & T. McEnery (Vol. Eds.), *Technical Papers 16: Proceedings of the Corpus Linguistics 2003 Conference* (pp. 835–844). Lancaster, UK: Lancaster University Centre for Computer Corpus Research on Language.

Villavicencio, A. (2003). Verb-particle constructions and lexical resources. In F. Bond, A. Korhonen, D. McCarthy & A. Villavicencio (Eds.), *Proceedings of the ACL 2003 workshop on multiword expressions: Analysis, acquisition and treatment*. Sapporo, Japan.

Villavicencio, A. (2006). Verb-particle constructions in the world wide web. In P. Saint-Dizier (Ed.), *Computational linguistics dimensions of the syntax and semantics of prepositions* (pp. 115-130). Netherlands: Springer.

Villavicencio, A., & Copestake, A. (2002). Verb-particle constructions in a computational grammar of English. In Proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar (HPSG-2002), Seoul, South Korea.

Walter, E. & Pye, G. (1997). *Cambridge International Dictionary of Phrasal Verbs*. Cambridge: Cambridge University Press.

Wang, L. (2001). Exploring Parallel Concordancing in English and Chinese. *Language Learning and Technology, 5*(3), 174-184.

Warren, M. (2010a). Identifying aboutgrams in engineering texts. In M. Bondi, & M. Scott (Eds.), *Keyness in texts* (pp.113-126). Amsterdam: John Benjamins.

Warren, M. (2010b). Online corpora for specific purposes. *ICAME Journal, 34*, 169-188.

Warren, M. (2011). Using corpora in the learning and teaching of phraseological variation. In A. Frankenberg-Garcia, L. Flowerdew, & G. Aston (Eds.), New *trends in corpora and language learning* (pp. 153-166). London: Continuum.

Warren, M. (2013). Corpora: Specialized. In C.A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Oxford: Wiley-Blackwell Publishing Ltd.

Warren, M. (2016). Signalling intertextuality in business emails. *English for Specific Purposes, 42*, 26-37.

Warren, M., & Leung, M. (2013). 'At any time after any proceedings…': Why is *any* so frequent in legal texts?. *International Journal of Law, Language & Discourse, 3*(2), 70-102.

Warren, M., & Leung, M. (2016). Do collocational frameworks have local grammars?, *International Journal of Corpus Linguistics, 21*(1), 1-27.

Webster, L. F. (1997). *The Wiley dictionary of civil engineering and construction*. New York: John Wiley & Sons.

Wei, N. & Li, X. (2014). Exploring semantic preference and semantic prosody across English and Chinese: Their roles for cross-linguistic equivalence. *Corpus Linguistics and Linguistic Theory, 10*(1), 103-138.

White, B. J. (2012). A conceptual approach to the instruction of phrasal verbs. *The Modern Language Journal*, 96(3), 419-438.

Xiao, R. (2009). Multidimensional analysis and the study of world Englishes. *World Englishes, 28*(4), 421-450.

Xiao, R., & McEnery, T. (2006). Collocation, semantic prosody, and near synonymy: A cross-linguistic perspective. *Applied Linguistics, 27*(1), 103-129.

Zarifi, A., & Mukundan, J. (2012). Phrasal verbs in Malaysian ESL textbooks. *English Language Teaching, 5*(5), 9-18.

Zarifi, A. & Mukundan, J. (2014) Grammatical presentation of phrasal verbs in ESL textbooks. *Pertanika Journal of Social Sciences & Humanities, 22*(2), 649-664.

Zeller, J. (2001). *Particle verbs and local domains*. Amsterdam/Philadelphia: John Benjamins.

Zipp, L., & Bernaisch, T. (2012). Particle verbs across first and second language varieties of English. In H. Marianne, & G. Ulrike (Eds.), *Mapping unity and diversity world-wide: corpus-based studies of new Englishes* (pp. 181-196). Philadelphia: John Benjamins.