DECOMPOSABILITY AND TRANSPARENCY:

A CORPUS-BASED ANALYSIS OF

THE USE OF PHRASAL VERBS IN ELF INTERACTIONS

ZEYNEP SENA ABAYLI

BOĞAZİÇİ UNIVERSITY

2016

DECOMPOSABILITY AND TRANSPARENCY:

A CORPUS-BASED ANALYSIS OF

THE USE OF PHRASAL VERBS IN ELF INTERACTIONS

Thesis submitted to the

Institute for Graduate Studies in the Social Sciences

in partial fulfillment of the requirements for the degree of

Master of Arts

in

English Language Education

by

Zeynep Sena Abaylı

Boğaziçi University

2016

Decomposability and Transparency:

A Corpus-Based Analysis of the Use of Phrasal Verbs in ELF Interactions

The thesis of Zeynep Sena Abaylı

has been approved by:

Prof. Yasemin Bayyurt
(Thesis Advisor)

Assoc. Prof. Leyla Martı
(Thesis Co-Advisor)

Assoc. Prof. Didar Akar

Assist. Prof. Senem Yıldız

Assist. Prof. Zeynep Çamlıbel
(External Member)

June 2016

DECLARATION OF ORIGINALITY

I, Zeynep Sena ABAYLI, certify that

- I am the sole author of this thesis and that I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution;
- this thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- this is a true copy of the thesis approved by my advisor and thesis committee at Boğaziçi University, including final revisions required by them.

Signature .....

Date 06.06.2016

ABSTRACT

Decomposability and Transparency:

A Corpus-Based Analysis of the Use of Phrasal Verbs in ELF Interactions

This study aims to investigate the decomposability and transparency nature of the uses of phrasal verb (PV) structures through two corpora representing two dichotomous ends of the speaker continuum, namely, native speakers in British National Corpus (BNC) and non-native speakers in Vienna Oxford International Corpus of English (VOICE). It examines the usage-based patterns of PVs through the natural uses of PVs in authentic oral communications of native and nonnative speakers, thereby providing explanatory adequacy to the phenomenon. To achieve this objective, a corpus-based analysis using the concordance programs BNCweb and AntConc 3.4.4 was conducted through BNC and VOICE to do a quantitative analysis on the frequency, decomposability and transparency characteristics of PVs. The findings point to both native and nonnative speakers demonstrating an interestingly similar use of PVs, displaying avoidance behavior in their PV use and a tendency to use one-word equivalent verbs instead of using PV correspondences. As for the decomposability nature, the analysis has pointed to a striking resemblance, namely the tendency to use PVs in their non-decomposable position and the avoidance of decomposable PVs in their natural decomposable position. Therefore, the current study has taken on a new significance by drawing attention to the existence of a similar processing system by two representative speakers of English, in their PV uses, evidenced by two spoken corpora, BNC and VOICE.

ÖZET

Ayrıştırılabilirlik ve Şeffaflık: Ortak Dil İngilizce Etkileşimlerinde Öbeksi Eylemlerin Kullanımının Derlem Temelli İncelenmesi

Bu çalışma, öbeksi eylemlerin kullanımlarının ayrıştırılabilirlik ve şeffaflık doğasını İngiliz Ulusal Derlemi'ndeki (BNC) anadili İngilizce olan konuşmacıların ve Viyana Oxford Uluslararası Derlemi'ndeki (VOICE) anadili İngilizce olmayan konuşmacıların oluşturduğu iki tür derlem aracılığıyla araştırmayı amaçlamaktadır. Araştırma, anadili İngilizce olan ve olmayan konuşmacıların otantik sözlü iletişimlerinde öbeksi eylemlerin doğal kullanımları aracılığıyla bu yapıların kullanım tabanlı kalıplarını incelemektedir. Böylelikle, bu olguya açıklayıcı bir yeterlik sağlayacaktır. Bu amaca ulaşmak için, BNCweb ve AntConc 3.4.4 bağlamlı dizin yazılımları kullanan derlem-temelli bir inceleme, öbeksi eylemlerin sıklık, ayrıştırılabilirlik ve şeffaflık özellikleri üzerine nicel bir çözümleme yapmak adına BNC ve VOICE derlemleri aracılığıyla yürütülmüştür. Sonuçlar anadili İngilizce olan ve olmayan konuşmacıların ikisinin de ilginç bir biçimde benzer öbeksi eylem kullanımı gösterdiklerini, öbeksi eylem karşılıklarını kullanmak yerine bir kelimelik eşdeğerlerini kullanma eğilimi ve öbeksi eylem kullanımlarında ise kaçınma davranışı sergilediklerini işaret etmektedir. Ayrıştırılabilirlik doğasına ilişkin, inceleme; öbeksi eylemleri genellikle ayrıştırılamaz konumlarında kullanma ve ayrıştırılabilir olanları ise kendi doğal konumlarında kullanmaktan kaçınma eğilimini belirten şaşırtıcı biçimde ilginç bir benzerlikle sonuçlanmaktadır. Bu yüzden, bu çalışma BNC ve VOICE derlemleriyle belgelenen öbeksi eylemlerin kullanımlarında İngilizcenin iki konuşmacısı tarafından öne sürülen benzer bir işlem sisteminin varlığına dikkat çekmesinde önem taşımaktadır.

ACKNOWLEDGMENTS

Writing this thesis has been an illuminating, inspiring and thrilling experience for me, by means of which I have gone through different emotions for the past five years of my life. Therefore, I would like to express my sincere thanks and heartfelt gratitude to the people who have helped me to realize the thesis. Although there have been many fluctuations in this process, I have made it to the till by taking the bull by the horns. Bertrand Russell says that "The great bulk of necessary work can never be anything but painful"; yet, as Dr. Anthony John Pavlik, my mentor and instructor, in whose footsteps I have followed my academic studies and educational perspective, always says, "No pain no gain"; thus, this thesis has been painful but also very rewarding for me.

First of all, I would like to express my warmest thanks and genuine gratitude to my thesis advisor, Prof. Yasemin Bayyurt, who provided me with excellent support and encouragement to make my thesis possible, motivated me all the time and tolerated my extensions due to challenges I experienced during this process.

This thesis would have never been completed without the help of my former co-advisor, Prof. Cem Alptekin, who always displayed a sincere attitude, a motivating perspective and never-ending sympathy towards me and my endless questions. Cem Hocam, you may rest in peace, because in line with how you wanted this thesis to be completed, it was completed in exactly the way it should be for you.

Moreover, I would like to thank my current thesis co-advisor, Assoc. Prof. Leyla Martı, who supported me with her excellent feedback and greater insights in the writing process of my thesis. I also owe special thanks to my thesis committee members, Assoc. Prof. Didar Akar, Assist. Prof. Senem Yıldız, and Assist. Prof. Zeynep Çamlıbel for their support and invaluable suggestions.

I would like to express my heartfelt gratitude to Dr. Judy Monthie Doyum for her guidance in the analysis of my thesis and for her being my instructor; she always had time to listen to me and illuminate my way through her magnificent constructive comments, not only on my thesis but also on everything that I ever asked her advice on.

My special thanks should be expressed to my friends Nur Başak Karataş, Duygu Çandarlı, Elifcan Öztekin, Ecehan Sönmez, Filiz Rızaoğlu, Talip Gülle and Yavuz Kurt for being my friends and always supporting me.

Above all, I would like to express my wholehearted thanks to my beloved family for the unconditional love, endless support, patience, endurance, everlasting trust and encouragement they have provided me throughout my whole life. I am indebted to them. They are always there around me, no matter what. Without them and their contributions, nothing would have been achieved. They are the ones always motivating me, urging me not to give up. My family has been the sole illuminating light directing and leading me to do the right thing.

Last but not least, I feel compelled to explain that this thesis has witnessed the death of two distinguished scholars whom I have admired most. One is my co-advisor, Prof. Cem Alptekin, whose sudden death affected me considerably. He was an inspirational academician with his enlightening and thought-provoking perspectives. The other is Prof. Richard Ruiz, who has offered me a place in the PhD program at Arizona University, insisting that I pursue a PhD degree there; he too passed away instantaneously. I am very saddened by both of these losses. Their deaths have showed me that in this life nothing is more important than our lives. Therefore, we must achieve our aims and must not give up the race until the bitter end and that we should never worry about anything in this ephemeral world.

To my family...

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AVP	Adverbial Particle
BNC	British National Corpus
COCA	Corpus of Contemporary American English
EFL	English as a Foreign Language
ELF	English as a Lingua Franca
ENL	English as a Native Language
ESL	English as a Second Language
L1	First Language
L2	Second Language
LV	Lexical Verb
MWE	Multi-Word Expression
NLA	Nondecomposable Lexical Annotation
NLP	Natural Language Processing
ODPV	Oxford Dictionary of Phrasal Verbs
PRP	Prepositional Phrase
PV	Phrasal Verb
SCL	Socio-Cognitive Linguistics
VOICE	Vienna Oxford International Corpus of English
VPC	Verb-Preposition Construction

CHAPTER 1

INTRODUCTION

As the nature of being a human, individuals have always felt the need for communication as part of their lives to be able to express their thoughts and feelings, thereby making languages inevitable. For centuries, languages have served as a means to allow individuals to interact with each other. They have performed as cornerstones of communication, arising out of human conversations. Some of the languages have evolved, some have died, but all of them have always changed.

No matter how many languages have evolved, people from different language backgrounds feel the need to interact for various reasons to convey their messages, to get them across to their interlocutors. Thus, a lingua franca has had to develop out of situations where speakers of different first languages need to talk (Mauranen, 2012, p. 15), mostly preferred as a medium of communication. With the advent of the globalization and internationalization of English (Saxena & Omoniyi, 2010) as 'the world language' (Crystal, 2003; Seidlhofer, 2004) and 'the global language' (Crystal, 2003; Graddol, 1997, 2004), English has been widely acknowledged as 'a lingua franca' (Jenkins, 2009; Seidlhofer, 2005) allowing individuals with different language backgrounds to maintain mutual understanding.

Whether due to the variations and play within the language as part of English as a lingua franca (ELF) spoken by non-native speakers of the language or due to the novelties arising from the use of native speakers of the language, in line with different uses, forms and functions of linguistic features, languages have come to possess various substructures related to syntax and different metaphorical formulations related to semantics. Thus, one of these language developments coming

up with the play performed by native or non-native speakers with the language itself is the use of formulaic expressions, named in general as ‘non-literal language’ (Rumelhart, 1993).

1.1 Background to the study

Formulaic expressions have been commonly accepted to constitute a considerable number of daily conversations as part of non-literal language in the interactions of individuals. Related to these expressions, one can refer to a number of studies trying to analyze the concept of formulaic expressions within the figurative language and categorize their processing structures (Ellis, Simpson-Vlach, & Mynard, 2008; Giora, 2002; Giora & Fein, 1999; MacArthur & Littlemore, 2008; Swinney & Cutler, 1979; Traxler, 2012; Van Lancker Sidtis, 2012; Van Lancker Sidtis & Rallon, 2004; Wood, 2010; Wray, 2002; Wray & Perkins, 2000). Among these studies, Traxler (2012) has shed light on non-literal language, emphasizing the existence of a distinction between the standard meaning, founded on how the words in the expression are commonly understood, and the formulaic one, established upon how the utterances are actually construed by the interlocutors drawing different pragmatic inferences to the front. In accordance with this view, Wray (2002) raises the issue of formulaicity, pointing to the word strings being processed without resorting to their lowest level composition. She puts forward a holistic perspective of formulaic expressions.

Moreover, the most substantial formulization of formulaic expressions is suggested by Van Lancker Sidtis (2012), who distinguishes the internal mechanisms of these structures from novel utterances in terms of stereotyped form, conventionalized meaning, sociolinguistic or appropriate contextual usage

conditions, and familiarity in the form of memorial knowledge. With regard to these perspectives on formulaic language, it can be put forward that formulaic expressions can be perceived and used as unitary segments, composed of words or word sequences, incorporating spontaneous interpretations based on the context of the utterances.

Analyzed under close scrutiny by many researchers, it can be said that formulaic expressions encompass various distinctive components. These include idiomatic expressions (Bobrow & Bell, 1973; Drew & Holt, 1988; Gibbs, Nayak, & Cutting, 1989; Titone & Connine, 1999), clichés and fixed expressions (Moiron, 2005; Prodromou, 2007), frozen metaphors (Balconi & Amenta, 2010; Glucksberg, 1989; Littlemore, 2001), phrasal verbs (Alejo-Gonzalez, 2012; Darwin & Gray, 1999; Gardner & Davies, 2007; Konopka & Bock, 2009; Side, 1990), collocations (Shin & Nation, 2008; Teubert, 2001), metaphors (Glucksberg & Keysar, 1993; Lakoff, 1993), speech formulas (Van Lancker Sidtis, 2010) and sentence builders (Wray, 2002) going from the fixed and frozen constructions to the free expressions. Whether these structures belong to the former or latter group does not change the fact that they are used for the purpose of communicative effectiveness and conciseness of the intended meaning instead of the use of many utterances to convey the same meaning. That is to say, all the formulaic expressions are utilized within the stream of conversations at varying degrees. However, the most debated question still remains as ‘to what extent’. This dilemma has been of concern to the researchers in the field. Many analyses have been done, but there are still more to do.

Among the formulaic expressions whose components have been recently identified, phrasal verb structures have attracted a great deal of attention. Researchers are particularly interested in the degree of the use of phrasal verbs in the interactions

of individuals and the variations of these structures in those interactions because of the confusing structures of their subcomponents. A phrasal verb (PV) is conventionally defined as a unified form of a verb and particle that works as a single verb but loses its original meaning to convey a new meaning (Darwin & Gray, 1999). PVs are widely used structures with regard to the conversations of individuals, both native and non-native speakers. To what degree these speakers deploy PV structures in their conversations at full length will be examined in great detail throughout the present study.

1.2 Rationale of the study

The aim of the current study is mainly to analyze the decomposability nature of the PVs by analyzing the frequency of each PV structure through the conversations of both native and non-native speakers of English demonstrated by two types of corpora, British National Corpus (BNC) and Vienna Oxford International Corpus of English (VOICE), comprising naturally-occurring interactions, thereby also assessing the transparency nature of the PVs in those interactions and providing explanatory adequacy to the usage patterns of PV structures in the field.

Since PVs are semantically concise elements, being used instantaneously in the course of conversations for the purpose of efficiency in conveying meaning, it is really difficult to keep track of them. They are not ready-made utterances that speakers of the language recite and employ during their conversations. They are spontaneous instances of actual language usage. Therefore, it is important to keep in mind that real PV usage of speakers is difficult to analyze.

Incorporating an idiosyncratic nature, PVs have become a highly investigated subject matter among a growing number of researchers. The researchers have analyzed these items in terms of their internal syntactic structure, their semantic feature, the ratio of their attainment by the students in classroom environments, their acquisition properties by native speakers, avoidance issues among both native and non-native speakers, and many more aspects (e.g. Azzaro, 2012; Blais & Gonnerman, 2013; Condon, 2008; Dagut & Laufer, 1985; Gonnerman & Hayes, 2005; Hulstijn & Marchena, 1989; Liao & Fukuya, 2002; Matlock & Heredia, 2002; Morgan, 1997). However, there are a few linguists trying to sort out PVs on a continuum based on their lexico-grammatical nature (e.g. Darwin & Gray, 1999; Dirven, 2001; Gardner & Davies, 2007; McCarthy, Keller, & Carroll, 2003) and only a few books on the subject of PVs being analyzed such as that of Rudzka-Ostyn (2003), all of which will be elaborated in the literature review part of the thesis.

Although the previously mentioned studies have illuminated PV structures in terms of many aspects, a corpus study comparing PV use between native and non-native speakers of English is required in the linguistic literature to provide an explanatory perspective for the analysis of decomposability and transparency distinctions of the use of PVs within a contrastive dimension. Thus, this study aims to fill this gap. It involves an exploratory corpus-based study on PVs being analyzed through two corpora representing two ends of the speakers of English, one as the norm-providing, native speakers, and the other as the norm-following, non-native speakers. It will yield a comparative analysis for English PVs, depicting natural occurrences, similarities and differences between natives and non-natives, thereby providing an explanatory adequacy for the phenomenon.

1.3 Significance of the study

The current study is significant in many respects, contributing to the PV literature in various ways. First of all, it evidences a corpus-based analysis of PV structures, which was conducted by a very limited number of studies. Since PVs are spontaneous instances of actual language usage, PV structures require corpus studies to be analyzed in depth. As this is a corpus-based study, the study deals with authentic oral communication, which eventually indicates spontaneous actual language use of PVs.

There is no controlled processing at work since PVs are analyzed in actual conversations. There is no intermediary position, either, such as teachers, exams, assessments as the confounding variables, affecting the interactions of the speakers, which provides the opportunity for the study to deeply examine the context in which the speakers are fully immersed in natural interactions, far away from anxiety. In line with this objective, the recognition factor of the target items by the speakers has been prevented. In this way, when the recognition of the use of PVs is eliminated from the conversations of the speakers, this leads to natural usage patterns of PV structures. Thus, these are meaningful interactions among themselves.

The significance of the current study, therefore, lies in the fact that as the method of investigation, it has not used any test types or tests which normally bring the second language (L2) knowledge of the learners to the forefront of their consciousness; yet, in daily oral communication, procedural (implicit) knowledge is more often activated than explicit conscious knowledge (Ullman, 2004). That is to say, the natural language use of PVs is actively involved in the actual language productions of the speakers, represented by corpora. Then, in the study, the corpus, in this case VOICE, has been scanned to detect and analyze natural uses of PVs in

oral communication between nonnative speakers without any native speaker either priming something or intervening as a measurement tool, as a guide, or as a yardstick that would activate conscious explicit knowledge. The interaction is more or less based on implicit knowledge of the language. In other words, their linguistic performance is whatever they have in terms of linguistic competence as Chomsky (1965) emphasized.

Apart from this unique nature of the study, what has equal importance is that BNC has been also scanned to provide native speaker stance for the natural uses of PVs. BNC has similar characteristics when compared to VOICE, showing difference only with respect to its speakers, being native. BNC shows the other side of the PV medallion, depicting the natural occurrences of language use among natives, perceived as the norm-providers of the language. Thus, the current study attempts to investigate the PV use of non-native speakers in comparison to that of native speakers based on corpus evidence.

From a contrastive point of view, the current study is important in addressing the decomposability and transparency paradigms in the analysis of PVs. Linguistic processing of PV structures undergoes both syntactic and semantic processing. With the aim of reaching a full-fledged analysis, it is essential for the study to examine PV structures through both of these processing systems. Thus, it analyzes PV structures within the scope of decomposability and transparency nature on a continuum, thereby further pointing out these features with additional implications.

It is also worth mentioning that the study will be an important step within the ELF framework. Since ELF has witnessed an unprecedented spread of English on the world scene, it has ended up with a lot of variations among the different structures of language. Regarding these structures, the current study comes in handy in the field of

PV processing with reference to the classification of these variations as part of ELF framework if there are any, and indicates the similarities or differences of PV usage in comparison to English as a native language (ENL) use. Therefore, the realization of ELF framework in the field necessitates an investigation in its own right, proposing a socio-cognitive approach. In this respect, this study becomes prominent in illustrating this approach to PV structures from an ELF perspective by providing the decomposability and transparency analyses of PVs through the interactions of ENL and ELF speakers.

Last but not least, the current study is of vital importance since it proposes an alternative approach to the PV analysis within the scope of decomposability and transparency distinctions of the use of PVs. The thesis formulates a new approach 'Nondecomposable Lexical Annotation' (NLA), which has been formed based on the results of the corpora analyses of the study. NLA puts forward that there is an inclination to process and use PVs in a nondecomposable position and with the involvement of lexical annotations based on the initial conceptualization of their opaque meaning. Thus, NLA is significant in bringing a new point of view to the use of PV structures.

1.4 Definition of the main concepts

Phrasal Verb (PV): It can be defined as multi-word expression, a phraseological entity, a structure including a lexical verb and an adverbial particle (either adjacent or not to the main lexical verb), losing the meanings of both of its constituents, thereby forming a new lexical unit. 'Put off', 'give up', 'take in', and 'move on' can be given as examples.

Decomposability: It refers to whether the particle constituent of a PV can be obligatorily adjacent to its lexical verb or not. Decomposability shows a cline in three stages in syntactic terms. 'Non-decomposability' is the situation in which PV has to have its lexical verb and particle always adjacent, not accepting any object, pronoun, adverb or any other lexical entity in between. 'Look into' and 'go through' can be given as examples to nondecomposable PVs because their lexical verbs and adverbial particles are always used in an adjacent position. 'Decomposability' is the situation in which PV can be witnessed having different forms in the use of lexical verb and particle (e.g. any object such as pronoun, noun or phrase can intervene between lexical verb and particle). 'Pick up', 'throw away' and 'turn down' can be given as examples to decomposable PVs because of their internal characteristic involving syntactic flexibility of lexical verb and particle. 'Both' stage is the circumstance of PVs suitable for the uses of both decomposability and non-decomposability. 'Bring up', 'try on' and 'fill out' can be given as examples.

Transparency: It refers to whether the meaning of a PV is literal or figurative. If the meaning of the PV can be inferred from the combination of the literal meanings of lexical verb and adverbial particle, then the PV can be denoted as transparent. 'Go back' and 'come in' can be given as examples to transparent PVs. If the constituents of PV have figurative meaning and the meaning of the PV cannot be inferred from its constituents, then the PV can be denoted as opaque. 'Let down' and 'take over' can be given as examples to opaque PVs. If the constituents of PV have both literal and figurative meanings, the PV can be denoted as semi-transparent as an intermediary structure. 'Come down' and 'take off' can be given as examples to semi-transparent PVs since these PVs have both transparent and opaque meanings in terms of their semantic nature.

Nondecomposable Lexical Annotation (NLA): This is an alternative approach which has emerged out of the findings of the thesis based on the analyses of the corpora within the scope of decomposability and transparency nature of the use of PVs. This approach mainly proposes that PVs are processed as holistic units and used in a nondecomposable position though they are even perfectly grammatical in a decomposable position in nature. Nondecomposable is accepted in terms of its syntactic approach. PVs are considered to possess lexical annotations comprising transparent or opaque characteristics, but when they are produced, the initial activation of opaque meaning is used in lexical annotations. As a semantic approach, lexical annotation is accepted. Therefore, NLA has arisen as an evidential finding of the thesis.

English as a Lingua Franca (ELF): With the advent of the globalization of English as the world language, ELF serves as a new phenomenon about the usage of English mostly signifying a preferred medium of interaction, in which two speakers of English with different mother tongues choose to speak in their common language, English, in order to maintain the mutual understanding in their conversation. ELF has been defined by Jenkins (2009) as "English being used as a lingua franca, *the common language of choice,* among speakers who come from different *linguacultural* backgrounds" [emphasis added].

Corpus: It refers to large collections of texts, comprising spoken or written data, sampled from various speakers with different backgrounds, mostly transferred to an electronic system for concordances to yield further linguistic analyses. A corpus is clearly defined by Sinclair (2004) as "collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language".

Multi-word expression: It refers to a unit of lexical structures, a meaningful unit of words, possessing some properties which are not predicted with its components individually, ranging from phrasal verbs to idioms and even proverbs. Grant and Bauer (2004) define a multi-word expression as "a fixed and recurrent pattern of lexical material sanctioned by usage" (p. 38).

Verb-Preposition Construction (VPC): It refers to a verb and a preposition, being processed separately, and not recognized as a chunk. The preposition only gives meaning to the object of the verb, and does not affect the meaning of the verb in any terms. ‘Look at’, ‘depend on’, ‘suffer from’ and ‘listen to’ can be given as examples to VPCs since these verbs retain the same meaning without the use of prepositions, but these prepositions have to be used in the presence of the objects of the verbs.

Socio-Cognitive Linguistics (SCL) / Socio-Cognition: It refers to the study and analysis of social interactions or situation of discourse with regard to cognitive approaches. Socio-cognition can be designated as a meaningful synthesis of sociolinguistics and cognitive linguistics. Kristiansen and Dirven (2008) define socio-cognitive linguistics as "a re-contextualization of social dimensions explained with a cognitive approach".

1.5 Organization of the thesis

The thesis is organized as follows: Chapter 1 introduces the topic of the study. It puts forward the general introduction and background to the study in a broad spectrum. It further explicates the rationale and significance of the study, depicting the reason why the current study has been conducted. The definitions of the main concepts are also situated at the end of Chapter 1.

Chapter 2 illustrates the ELF framework. It explains the globalization of English and gives information about the spread of English. It expresses the terms English as a Second Language (ESL) and English as a Foreign Language (EFL). Chapter 2 also clarifies the layers of analysis in ELF framework, describing the conceptualization of ELF. Then, it further outlines the previous studies conducted in the ELF literature.

Chapter 3 deals with the issue of decomposability. It analyzes the decomposability phenomenon in first language and second language use. It details the previous studies conducted in the decomposability literature. It lists the hypotheses trying to clarify the decomposability paradigm put forward by researchers.

Chapter 4 examines phrasal verb (PV) structures. It outlines multi-word unit components such as idiomatic expressions and collocations. Then, it reviews a number of different perspectives on the linguistic processing of idiomatic expressions, listing the earlier research carried out in the literature. It also places a special emphasis on idiomatic expressions in ELF encounters. Since analyzing all the multi-word unit components in one study is not a possible thing to realize, the study centers upon PVs. Having its focus on PV structures prompts the current study to propose its own understanding of what PVs really are. In line with this objective, the study clearly explicates the meaning and usage distinction between the structures ‘PVs’ and ‘verb-preposition constructions (VPCs)’. It further discusses PV taxonomies in the literature. After presenting the theoretical background on PV structures adopted by the previous studies, the study depicts the use of PVs in ELF and British English, giving a comparative analysis between natives and non-natives.

Chapter 5 proceeds with the data and methodology. It proposes its research questions and hypotheses. It presents participants referring to the corpus speakers as the native speakers in BNC and the non-native speakers in VOICE. Then, it delves into the corpus data, displaying the internal structures of the two corpora on which the study has based its research. It also outlines the procedure of the research. Lastly, the data analysis section in Chapter 5 proposes the steps of analysis to indicate what has been involved in the examination of the corpus data. Besides, Chapter 5 explains the study's own approach, NLA, to the decomposability and transparency distinctions of the use of PVs.

Chapter 6 illustrates the results of the study, pointing to the findings of the analysis in the research. It depicts the frequency of the target items, namely the PVs in comparison to VPCs. It further illustrates the findings of syntactic analysis, expanding on the decomposability nature of PVs. It also provides the findings of the semantic analysis of PVs, drawing attention to the transparency nature of PVs.

Chapter 7 lays out the discussion part of the study. It ponders upon the hypotheses the study proposes and their congruence with the results of the study. It addresses the issue of the use of PVs from the viewpoints of native and non-native speakers.

Chapter 8 designates the key ideas and general understanding of the main concepts in the study. It summarizes the main findings of the study. It further discusses the pedagogical implications of the study. It lists the limitations in the research. Finally, it makes recommendations for future studies.

CHAPTER 2

ENGLISH AS A LINGUA FRANCA

English is an international language, playing a crucial role in the interactions of individuals who are using it either as their native language or as their second or foreign language, growing in importance at the global level. Within the last three decades, English has incorporated several terminologies, particularly, ‘the global language’ (Crystal, 2003; Graddol, 2006; Pennycook, 2007), ‘the dominant language’ (Phillipson, 2003), ‘the world language’ (Crystal, 2003; Graddol, 1997, 2004; Kachru, 1992a, 2005; Seidlhofer, 2004), ‘the powerful language’ (Mesthrie & Bhatt, 2008). All these conceptualizations of English lead us to raise the question, "Why English?". It can be accepted that by means of advancements in the realm of the mass media, economy, education, technology through the medium of English, English secures its position, making its omnipresence impossible to neglect. This circumstance leads to a necessity to review the status of English in the world with regard to its globalization and internationalization.

2.1 Globalization of English as a world language

Although there are many divergent perspectives on what globalization means, the convergent thinking on the term ‘globalization’ can indicate the phenomenon that globalization causes the world, which individuals live in, to decide on a common language for worldwide communication. Thus, globalization leads to change in many domains of life such as economy, technology, international trade, educational policies, politics and many others throughout the world. As a result of globalization, English emerges as a common language of communication all over the world.

Scholte (2000) has categorized globalization into different spheres, namely, globalization as internationalization, liberalization, universalization, westernization or modernization, respatialization, proposing globalization as a reconfiguration of social space where the local is constantly interacting with the global. Concerning the local and global interactions, Dewey (2007) puts forward the importance of having an interconnected position for globalized language. Therefore, English as the mostly preferred, the most common and highly involved in many aspects of nearly all the countries is the global language, being widely used all around the world.

With its unprecedented spread in the world, English is 'the' global language. It is used to manage international communications for countries as the medium of language in many areas, mainly political issues, educational systems, and economic spheres. Crystal (2003) lists two conditions which are necessary for a language to be global, namely being an official language and being taught as a foreign language. Regarding these conditions, English as the global language has the status of being an official language for more than seventy countries, including Ghana, Nigeria, India and Singapore, which makes itself the global language. English has the highest prevalence as the global language, being incorporated into foreign language teaching programs of the countries, the language of which does not have an official status.

To briefly outline why English is the global language, the criteria of predominance of a language in many domains in the world and being both official language and preferred foreign language to manage mutual communication for the countries can point out the globalization of English. Global English yields a necessity to analyze the phenomenon of the spread of English in detail, which is perceived differently by several researchers.

2.2 The spread of English as ESL and EFL

With globalization and internationalization of English (Crystal, 2003; Dewey, 2007; McKay & Bokhorst-Heng, 2008; Omoniyi & Saxena, 2010), English is the most widely spoken and preferred language in the world. Besides the ubiquitous status of English in the developed countries, English, as a super-strate language (Mesthrie & Bhatt, 2008), has spread its use and depth to many developing countries. This perspective indicates the fact that there are more non-native speakers of English than native speakers.

Considering the unprecedented speed of its becoming prevalent as the global language, English will proceed to spread in a far-reaching manner more than ever in the future, according to what has been put forward in the studies by Kachru (2005) and Yano (2001), calling for the use and spread of English in both ESL (English as a Second Language) and EFL (English as a Foreign Language) settings. These settings have varying degrees of influence, posing different kinds of approaches to classify the domains through which English has been situated in terms of 'the circle of World English' by McArthur (1987), 'circle model of English' by Görlach (1990), 'three circles model' by Kachru (1992b), as Inner Circle (ENL), Outer Circle (ESL) and Expanding Circle (EFL), and lastly 'diaspora model' by Omoniyi and Saxena (2010).

Three scholars have tried to formulate the models for the characterization of the spread of English language during the same years more or less with the aim of forming a conceptual map as a way of clarification. First, McArthur (1987) developed the circle of World English model, situating English in one wheel being composed of several subcomponents. In his model (Figure 1), World Standard English is in the center, representing the idealization of English as the standard form in written international English and Received Pronunciation.

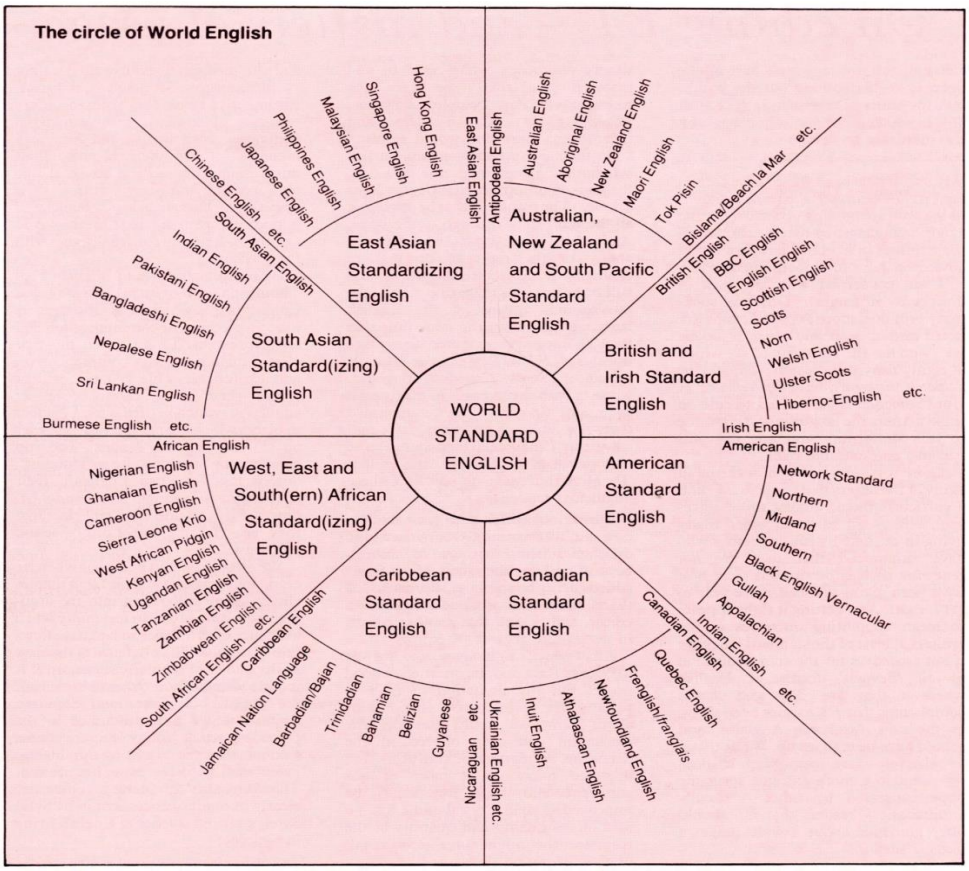

Fig. 1 Tom McArthur's circle of World English

(McArthur, 1987)

Though there are some small variances in the written norms of British English and American English, World Standard English illustrates the skeleton of English. As can be seen in Figure 1, the next layer around the center is divided into 8 subcomponents. These subcomponents are composed of different Standard English and Standardizing English. Standardizing English varieties indicate that they are not standard for the time being but they are on the verge of transforming into a standard one. The next outermost layer consists of localized varieties. These varieties can be regarded as possessing resemblances with their adjacent layer including standard and standardizing Englishes. In his article in *English Today*, McArthur (1987) states his aim in developing this circular representation of World English as "to highlight the broad three-part spectrum that ranges from the *'innumerable' popular Englishes*

through the various national and regional standards to the remarkably homogenous but *negotiable 'common core'* of World Standard English" (p. 11) [emphasis added].

McArthur (1987) describes the spread of English and different varieties of English in his circle of World English presenting the position of English at that time; yet, the interpretation of the model raises questions about ENL, ESL and EFL distinctions. In the core, the placement of World Standard English brings American and British English to mind as the standard forms of English, representing ENL. However, the penultimate layer which is placed round the core also includes American and British English. That is to say, both ENLs and ESLs are situated in this intermediary layer. This fact discloses the idea that McArthur (1987) thought that standard English and standard(izing) ones should belong to the equal layer. However, the situation that they are recognized at the same layer also raises questions for the ESL varieties and ENLs in the literature. In fact, it is an interesting side of the model to regard the term standard(izing) with parentheses to imply that standardizing varieties are in the process of standardization to be seen as standard. Apart from that, the circle of World English by McArthur (1987) has localized varieties of the languages presented in the penultimate layer divided into branches as the outermost layer. With respect to these branches in the circle, McArthur (1987) emphasizes ". . . because of the fluidity and fuzziness wherever the language is used, the demarcation lines are all discontinuous, and at the outer limits of the 'circle' the circumference is open to intermingling with other languages . . . " (p. 11). Hence, this leads to the interpretation that McArthur has designed his representational model of World English by thinking that English varieties are in interaction with the other varieties, changing them and being changed by others, so they should be represented within discontinuous lines to denote the cross-linguistic influence among the

varieties. Nevertheless, the utmost layer also displays a mixture of different types of English as ESL varieties and even pidgin and creoles all together. Thus, these issues arouse a feeling of clarification for the model in terms of the representations of ENL, ESL and EFL distinctions.

With the beginning of the curiosity into language, English, and its spread and prevalence in the world, other scholars also developed different models of English. Görlach (1990) was one of them. Görlach (1990) established his own model of English, namely circle model of English, situating English into different layers of circles, each of which denotes separate English varieties and distinctions. In his circle model of English (Figure 2), International English is in the center of the circle.

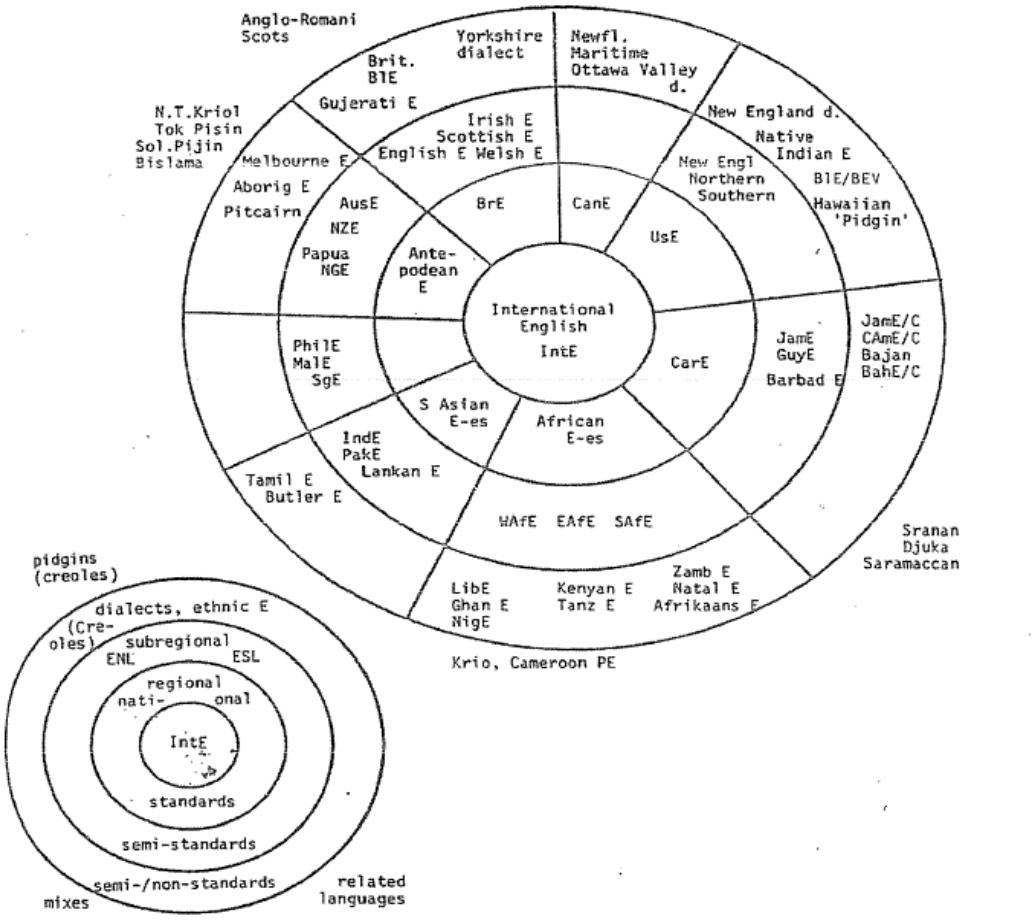

Fig. 2 Manfred Görlach's circle model of English

Note: (Görlach, 1990). Adapted from Mesthrie & Bhatt (2008)

The circle is surrounded by regional and national English as the standards, subregional ENL and ESL as the semi-standards, dialects and ethnic English as semi- and non-standards within nested circles intertwined and also surrounded by pidgins, mixes and related languages around the circle at the outermost part. As can be seen in Figure 2, International English being in the center of the circle is enclosed by the next layer involving a range of Englishes from British, Canadian and American English to South Asian Englishes, as depicted in the holistic interpretation of the circle in Figure 2 (adapted from Mesthrie & Bhatt, 2008). It can be interpreted that Görlach (1990) thought both regional and national English varieties would form International English as the backbone of the circle model of English.

Compared to the model 'the circle of World English' by McArthur (1987), 'circle model of English' by Görlach (1990) illustrates a similar way of representation and categorization of English to that of McArthur (1987). Taking International English as the center, ENLs and ESL varieties are situated in the next layer. They are also surrounded by subregional English varieties which are actually elaborations of the aforementioned layer. Then, the next layer is also composed of the explication of the varieties, being placed in terms of their dialects. In addition, though they are not even inside the circle, as part of the shape the utmost part include pidgins, creoles and mixed languages surrounding the circle. This means although they are not part of the circle, they still belong to English. Regarding the circle, it can be interpreted as the cone-shaped illustration of the English languages with all elaborations of the varieties. Thus, in comparison to the model devised by McArthur (1987), Görlach (1990) provides a more visualized form of English languages and their varieties in terms of their categorizations, indicating a more consistent illustration, but it still does not include EFL varieties, requiring a clearer depiction.

Moreover, Kachru (1992b) proposed a conceptualization of a model of English more broadly and differently as opposed to the representational models (the aforementioned models by McArthur, 1987 and Görlach, 1990) of other scholars at the time when they also proposed different models of World Englishes. Kachru (1992b) constructed 'three circles model' involving three concentric circles, which are actually no longer concentric, representing the type of the spread of English, the historical and political factors of the use of English across countries. These circles include Inner, Outer and Expanding Circle, which can be seen in Figure 3.

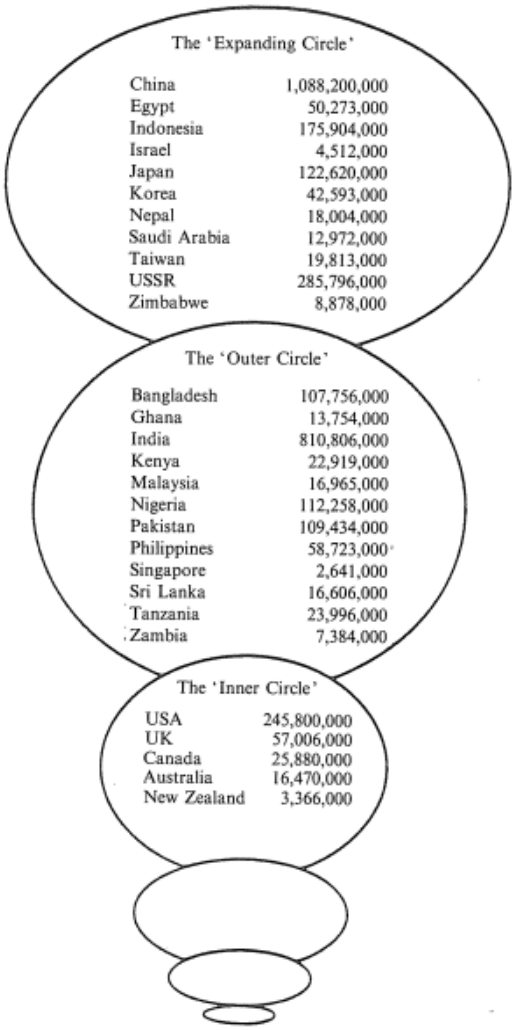

Fig. 3 Braj Kachru's three circles model of English

(Kachru, 1992b)

According to Kachru (1992b), these circles of World English are created with respect to 'the historical, sociolinguistic and literary contexts'. The Inner Circle involves 'norm-providing' varieties, which constitute ENLs (English as a native language) (Bhatt, 2001; Kachru, 1992a, 1992b, 2005; Mesthrie & Bhatt, 2008). It involves the traditional, cultural and linguistic bases of English (Kachru, 1992b). In terms of varieties, Kachru (1992a) has considered American English and British English more appropriate than the varieties used in Australia and New Zealand in the Inner Circle. However, Kachru (1992b) has also redefined Inner Circle countries as USA, UK, Canada, Australia and New Zealand. Besides, under the Inner Circle, the other empty three circles are also eye-catching, suggesting the core of English.

The Outer Circle involves 'norm-developing' varieties, which comprise ESLs (English as a second language) according to Kachru (1992b). They symbolize the institutionalized non-native varieties in the regions which went through the extended periods of colonization (Kachru, 1992a). ESL users do not possess "identical attitudes about an endocentric (locally defined) norm" according to Kachru (1992b). The users of these varieties have some confusion between linguistic norm and linguistic performance, but they have developed their norms and these localized norms reflect well-established linguistic and cultural identities (Kachru, 1992b). Singaporean English, Nigerian English and Indian English can be given as examples. However, although the speakers in the Outer Circle have formulated their own spoken norms, they show an inclination to depend on written forms of English for the linguistic standards.

The Expanding Circle involves 'norm-dependent' varieties, which includes EFLs (English as a foreign language) according to Kachru (1992b). EFL users have not developed internal norms according to Mesthrie and Bhatt (2008), and they rely

on the external norms, so they are norm-dependent. Kachru (1992a) has also emphasized the status of English in the Expanding Circle in his book *The Other Tongue* since EFL varieties lack official status and only used in restricted terms. According to Kachru (1992a), in the Expanding Circle, the performance varieties of the language have been used in EFL environments. What can be understood from the Expanding Circle is that in this circle, English is spoken as the 'foreign language' rather than the 'second language' as in the Outer Circle. In the Expanding context, English is important for international communication and in education such as for further studies in the universities or in the academic life of learners.

There is also a diasporic representation by Omoniyi and Saxena (2010) as the extended and reconsidered version of Kachruvian Circles. Omoniyi and Saxena (2010) propose the existence of "three broad diaspora Englishes that are relevant to our understanding of the interface between the sociolinguistics of colonization and that of globalization" (p. 4) as the representational vision of the English language. Diaspora Type 1 is composed of ENL speakers whose native language is English and who have relocated from an English-speaking homeland, termed as "speakers of neo-local Diaspora English" (Omoniyi & Saxena, 2010) such as Canadian English, Australian English and New Zealand English. Interestingly, they do not mention British and American English as belonging to Diaspora Type 1. According to Saxena and Omoniyi (2010), English is the 'de facto' language.

Diaspora Type 2 consists of "the Englishes that have emerged out of colonial enterprise" (Omoniyi & Saxena, 2010, p. 4), which include ESL speaking countries. Omoniyi and Saxena (2010) give the example of former British colonies like Nigeria, Kenya, Jamaica, India and Hong Kong, and also former colonies of the United States such as the Philippines, Guam and Porto Rico. In these countries belonging to

Diaspora Type 2, English language has the 'institutionalized' variety, meaning English is formally used in the educational, civil, administrative and governmental domains of the countries.

Diaspora Type 3 consists of EFL speaking countries such as China, Japan, South Korea, Taiwan and many others. In these countries, English has developed with the consequence and the effect of the global market place, where English does not have the historical or governmental role but does have an important place in international communication.

What is striking about Omoniyi and Saxena (2010) is that they do not include American and British English to any Diasporas, which may indicate that they perceive American and British English as a super-strate form of language as suggested in Mesthrie and Bhatt (2008). It can be interpreted that American and British English are perceived as in the core of their Diaspora representation of English, separated by the other varieties whether they belong to ENL, ESL and EFL.

Among these stratifications of English, what sounds more plausible belongs to Kachru (1992b). Since English has been developing constantly, the countries that previously seemed as norm-developing may actually attain a role of norm-providing regarding the developments in English. I agree with Kachru (2005) and Yano (2001) on their view of redefinitions of the boundaries in 'Kachruvian Three Circles'.

Kachru (2005) has emphasized the distinction between the concepts of nativeness and otherness. What Kachru (2005) has put forward is important because the necessity for a reconceptualization of the distinction between genetic nativeness and functional nativeness. Kachru (2005) proposes that genetic nativeness includes ENL as Inner Circle varieties. There must be a distinction as functional nativeness which further points out that although English was not the ESL countries' actual

native language, it has developed to become like one, not only within the genetic geographical mapping but also within the dimension of the language in a society like range and depth of the language as the status of the language. Thus, Kachru (2005) has stated that functional natives have the nativeness of the language according to all the domains of the language such as genetic mapping, range, depth, types of acculturation of languages, the creative processes used at various levels of local identities. Therefore, according to Kachru (2005), there is a change and no strict borders between the Inner Circle and the Outer Circle. By explaining this, he has also revised what he proposed in his 1992 article as three ‘concentric’ circles because, although they are from the same origin, they are no longer concentric.

Similarly, Yano (2001) supports the view of the reconceptualization of the nativeness by Kachru (2005), and he has interestingly pointed out that the "functionally native ESL speakers in the outer circle are expected to far exceed those genetically native English speakers in the inner circle not only by their numbers but also by their economic and technological power" (Yano, 2001, p. 122). Yano (2001) also indicates the fuzziness between the Inner Circle and the Outer Circle.

Most importantly, the idea of GLOCAL language put forward by Yano (2001) is the most plausible formulation in my opinion. According to Okushima (1995, as cited in Yano, 2001), GLOCAL means ‘global and local’. He proposes that glocal involves being international and retaining the local self-identity as well. It is also suggested in Yano (2001) that "GLOCAL language has an international status in its spread but expresses local conditions in the Kachruvian outer circle and expanding circle." Thus, through the developments and changes (Kachru, 2005) and an international spread (Yano, 2001) of the English language, the idea that English will be the glocal language of the future can be set forth in line with these studies.

2.3 The beginnings of English as a lingua franca phenomenon

English as 'the' global language has been not only influential in the countries where it is spoken as a native language, i.e., ENL, but also has been influential in the countries where it is a foreign or second language, i.e. EFL and ESL. In EFL/ESL contexts, English functions as an institutionalized variety through which speakers communicate in various domains such as education, administration, and media.

With the developments in English, the circles are also reconceptualized and redefined by Kachru (2005) since the people speaking ESL in Outer Circle have also developed their norms of English and have been regarded as functional native speakers of English (Kachru, 2005). This huge spread of English into Outer and Expanding Circle has provided those people from different first languages (L1s) with the opportunity to communicate and maintain mutual understanding in their interactions via a common language, functioning as lingua franca, which in the end formulates English as a lingua franca (ELF) phenomenon.

There are many studies investigating the conceptualization of ELF. In line with the current study, ELF can be defined as a 'shared' language used when two non-native speakers of English – whose native languages are not the same – or one native and one non-native speaker of English interact to communicate and maintain mutual intelligibility. The key issue is not to possess the same native language. The maintenance of a successful communication is of vital importance in ELF.

According to Jenkins (2007), "ELF is a contact language used among people who do not share a first language, and is commonly understood to mean a second (or subsequent) language of its speakers". Seidlhofer (2001, 2004) has stressed the occurrence of ELF interactions when English is the "preferred option for communication" among people from different L1 backgrounds. Jenkins (2009) also

defined ELF as “English being used as a lingua franca, the common language of choice, among speakers who come from different linguacultural backgrounds” (p. 200). This definition is similar to what Firth (1996), House (2003a, 2003b) pointed out ELF interactions as “between members of two or more different linguacultures in English, for none of whom English is the mother tongue”.

I agree with the views of Jenkins and Seidlhofer on the types of ELF interactions that ELF can also involve speakers from native English, but ELF occurs when they are also interacting with other speakers whose L1 is not English, so ELF can occur between both non-native to non-native and native to non-native, in which English is the only way to communicate. Thus, ELF acts between two non-native English speakers of different L1s as ‘tool language’. As for the position of native speakers, Jenkins (2008) states in her conference presentation that "ELF interaction can include native English speakers, but in most cases, it is a contact language between people who share neither a common native tongue nor a common national culture, and for whom English is an *additional* language" [emphasis added].

Since people feel the need to communicate in order to express their feelings and thoughts, there must be a common language for a successful communication. Thus, that is when ELF comes out as the contact language. Many developments in English and other languages around the world also enable ELF to develop for communication and it will be “the glocal language” of the future (Yano, 2001).

Moreover, Seidlhofer (2005) highlights the reason of the conceptualization of ELF as "a way of referring to communication in English between speakers with different first languages" (p. 339). Through the globalization and internationalization of English, almost 75% of non-native speakers of English in the world produce most of the ELF interactions since almost one out of every four speakers of English is a

native speaker in the world (Crystal, 2003; Graddol, 2006). This signifies the reason for the huge percentage of people in the Outer and Expanding Circle varieties (Davis, 2011; Kachru, 2005; Seidlhofer, 2005; Yano, 2001). Thus, the constant influence and spread of English as a lingua franca may lead to the immense increase in the English-speaking individuals as non-native speakers, which requires a need for the reformulation of 'Kachruvian Three Circles'.

In this technological era, with the developments in education, economy, business relations, technology, and especially social media, globalization has evolved and caused English to be used as the most common 'communication' language as the first choice, allowing an increase in ELF interactions all over the world. Thus, ELF will continue to be the main tool language for getting to the common ground among the speakers who want to maintain mutual intelligibility.

2.4 Description of the ELF framework and its layers of analysis

As a consequence of the continuing process of development and spread of English as a means of communication, the ELF framework has burst onto the scene as the dominant lingua franca at the global level throughout the world. Since the ELF framework brings billions of speakers around the world together, this has attracted a lot of attention.

Several researchers have investigated the ELF context, ELF interactions and the influence of ELF on many different domains of life in the world (Dewey, 2007; House, 2003a; Jenkins, 2006, 2007, 2008, 2009; Jenkins, Cogo, & Dewey, 2011; Pickering, 2006; Seidlhofer, 2000, 2001, 2002, 2004, 2005, 2006, 2009a, 2009b; Seidlhofer & Berns, 2009; Seidlhofer, Breiteneder, & Pitzl, 2006; Wacker, 2011).

Many other researchers have focused on one specific point in ELF and have made investigations (Breiteneder, 2005, 2009; Cogo, 2008; Mauranen, Hynninen, & Ranta, 2010; Saraceni, 2008; Smit, 2010), and there are many other researchers to count.

ELF has been debated so much that while Seidlhofer (2001, 2004, 2005) supported that there is a need to find a conceptualization of ELF and native – non-native speaker distinction in the conceptual gap, and tried to close this gap, Jenkins (2006, 2007, 2008) analyzed the distinctions between World English and ELF (cf. Seidlhofer & Berns, 2009) for current perspectives on teaching ELF and what the attitudes and perceptions of people about ELF are. After this conceptualization process by Seidlhofer (2006) clarifying misconceptions about ELF, many studies started to flourish about ELF interactions in terms of globalization, multilingualism, business, education, form and function in ELF, only a few as its layers of analysis.

ELF has been investigated within the globalization perspective by Dewey (2007). Dewey (2007) emphasized the linguistic diversity in ELF from a multi-dimensionalist perspective and stated that ELF is in a continuous change. While the effect of globalization is on its rise, House (2003a) investigated whether ELF is a threat to multilingualism, stressing that ELF works hand in hand with multilingualism, as complementary to the development of the local varieties.

In terms of the distinction of the conceptions in ELF, there are studies conducted by Pakir (2009) and Seidlhofer (2009b) analyzing EIL (English as International Language), WE (World Englishes) and ELF. About the educational domain in ELF, there are many investigations (Björkman 2011; Decke-Cornill, 2002; Griva & Chostelidou, 2011; Knapp, 2011; Kuo, 2006; Ljosland, 2011; Mauranen, Hyninnen, & Ranta, 2010; Smit, 2010). Björkman (2011), Knapp (2011), Kuo (2006), and Smit (2010) analyzed ELF with a focus on higher education in teaching

ELF. Mauranen, Hyninnen, and Ranta (2010) investigated English as an Academic Lingua Franca (ELFA) within the ELFA project[1]. In terms of the business domain in ELF, Bohrn (2008), Louhiala-Salminen and Kankaanrata (2012), and Pitzl (2005) investigated ELF in the business context. Louhiala-Salminen and Kankaanrata (2012) discussed ELF in the business sector as Business English lingua franca.

Lastly, about form and function, Saraceni (2008) questioned the form and functions of ELF on different levels, namely, functional, cultural and phonological, lexico-grammatical. Saraceni (2008) emphasized that there should be functional focus on ELF rather than heavily on form. Cogo (2008) also referred to what Saraceni (2008) discussed and she supported the importance of both form and function. In terms of forms in ELF, there are three linguistic levels according to Jenkins, Cogo and Dewey (2011), namely, phonology, lexis / lexico-grammar, and pragmatics. From a functionalist perspective, Seidlhofer (2002) examined ELF as endonormative in terms of being norm-developing, not exonormative in terms of being norm-dependent, thereby viewing ELF on an empirical base.

All in all, what the current study has aimed to examine in the ELF framework is to analyze the target items of the study by incorporating both form and functionalist perspective, thereby providing an explanatory adequacy to the phenomenon within the scope of the ELF framework.

[1] For further information, please visit http://www.helsinki.fi/englanti/elfa/project.html

CHAPTER 3

DECOMPOSABILITY

Decomposability can be identified as whether the parts of a whole are connected to each other as the constituents in a fixed and stable way, especially on syntactic terms without undergoing any change. These parts can be separated from each other by other lexical items or phrases without losing any meaning and usage, belonging to the same whole in a loose and flexible connection. The term has been highly used in linguistic analyses of the studies and it can be seen as a heavily debated issue among scholars since a consensus could not be reached in their analyses on how the linguistic item shows decomposability. Because of its fluidity and fuzziness, there is no single viewpoint about the decomposability feature in the literature, but there are many to count, in need of clarification and mutual agreement (e.g. Bobrow & Bell, 1973; Caillies & Declercq, 2011; Fischer & Keil, 1996; Gelder, 1990; Gibbs, 1992; Gibbs et al., 1989; Grant & Bauer, 2004; Kamp & Partee, 1995; Lewis, 1993; Molinaro, Canal, Vespignani, Pesciarelli, & Cacciari, 2013; Sag, Baldwin, Bond, Copestake, & Flickinger, 2002; Titone & Connine, 1994; Van Lancker Sidtis, 2012; Wray, 2012). However, for the sake of presenting a concise perspective and a consolidated account, all the explanations proposed by many researchers will not be explicated here one by one, but the stance of the thesis for the decomposability feature will be explained and a core definition of decomposability will be presented and addressed towards the end of the chapter by pondering on a different kind of approach to the decomposability distinction.

The general statement about decomposability nature in the linguistic field to be mentioned can signalize the situations in which whether or not there is a

constraint on the use of the items to their adjacent ones being involved within the context of the utterance or not. If the constraint exists on the use of the items, the items are called ‘non-decomposable’. However, if there is not a constraint on the use of the items, these are called ‘decomposable’.

3.1 Decomposability issues in L1 and L2 use

Decomposability feature has been studied in terms of its influence on the use of structures in the first language (L1) and second language (L2) of individuals. In broad terms, these structures have been mostly involved in the investigations of formulaic expressions, namely idiomatic expressions, phrasal verbs and collocations. The decomposability nature serves as part of the syntactic analysis in the literature.

Incorporating two dichotomous categories, namely, holistic and analytical examination of formulaic expressions, Wray (2002) analyzes formulaicity in L1 acquisition of children and L2 acquisition of very young learners, older children, teenage and adult learners. Wray (2002, p. 206-209) makes a special emphasis between the first language acquisition and the classroom-taught L2 (after childhood), illustrating these two in separate models, indicating the holistic processing of formulaic expressions. This points out that she has approached the formulaicity by adopting a non-decomposable perspective, because of her holistic attitude in the analysis. Furthermore, Wray (2012), with a reformulated point of view, draws attention to certain features associated with formulaic language and formulates a model on the issue (Figure 4) as the complex discussion space of formulaic language.

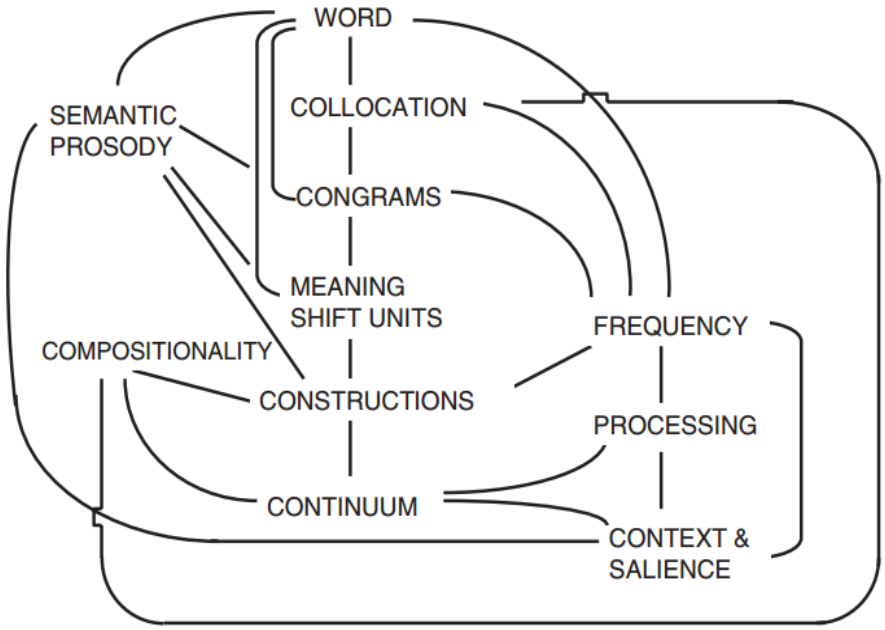

Fig. 4 The complex discussion space of Wray's conceptualization on formulaic language

(Wray, 2012, p. 238)

In her review article, Wray (2012) interestingly divides the decomposability[2] feature into two parts as constructions and continuum which can be seen in Figure 4. Based on this figure, we can assume that if the structures are perceived with a non-decomposable perspective as a whole, they will yield to the creation of constructions, which in the end denotes frequency analysis. However, if these structures are perceived with a decomposable view, which can make analytical analysis possible, then they will yield to the continuum, calling on processing, context and salience features for further analysis. Wray (2012) regards decomposability as a factor combined with frequency, having influence on the formulaicity of the structures within the scope of their faster processing, being further indicated in Figure 5.

[2] 'Decomposability' and 'compositionality' are synonyms. For the purpose of attaining coherence, the term 'decomposability' will only be used throughout the study.

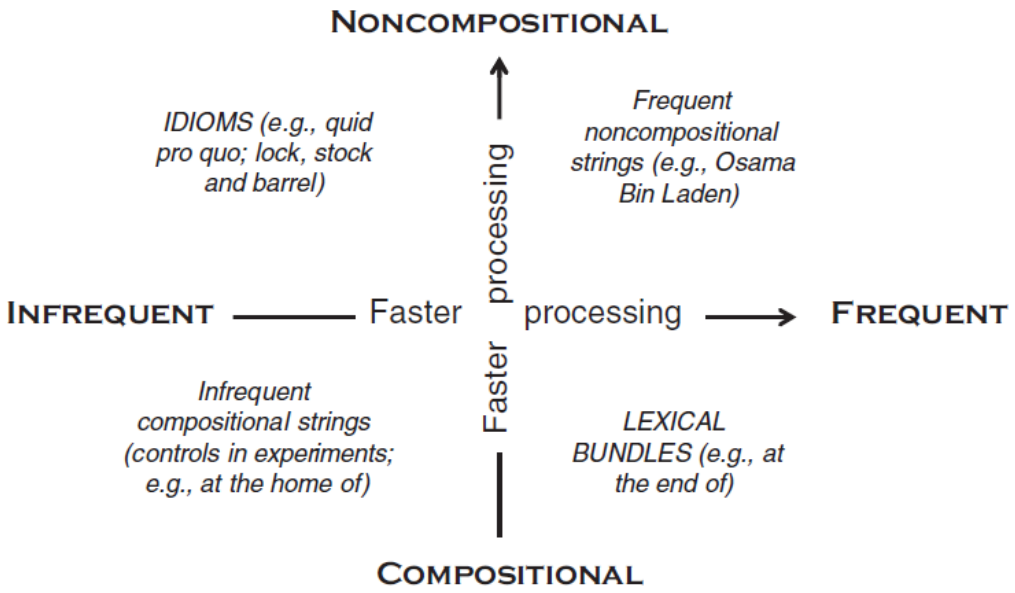

Fig. 5 Wray's fast (or holistic) processing on two axes

(Wray, 2012, p. 241)

According to what Wray (2012) suggests about Figure 5, because of the advantage of frequency and non-decomposability features with respect to formulaicity, frequent non-decomposable word structures such as the commonly used proper names are expected to be processed faster than lexical bundles and idioms which are more decomposable and less frequent, respectively. However, Wray's assumptions remain at the descriptive level, in need of investigation since Wray (2012) also identifies herself as "pattern spotter" (p. 249), indicating the necessity to provide further evidence through studies.

As for the decomposability nature of the analyses of idiomatic expressions, Sag et al. (2002) analyze idiomatic expressions as a category under multi-word expressions and divide them into non-decomposable and decomposable idioms, which are accepted to belong to semi-fixed expressions and syntactically-flexible expressions in Sag et al. (2002) respectively. According to Sag et al. (2002), idiomatic expressions should be analyzed in terms of semantic decomposability,

which is determined whether the overall meaning of idiomatic expression can be accurately interpreted by understanding the meaning of its parts separately, forming the idiom.

Likewise, Fischer and Keil (1996) classify idioms into decomposable and non-decomposable[3], analyzing them in terms of their parsing nature. They have perceived idioms belonging to two separate categories, which is also true in the aforementioned study. Fischer and Keil (1996) indicate that decomposable idioms can be decomposed into components while non-decomposable ones are syntactically frozen, functioning as chunks.

Moreover, Siyanova-Chanturia, Conklin, and Schmitt (2011) have investigated idiom processing in native and non-native speakers, comparing the processing nature in the analyses of decomposability and non-decomposability within the decomposability nature of idioms, which is an influential factor in the study. Siyanova-Chanturia et al. (2011) have pointed out that decomposability nature affects the processing nature of non-decomposable idioms among the non-native speakers negatively, yet it remains neutral for that of non-decomposable idioms among the native speakers. However, they again analyze the idioms with respect to the two dichotomous ends of decomposability.

The decomposability nature of idioms was efficiently used in the experiment conducted by Titone and Connine (1994). They investigated the decomposability nature of idiomatic expressions through experiments in L1 use. In terms of decomposability feature, Titone and Connine (1994) perceived idioms in three ranking scales, namely normally decomposable, abnormally decomposable and non-decomposable. Normal decomposability refers to idioms whose meanings are

[3] The study by Fischer and Keil (1996) uses the term 'non-compositional', but since 'non-compositional' and 'non-decomposable' are synonyms, the current study will use 'non-decomposable' for coherence.

directly related to their figurative interpretations, whereas abnormal decomposability refers to idioms whose words have a metaphorical relation to the idiom meaning in a less literal and more figurative manner.

Among the investigations of PV structures, the studies of Gonnerman and Hayes (2005), Konopka and Bock (2009) and Cappelle, Shtyrov, and Pulvermüller (2010) draw attention to the decomposability analysis of PVs. Gonnerman and Hayes (2005) analyze decomposability of PVs in L1 use, addressing the term adjacency. The constituents of PVs, namely verb and particle, are produced adjacently or separately (Gonnerman & Hayes, 2005). On the other hand, Konopka and Bock (2009) investigate L1 use of PVs in two conditions, post-object and post-verb priming, which refer to decomposability and non-decomposability respectively. They divide the items as frozen and flexible. Under these categories, two subcategories are formulated as idiomatic and non-idiomatic. They look into the decomposability nature though priming experiments.

From a different perspective, Cappelle et al. (2010) examine PV decomposability by using magnetoencephalography (MEG) evidence, thereby providing a neurophysiologic explanation. The PVs "rise up, fall down, heat up, cool down" have been used in the neurophysiological experiment of language processing among native speakers of English. They have been analyzed as "congruent and non-congruent verb-particle combinations" (Cappelle et al., 2010) in terms of L1 use, which again denotes decomposable and non-decomposable PVs in the end.

Finally, the study conducted by Molinaro et al. (2013) is enlightening within the scope of analysis of collocations from the decomposability perspective. Molinaro et al. (2013) have analyzed collocational complex prepositions, basically consisting of Italian collocations, in two conditions, namely Standard condition which actually

symbolizes non-decomposability of the structures and Insertion condition, which actually symbolizes decomposability nature among the native speakers of Italian. Their study involves the analysis of collocations in L1 use through decomposability.

No matter how variant all the aforementioned studies are in terms of their focus of analysis such as formulaicity, idiomatic expressions, phrasal verbs and collocations, their methodology structure of them shows strong resemblance in the division style of the target items in their studies. That is to say, these studies adopt the decomposability nature of the items more or less in the same manner, with a different terminology. Most of the studies divide the decomposability issue into two dichotomous categories as decomposable and non-decomposable in their analyses of L1 and L2 use of the target items.

3.2 Hypotheses on resolving the decomposability issue

There are several hypotheses put forward by scholars in the field trying to provide clarification for the decomposability issue. These hypotheses present different perspectives to deal with the decomposability paradigm of idiomatic expressions, phrasal verbs, metaphors and collocations as multi-word expressions by and large.

The lexical representation hypothesis is one of the effective decomposability models being highly used within the scope of the decomposability nature of multi-word expressions. The hypothesis is put forward by Swinney and Cutler (1979). They perceive idiomatic expressions as a large word-like unit and its individual constituents do not possess any characteristics on their own as individual lexical items. These expressions lose their individual meanings of each of their constituents about syntactic, semantic features and grammatical features of the structures.

Therefore, what lexical representation hypothesis supports is that individuals should process figurative idiomatic expressions faster and make decisions about them without fixation rather than literal idiomatic expressions. Under this hypothesis, Swinney and Cutler (1979) point out the following idea between the items having literal and figurative interpretation:

> Under [lexical representation hypothesis] . . . the computation of a literal meaning and the access of a lexical (idiomatic) meaning should be undertaken simultaneously for the idiom string. The access of the lexical interpretation should conclude far more quickly than the access and computation of the relationships among the several lexical items in the literal interpretation of the idiom. Similarly, the lexicalized idiom meaning should also be recovered before the access and computation of relationships between words in the literal control phrase. (p. 526)

In the light of what Swinney and Cutler (1979) have proposed, it can be acknowledged that individuals store idiomatic expressions in their mental lexicon, asserting the simultaneous initiation of the retrieval of lexicalized and literal idiomatic expressions, but the early activation of lexicalized expressions and the late retrieval of literal idiomatic expressions. Swinney and Cutler (1979) have regarded idiomatic expressions as literal and lexicalized/figurative which signify decomposable and non-decomposable forms of understanding within the decomposability context. The lexical representation hypothesis by Swinney and Cutler (1979) is also readdressed in the study conducted by Cutting and Bock (1997) as strong unitization hypothesis.

The idiom decomposition hypothesis is another major hypothesis, which has gained much popularity recently. The hypothesis suggests that in terms of lexical and semantic processing, idiomatic expressions are not single lexical units, but these expressions are analyzable (Gibbs, Nayak, & Cutting, 1989). It analyzes the decomposability phenomenon as a basis, pointing out the idea that the decomposition of idiomatic expressions displays variance to the degree that they are decomposable

or non-decomposable. In fact, Gibbs et al. (1989) explain the processing of idiomatic expressions as being dependent on whether they are decomposable, which stipulates that the representation and analysis of the individual constituents of the idiomatic expressions are associated with the overall meaning of these expressions, or they are non-decomposable, which suggests that the overall meaning of idiomatic expressions are not related to the meaning of their individual parts.

Gibbs et al. (1989) further point out in terms of the idiom decomposition hypothesis that the decomposable idiomatic expressions are processed faster than their novel counterparts because people can process the meaning of the word parts of the decomposable idiomatic expressions one by one and assign separate meanings to each individual word constituent of them, for the formulation of the meaning of the target item in their mental lexicon. On the other hand, there is no processing advantage for the comprehension of the non-decomposable idiomatic expressions within the decomposability nature and these non-decomposable structures are processed and retrieved as a whole, thereby necessitating the acquisition of their figurative meaning as chunks. Gibbs et al. (1989) also indicate the semantic analyzability of idiomatic expressions to be bound to their syntactic flexibility.

Standard pragmatic model is one of the most influential models in the literature. It was proposed by Bobrow and Bell (1973). It poses the idea that idiomatic expressions are stored as holistic units in the mental lexicon, which forms the separate part of semantic memory of individuals. Thus, when individuals are exposed to an expression, the first interpretation these individuals directly retrieve from their semantic memory will be the one most closely related to concrete and literal meaning of the expression (Bobrow & Bell, 1973). Thus, the first literal interpretation is checked and considered for a brief period to make sense of the

context, but if the literal meaning is not integrated in a suitable manner for the context, then the idiomatic mental lexicon comes in handy for the retrieval of figurative idiomatic interpretation which is processed as the non-decomposable structure. Within the scope of this model, Bobrow and Bell (1973) have supported the following as the main approach in their study and have stated the retrieval process of the idiomatic expression:

> [Idiomatic expressions are investigated] from the point of view of information processing rather than linguistic analysis and is concerned with comprehension rather than production. . . . the idiom "John let the cat out of the bag" [as an example] cannot be understood to mean "John told the secret" by simply changing the meaning of one word, by regrouping adjacent words into a new phrase structure, or by reinterpreting the relationships between some of the words. Instead, the idiomatic meaning seems to be understood by combining several words into a complex "idiom word" and finding the meaning of the phrase by a search through a mental "idiom word" dictionary. (p. 343)

What can be set forth related to the idea proposed by Bobrow and Bell (1973) is that if the literal meaning always precedes the idiomatic figurative one, then the non-decomposable structures may turn into the structures which are accessed with the longer fixation period, which may lead to a situation available that is vice versa (cf. Siyanova-Chanturia et al., 2011).

Configuration hypothesis, being regarded as another major hypothesis, was postulated by Cacciari and Tabossi (1988). As opposed to standard pragmatic model in suggesting that idioms are separately located in the "mental idiom word dictionary" (Bobrow & Bell, 1973), what Cacciari and Tabossi (1988) propose in configuration hypothesis is that "idioms are not encoded as separate entries in the mental lexicon. Rather, their meaning is associated with particular configurations of words and becomes available . . . whenever sufficient input has rendered the configuration recognizable" (p. 678). Regarding this perspective, it can be put forward that multi-word expressions are processed one by one as activating

individual word meanings until enough information has been attained to figure out the word sequence as an idiomatic expression because Cacciari and Tabossi (1988) have stressed uninitiated recognition of a configuration such as idiomatic expressions "before a certain amount of information has been received" (p. 678). In order to indicate to what extent the certain amount of information will be sufficient for an idiomatic expression to activate the idiomatic meaning, Cacciari and Tabossi (1988) have formulated the term "idiomatic key" (p. 678-679), which determines the appearance of the figurative context available.

Finally, possessing conflicting characteristics with configuration hypothesis in terms of the decomposability nature of multi-word expressions, the production hypothesis is conceptualized by Cutting and Bock (1997), formulating this hypothesis taking either standard pragmatic model or lexical representation hypothesis as the baseline for the derivation of the concept. According to Cutting and Bock (1997), production hypothesis may signify that "idioms are represented and accessed as whole units" (p. 58). It can be acceded that what Cutting and Bock (1997) propose in this conceptualization is that a clear-cut distinction in terms of the separate constituents of an idiomatic expression which displays a literal interpretation and a figurative interpretation of the idiomatic construction might suggest that an individual should be able to retrieve the figurative meaning of an idiomatic construction during the course of speaking, notwithstanding the literal meanings of its separate constituent words. For this reason, "speakers have in mind the idea that is to be conveyed prior to producing an idiom, and that idea presumably *maps onto the figurative meaning*" (Cutting and Bock, 1997) [emphasis added].

CHAPTER 4

PHRASAL VERBS

"Formulaic language is like the elephant differently described by blind men with access to different parts of its huge mass." (Wray, 2012, p. 239)

Formulaic language has been investigated and interpreted differently by several scholars in the field as Wray (2012) has stated in the quotation. It has been analyzed in terms of various aspects throughout academic history. Formulaic language involves multi-word expressions, idiomatic expressions, phrasal verbs, metaphors and collocations in broad terms. In fact, it is like a huge elephant, consisting of many subcomponents. The present chapter will deal with the first three main components, namely, multi-word expressions, idiomatic expressions and phrasal verbs in line with the study. However, to narrow down the topic even more, the current study will specifically dwell on ‘phrasal verbs’ at full length as its main objective.

4.1 Multi-word expressions

Multi-word expression (MWE) is a lexical structure which is composed of more than one word, actually a unit of words, referring to a meaningful structure, displaying variance from fixed expressions to flexible ones within the scope of formulaic language. MWE is used as an umbrella term to address various types of linguistic structures, ranging from proverbs to collocations and even compounds.

Sag et al. (2002) broadly define MWEs as "idiosyncratic interpretations that cross word boundaries (or spaces)" (p. 2). Because of their idiosyncratic nature, MWEs pose a challenge for many researchers at the different layers of analysis in

terms of Natural Language Processing (NLP), leading to difficulties in their decomposability nature as "a pain in the neck for NLP" (Sag et al., 2002). Moreover, Grant and Bauer (2004) perceive MWE as "a fixed and recurrent pattern of lexical material sanctioned by usage" (p. 38), focusing on the non-decomposable nature of MWEs. In line with these perspectives, the study can conceptualize MWEs as holistic semantic units being constrained by their syntactic features to some extent.

MWEs incorporate a wide range of expressions, fixed or flexible, analyzed in terms of semantic and syntactic levels, including proverbs, idioms, phrasal verbs, collocations and compounds as the main components. The fixed MWEs can be basically exemplified as proverbs (e.g. When the going gets tough, the tough get going, to mean 'be determined and do not give up'), idioms (e.g. kick the bucket, to mean 'die'), phrasal verbs (e.g. call on, to mean 'visit'), collocations (e.g. take for granted, to mean 'believe and accept'), compounds (e.g. hotdog, to mean 'sausage'), all of which are perceived to behave in a nondecomposable and opaque manner in terms of their syntactic and semantic nature. However, in other cases, the flexible structures such as proverbs (e.g. Hope for the best, but prepare for the worst, to mean 'Bad things might happen, so be prepared'), idioms (e.g. spill the beans, to mean 'reveal secret'), phrasal verbs (e.g. eat up, to mean 'finish eating'), collocations (e.g. make a decision, to mean 'decide'), compounds (e.g. freezing cold, to mean 'extreme cold') are accepted to behave in a decomposable and transparent manner. What is essential is that these expressions are widely used within some structural patterns to denote certain lexical meaning. The examples show a few essential structures under MWEs, but there are more to count. However, the study will continue to explain idiomatic expressions, as one of the most-debated issues, and more importantly, it will particularly focus on phrasal verbs, the most widely-used expressions.

4.2 Different views on idiomatic expressions

Idiomatic expressions have aroused a great deal of interest, being regarded as a highly debated issue for their definitions and classifications. Several researchers have provided different dimensions for the classification of idiomatic expressions (Bobrow & Bell, 1973; Caillies & Declercq, 2011; Ciéslicka, 2012; Cutting & Bock, 1997; Gibbs & Gonzales, 1985; Gibbs, Nayak, & Cutting, 1989; Grant and Bauer, 2004; Fraser, 1970; Hillert & Swinney, 2001; Kovecses & Szabo, 1996; Nunberg, Sag, & Wasow, 1994; Swinney & Cutler, 1979; Tabossi, Fanari, & Wolf, 2009; Titone & Connine, 1994; Vega-Moreno, 2003). These dimensions can be listed as decomposability, conventionality, frozenness, transparency, analyzability, transparency, familiarity, predictability and literality, used by different scholars. There are also a few more dimensions which will be taken into consideration.

In all the aforementioned dimensions, two basic aspects of idiomatic expressions draw focal attention. The first one indicates the relation between a certain string of words and a certain semantic representation, displaying arbitrariness to express certain conceptual representations, referring to familiarity, transparency, frozenness and conventionality as a holistic approach. The second one is about the accessibility of the figurative meanings of the idioms or the intended implication of idiomatic expressions, referring to decomposability, predictability, analyzability and literality as an analytical approach.

Nunberg et al. (1994) perceive idiomatic expressions as analyzable decomposable structures by dividing them into orthogonal properties and semantic properties. Nunberg et al. (1994) have listed conventionality, inflexibility, figuration, proverbiality, informality, affect within ‘the orthogonal properties’ and compositionality, conventionality, transparency within ‘the semantic properties’.

According to Nunberg et al. (1994), none of these properties is obligatorily pertinent to all idioms except for the conventionality property because most of the idiomatic expressions can have identifiable parts in terms of meaning, related to the constituents of the idioms, and they can be analyzed as phrases.

Likewise, Titone and Connine (1994) view the idiomatic expressions within the categories of familiarity, compositionality, predictability and literality for the idiom processing, and these dimensions have been claimed to affect the comprehension of idiomatic expressions. Vega-Moreno (2003) identifies idiomatic expressions with the distinctions 'conventionality, analyzability, transparency' with a focus on the comprehension of idioms in relation to relevance theory.

Some other researchers consider idiomatic expressions within the scope of frozenness. According to Kovecses and Szabo (1996), "idioms are linguistic expressions whose overall meaning cannot be predicted from the meanings of the constituent parts". Similarly, Gibbs and Gonzales (1985) point out that an idiomatic expression can be analyzed according to their degree of syntactic frozenness. The level of syntactic frozenness has an influential factor in processing and remembering idiomatic expressions. Swinney and Cutler (1979) emphasize that "in its simplest form, an idiom is a string of two or more words for which meaning is not derived from the meanings of individual words comprising that string". Swinney and Cutler (1979) have analyzed idiomatic expressions in terms of linguistic frozenness. Moreover, Langlotz (2006) has identified the idiomatic expression as the preliminary definition in the following:

> An idiom is an institutionalised construction that is composed of two or more lexical items and has the composite structure of a phrase or semi-clause, which may feature constructional idiosyncrasy. An idiom primarily has an ideational discourse-function and features figuration, i.e. its semantic structure is derivationally non-compositional. Moreover, it is considerably fixed and collocationally restricted. (p. 5)

From a non-decomposable view, Langlotz (2006) classifies idiomatic expressions as non-decomposable structures like "a hard nut to crack" (p. 15) to point out the unpredictability of some idiomatic expressions by thinking over the fixedness and restrictedness of these items and situating them within a bottom-up perspective; yet later, Langlotz (2006) interestingly delves into the topic "idiomatic creativity and variation" in his book *Idiomatic Creativity*.

Grant and Bauer (2004) redefine idioms as a particular type of multi-word expression, and classified them semantically, syntactically, and functionally. What Grant and Bauer (2004) point out is that idiomatic expressions are frozen metaphors and there are problematic issues with the decomposable theories trying to analyze the constituents of them separately, so ‘in order not to bark up the wrong tree’, idiomatic expressions should be perceived as 'core idioms' and used in a non-decomposable manner with their figurative meaning.

Moreover, idiomatic expressions have been recognized as lexical bundles according to some researchers (Biber & Conrad, 1999; Hyland, 2012; Tremblay, Derwing, Libben, & Westbury, 2011). These researchers have observed idiomatic expressions which are stored and processed as single units and non-decomposable structures, adopting a holistic approach. Biber and Conrad (1999), for instance, analyze lexical bundles in conversational and academic prose. This is also the focus of the study conducted by Hyland (2012), emphasizing the importance of frequent use of lexical bundles in academic discourse and perceiving them as fixed phrases, the learning of which contributes to communicative competence.

Regarding the decomposability of idiomatic expressions, the studies conducted by Cutting and Bock (1997), Riehemann (1997), Tabossi et al. (2009) and Caillies and Declercq (2011) are worth mentioning. Cutting and Bock (1997)

conceptualize idiomatic expressions as unitized phrases with interpretations that are independent of the literal meanings of their individual words. Even so, Cutting and Bock (1997) further their analysis on idiom blends, concluding that "idioms are not produced as *frozen phrases*, devoid of information about their internal syntax and semantics" (p. 69) [emphasis added].

Riehemann (1997) focuses on the phrasal types of idiomatic expressions, specifying the semantic relationship between the constituents of idiomatic expressions involved. Idiomatic expressions have been analyzed in the head-driven phrase structure grammar framework. According to Riehemann (1997), “idioms have two main properties that are hard to account for in various approaches: they tend to be syntactically variable, and they sometimes involve fixed items that go beyond simple head-complement relationships” (p. 2); and thus, within the decomposability perspective, Riehemann (1997) view idiomatic expressions as decomposable (the former) and non-decomposable (the latter), two different ends. Riehemann (1997) has further divided the alternative dimensions into two, namely, whether they represent idioms at the word or phrasal level, and whether the kind of information that gets specified is syntactic or semantic, analyzing idiomatic expressions through these dimensions, especially making word-level versus phrase-level comparisons.

Similar to the study by Riehemann (1997), Tabossi et al. (2009) have conducted a study on the recognition levels of idiomatic expressions and classified these expressions into two groups, namely, decomposable and non-decomposable within the decomposability perspective, subcategorizing them into two subgroups, literal and nonliteral. According to Tabossi et al. (2009), "findings indicate that idiomatic expressions are not treated holistically; rather, they undergo full analysis, like literal expressions. This is the case even though they are semantically

noncompositional" (p. 538). Thus, even though Tabossi et al. (2009) classify idiomatic expressions as non-decomposable and decomposable, they perceive these expressions as fully decomposable convenient for semantic and syntactic analyses.

Moreover, Caillies and Declercq (2011) have recently analyzed the semantic processing of decomposable idioms and predicative metaphors to indicate to what extent there is a processing difference between these structures. Idiomatic expressions have been perceived as decomposable idioms which go through semantic operations, and the verb and the object of the verb have semantic relations. Caillies and Declercq (2011) have classified idiomatic expressions according to familiarity, predictability and literality. Nevertheless, the idiomatic expressions, termed as decomposable, are expected and found to be processed and retrieved as holistic units from the semantic lexicon. Thus, decomposable idiomatic expressions in the study of Caillies and Declercq (2011) are not fully decomposable.

Finally, idiomatic expressions have been addressed in terms of the idiomatic creativity and variation (Cutting & Bock, 1997; Langlotz, 2006; Pitzl, 2011; Prodromou, 2007). Cutting and Bock (1997) term idiomatic expressions as idiom blends to signalize the creativity of the formation of idioms in the production phase. Idiom blend errors can occur at the word or phrase level (Cutting & Bock, 1997), but even so, idioms are decomposable structures and susceptible to variation despite being recognized as errors. Prodromou (2007) investigates creative idiomaticity as the creative use of idiomatic expressions among the L1 and L2 users of English from an ELF perspective, perceiving idioms as ‘word play’. What Prodromou (2007) indicated is that though speakers highly use idiomatic expressions as verbal plays to be witty in the stream of conversations, the word plays of L1 speakers are perceived as creativity, but those of L2 users are interestingly perceived as deviations.

4.3 Idiomatic expressions in ELF encounters

The immense increase of English language within the scope of the number of individuals speaking it as their L1, L2, L3 and maybe even L4, taking the globalized status of English into consideration, has prompted native-like competence in the language to take precedence among these individuals aiming to have credentials to use formulaic language, thereby contributing to an increasing attention to idiomatic expressions. Therefore, only a few researchers have recently started to analyze idiomatic expressions in the ELF context, trying to shed some light on the use of idiomatic expressions by ENL and ELF speakers (Seidlhofer, 2009a; Pitzl, 2009, 2011, 2012; Prodromou, 2008).

Seidlhofer (2009a) analyzes the place of idiomatic expressions in ELF framework among ELF users who have a tendency to focus on the open-choice principle, co-constructing idiomatic expressions in the stream of conversation, whereas the constructional patterns of idiomatic expressions are dependent on the idiom-principle that is consulted by native speakers of the language to maintain effective communication. The idiom principle forms the single choices of individuals in the processing and retrieval of idiomatic expressions from "a large number of semi-preconstructed phrases" (Sinclair, 1991, p. 110). This calls for the conventional use of pre-constructed idiomatic expressions with respect to the cooperative imperative, which is termed as the small adjustments and continuous modifications of language structures by the interlocutors to effectively manage conversations. It reinforces accessibility (Seidlhofer, 2009a). However, Seidlhofer (2009a) also indicates the existence of ‘territorial imperative’ in which native speakers adjust their language accordingly for their protection of personal space and reinforcement of their social identity.

Being influenced by the two distinctions, namely cooperative and territorial imperative, the idiom principle was perceived as having an influential role for ELF users to accommodate themselves to their interlocutors while engaging in their own 'idiomatic realizations' in line with communicative convergence (Seidlhofer, 2009a). ELF users focused on co-constructing idiomatic expressions to attain convergence in their interactions by trying to make their language appropriate within the idiom principle though their construction of idiomatic expressions does not entirely conform to L1 users' idiomatic expressions, which does not have to be exactly similar, either. Seidlhofer (2009a) emphasized that "In using English *on* their own terms, ELF users will quite naturally use English *in* their own terms" (p. 211) [emphasis added].

As for the analysis of idiomaticity though corpus studies in ELF, Prodromou (2008) and Pitzl (2009, 2011, 2012) have recently investigated the idiomatic expressions via corpus analysis. Pitzl (2009) has illustrated the existence of idiomaticity in ELF, analyzing the metaphorical functions. Pitzl (2009) has emphasized the conceptualization of re-metaphorization which is available in the use of idiomatic expressions of ELF speakers. As an extension of Pitzl (2009), Pitzl (2011) looks into idiomatic creativity in ELF through a corpus study of VOICE in her dissertation. Pitzl (2011) has analyzed the use of idioms and metaphorical expressions in VOICE. ELF speakers are found to use idioms and metaphorical creativities in terms of linguistic creativities to the extent of conventional language use by ENL speakers according to Pitzl (2011). Pitzl (2012) has detailed her findings and analyzed dimensions of metaphorical creativity, addressing the internal features of idiomatic expressions. However, since Pitzl (2012) focused only on metaphors, there is the necessity of a cross-linguistic corpus study on idiomatic expressions.

Prodromou (2008) has analyzed the idiomatic expressions, namely ‘you see’ and ‘sort of’ through native and non-native corpus. As the native corpus, Prodromou (2008) has analyzed these two expressions in small parts of BNC and CANCODE (The Cambridge and Nottingham Corpus of Discourse in English). As the non-native corpus, Prodromou (2008) has analyzed his own corpus that he collected from successful users of English with different L1s. Prodromou (2008) has demonstrated the frequency of ‘you see’ and ‘sort of’ in the L1 and L2 use. The importance of idiomaticity has been stressed in ELF context. Prodromou (2008) has found that L2 speakers used ‘you see’ and ‘sort of’ fewer than L1 speakers. The moderator role of speakers has been also emphasized within the interactional atmosphere according to Prodromou (2008). "The linguistic choice is not determined by the context, but in part creates the context" (Prodromou, 2008, p. 145). In line with this perspective, Prodromou (2008) concludes that "even proficient L2-users" (p. 238) have avoided certain idiomatic creativities, which is "the result of not lack of ‘competence’, but an intuition . . . as to where the limits are" (Prodromou, 2008, p. 239), questioning the performance of successful users of L2 from an ELF perspective.

All in all, considering all the aforementioned studies, the studies conducted by Pitzl (2011) and Prodromou (2008) specifically signalize a crucial gap in the literature for me, which has contributed to this thesis. A question arose in my mind: "What about phrasal verbs?" This question has urged me to design the current study, delving into the phrasal verb constructions, analyzing phrasal verbs through a corpus-based study cross-linguistically both in a standardized ENL corpus, BNC, and in a full-fledged ELF corpus, VOICE, and formulating an alternative approach based on the analysis.

4.4 Definition of phrasal verb and verb-preposition construction

Phrasal verb (PV) can be defined in broad terms as one of the multi-word expressions, commonly viewed as a structure, consisting of a verb and an adverbial particle in syntactic terms and functioning as a single verb, which in the end loses the meanings of both of its constituents in semantic terms, thereby yielding the construction of a new lexical unit to an extent (Darwin and Gray, 1999). PVs are the combinations of a lexical verb and a particle, denoting a new meaning together, being highly used for successful communication. The definition of PVs according to Gardner and Davies (2007) is that two-part verbs in the BNC being composed of a lexical verb proper and followed by an adverbial particle which is either contiguous or noncontiguous as for the lexical verb, to mean that the adverbial particle can directly follow the lexical verb or it can be used after more words coming in the intervening position between the verb and particle and situated after those words. ‘Turn off the lights’ and ‘turn the lights off’ can be given as examples for the former and the latter conditions, respectively.

The current study regards PVs as tiny structures having a small impact on one's language, but they are the actual representations of language competence as an invaluable part of communicative competence strategies, enhancing the comprehension and successful command of the language.

Verb-preposition constructions (VPCs), on the other hand, are composed of a verb and a preposition, the function of which is to provide the object of the verb with a morphological dimension in the sentence. There are some linguistic constraints in English, which stipulates the existence of a special preposition following a verb, morphologically complementing the verb, but not semantically, in the formulation of syntactic structure of the sentence. For instance, the verbs ‘look’ and ‘listen’ can be

regarded as the examples of VPCs, being followed by ‘at’ and ‘to’ respectively, thereby forming "Subject + look (verb) + at (prep) + somebody / something (object)" and "Subject + listen (verb) + to (prep) + somebody / something (object)".

The main distinction between PVs and VPCs as the cardinal rule is that PVs cannot exist without their adverbial particles and the lexical verb constituents of PVs cannot stand on their own without the particle structures in order to designate the same intended meaning whereas VPCs can perfectly yield a normal usage of the structure without their prepositions having the same intended meaning on their own on the condition that they do not take any objects. That is to say, an adverbial particle has to be present in the sentence either adjacent to the lexical verb or not, which is determined by the syntactic constraints of PV; on the contrary, the preposition of VPC does not have to be present in the sentence, so if there is no need for the object, the preposition does not have to be used and should not be used, as a matter of fact. The preposition item actually functions as a morphological connection between the verb and object, and in the cases where there is no object of the verb, the preposition should be absent in the sentence.

More interestingly, within the scope of the decomposability perspective, PVs can be used in a decomposable and non-decomposable manner; being constrained by some syntactic terms (see section 4.5). The adverbial particle of PVs can function either in an adjacent position or in a separate one, having the objects of the lexical verb being used in the intermediary position between lexical verb and particle. Even so, the lexical verb and the particle are closely associated with each other in either case. However, when considered from the aspect of VPCs, the preposition cannot be used separately and it always has to be after the verb in an adjacent position, being used as non-decomposable structure all the time within the decomposability nature.

Take the PV ‘look up’ and the VPC ‘look at’ as an example for the distinction, analyzing the issue considering two different structures with the same lexical verb. ‘Look up’ has a special meaning, being denoted by the combination of lexical verb and particle as PV, being used to express the meaning of specifically searching something from a directory, dictionary or huge collections of records. An individual can use it as, ‘OK, I will look up the word’ or ‘OK, I will look the word up’ within the context. However, ‘look at’ as VPC cannot be used in this way. An individual can state ‘I will look at the newspaper.’, but s/he cannot use the structure as ‘*I will look the newspaper at’[4]. In addition to the decomposability view, as a semantic entity, ‘look up’ can be used on its own. ‘I will look up’ is a grammatical usage whereas ‘*I will look at’ is ungrammatical and ‘at’ in the second sentence is redundant and semantically meaningless without its object since VPCs do not allow the preposition to stand on its own as a semantic entity in the sentence. The preposition has only a morphological function.

The current study delves into PV structures while adopting the perspectives stated in the aforementioned definitions of both of the studies by Darwin and Gray (1999) and Gardner and Davies (2007) by contributing to the literature through its analyses of these structures with an alternative approach. As a matter of fact, Darwin and Gray (1999) analyzed PVs, realizing PVs and VPCs belonging to the same category, at which point the current study diverges from. Interestingly, Gardner and Davies (2007) did not make any distinction between these two structures, either. Thus, by adopting the view that PVs are combinations of a lexical verb and an adverbial particle used directly after the verb or separately from the verb, the PVs will be perceived as a lexical entity in terms of their semantic analysis, having been

[4] The symbol asterisk (*) is used to indicate ungrammatical structure in the sentence. All the asterisks in the current study will underline the same meaning.

investigated within the transparency cline and as a syntactic entity, being exposed to the uses of decomposable, non-decomposable and both positions in terms of their syntactic nature within the decomposability cline, causing a clear-cut distinction of PVs from VPCs.

4.5 Phrasal verb taxonomies in the literature

Phrasal verbs have recently attracted a great deal of attention of researchers analyzing the phenomenon in certain dimensions, under different taxonomies in the literature. Although PVs have gathered increasing attention for the last three decades, the concept of PV dates back to the 1920s. Smith (1925) referred to PVs as a term in his book, as stated in the following:

> The term "phrasal verbs" was suggested to me by the late Dr. Bradley; not, as he wrote, that he was satisfied with it, or would not welcome any alternative that he could feel to be an improvement. But, as he said, one cannot write of these verbs without some workable description; and although the word "phrasal" is perhaps objectionable in information, it fills a want, and is sometimes indispensable. (p. 172)

The prevalence of PVs, however, has become highly observable since the 1990s, being affected by the spread of English at a global level thanks to rapid advancements in technology, economic, political and educational globalization.

PVs have been analyzed by being divided up into several taxonomies in the literature. These taxonomies include decomposability perspective, transparency perspective, avoidance behavior, teaching and learning perspectives, nativeness criterion, material design effects, psychotypology perspective, and corpus-based perspective which have channeled the researchers' analyses. The taxonomies which the current study will specifically ponder are decomposability, transparency and corpus-based perspectives; yet, the other dimensions will be mentioned briefly.

4.5.1 Decomposability taxonomy

Decomposability taxonomy is a type of classification for PVs which is used for the purpose of making distinctions in the studies of Matlock and Heredia (2002), Darwin and Gray (1999), Konopka and Bock (2009).

In terms of decomposability taxonomy, Matlock and Heredia (2002) have analyzed PVs through 4 techniques. The first one is to make a distinction between phrasal verbs (PVs) and verb-preposition constructions (VPCs) through syntactic constraints, involving particle movement (Matlock & Heredia, 2002).

i. Bob ate up the lasagna. [PV]

ii. Bob ate the lasagna up. [PV]

iii. Bob ate up the hill. [VPC]

iv. * Bob ate the hill up. [5] [VPC] (Matlock & Heredia, 2002)

The second one is about anaphoric referents. According to Matlock and Heredia (2002), the pronoun cannot be used after the adverbial particle in PVs:

v. Bob ate it up (it = lasagna)

vi. * Bob ate up it (it = lasagna) (Matlock & Heredia, 2002)

The third one is about the insertion of adverbs between lexical verb and adverbial particle. Although adverbs can be inserted between verb and preposition in VPCs, PVs do not welcome adverbs in the intermediary position between verb and its adverbial particle. The examples are as follows (Matlock & Heredia, 2002).

vii. * Bob ate quickly up the lasagna.

viii. Bob ate quickly up the street at Romano's, and then rushed off to work.

The last one is about topicalization. As for Matlock and Heredia (2002), VPCs allow topicalization which removes the prepositional phrase to the sentence

[5] The asterisk (*) shows that the sentence is ungrammatical.

initial position with some syntactic formulation, whereas PVs do not welcome such variation in the sentence. The following examples (ix) and (x) are the syntactic variants of the sentences (iii) and (i), respectively.

ix. Up the hill is where Bob ate.

x. * Up the lasagna is what Bob ate.

With respect to decomposability taxonomy, what Matlock and Heredia (2002) have indicated is that PVs are "highly conventionalized linguistic forms" (p. 265), displaying idiosyncratic syntactic constraints.

Konopka and Bock (2009) have recently analyzed PVs, conducting a priming study about the particle placement of PVs through structural priming in three experiments, displaying prime-target pairing sentences in post-verb and post-object position causing the adverbial particle to be adjacent to the verb or separate from the verb, as shown in the following examples:

xi. A celebrity threw in the first ball. (Post-verb)

xii. A celebrity threw the first ball in. (Post-object) (Konopka & Bock, 2009)

Interestingly enough, although Darwin and Gray (1999) have resorted to several syntactic constraints to distinguish PVs as for the decomposability taxonomy such as replacement, formation of passives, formation of action nominal (making PVs gerund), object movement, pronoun placement, adverbial insertion, stress, definite noun phrases and listing, they do not exclude VPCs from PV category and accept VPCs as potential PVs until VPCs are proven to be otherwise. Thus, though the theoretical definition of PVs by Darwin and Gray (1999) is in line with the perception of PVs by the current study, the categorization structure of Darwin and Gray (1999) does not correspond to the objectives and categorization of this study.

4.5.2 Transparency taxonomy

Transparency taxonomy is another type of classification for PVs in semantic terms used by the studies (Dagut & Laufer, 1985; Gonnerman & Hayes, 2005; Laufer & Eliasson, 1993).

With respect to transparency taxonomy, Dagut and Laufer (1985) have analyzed PVs under three semantic distinctions which are composed of "literal, figurative and completive" (p. 74). Literal PVs are those whose meanings can be derived from the semantic interpretations of each of their constituents directly (e.g. go out, come in) according to Dagut and Laufer (1985). Figurative PVs are the structures whose constituents have inherent meanings, but are interpreted differently due to "a metaphorical shift of meaning and the semantic fusion of the individual components", causing a new meaning to build up (e.g. let down, turn up, show off). The third and last one is completive PVs "in which the particle describes the result of the action" (e.g. cut off, burn down) as exampled by Dagut and Laufer (1985, p. 74).

Laufer and Eliasson (1993) have distinguished three types of PVs. These are "semantically transparent" in which PV can be interpreted based on the meaning of its parts, "semitransparent" in which PV is transparent when put into context, and "figurative" or "semantically opaque" having lexicalized meaning (p. 37).

Much recently, regarding transparency taxonomy, Gonnerman and Hayes (2005) have analyzed PVs through semantic similarity judgment tasks, measuring the priming effects of different PVs according to semantic similarity degree. They have categorized PVs based on the similarity of particle to the meaning of the lexical verb of PV. According to Gonnerman and Hayes (2005), the PV ‘bring in’ to ‘bring’ is more semantically similar than ‘chew out’ to ‘chew’. They have classified PVs in terms of their transparency, belonging to more similar or less similar levels.

4.5.3 Avoidance behavior

Avoidance behavior is another criterion influencing the analyses of the researchers on PVs (Chen, 2007; Dagut & Laufer, 1985; Ghabanchi & Goudarzi, 2012; Hulstijn & Marchena, 1989; Laufer, 2000; Laufer & Eliasson, 1993; Liao & Fukuya, 2002). Avoidance behavior can be defined as the underuse of linguistic constructions or the tendency not to use the linguistic items. Hulstijn and Marchena (1989) define avoidance behavior as "a tendency to adopt a play-it-safe strategy" (p. 241), viewing the avoidance behavior of L2 learners "as a cognitive strategy, implying a choice" (p. 243). Moreover, Dagut and Laufer (1985) highlight the role of avoidance behavior by L2 speakers as "the reverse side of negative transfer, since learners tend to avoid using in L2 . . . structures that have no parallel in their L1" (p. 73). Therefore, in line with these perspectives, avoidance behavior can be regarded as a language strategy among L2 speakers with the tendency not to use the target linguistic structures where necessary by feeling secure about their language performances.

The avoidance behavior of PVs is influenced by the transparency of PVs, and especially avoidance is more observable when these PVs are opaque in their semantic nature, which is indicated by the studies of Dagut and Laufer (1985) and Liao and Fukuya (2002). However, according to Hulstijn and Marchena (1989), not only the transparency difference of PVs is what matters in the avoidance of PVs, but also there are other factors in effect in PV avoidance.

Much recently, Ghabanchi and Goudarzi (2012) have investigated PV avoidance being affected in terms of PV transparency and have classified PVs as figurative (e.g. give up), denoting opaqueness, and literal (e.g. get up), denoting transparency. Ghabanchi and Goudarzi (2012) have pointed out avoidance behavior is a determining factor in the PV use for non-native speakers of English.

4.5.4 Teaching and learning perspectives

There are a few studies straightforwardly investigating the issues of teaching PVs and learning them, which are worth mentioning (Condon, 2008; Kurtyka, 2001; Nassaji & Tian, 2010; Pozdnyakova & Gunina, 2011; Rudzka-Ostyn, 2003).

Condon (2008) analyzes the learning of PVs and explains how cognitive linguistics affects the learning of PVs, indicating the positive effect of cognitive linguistics on PV learning as a result of English classes in the experimental design.

For the learning of PVs, Rudzka-Ostyn (2003) examines semantics of PVs in her book *Word Power*, in which PVs and compounds are gathered together related to each adverbial particle. Learning techniques of PVs are explained in a cognitive linguistic approach, leading learners to learn PVs by themselves through 'exetests' (Rudzka-Ostyn, 2003). The 'exetests' (EXErcise + TEST) provide a practice material for the target PVs, by also testing learning of PVs (Rudzka-Ostyn, 2003). Rudzka-Ostyn (2003) divides her book based on each particle (e.g. out, in, up, down, etc.) and provides exetests[6] for each particle after explaining their prototypical and metaphorical meanings. Figure 6 shows the prototypical representations of particles.

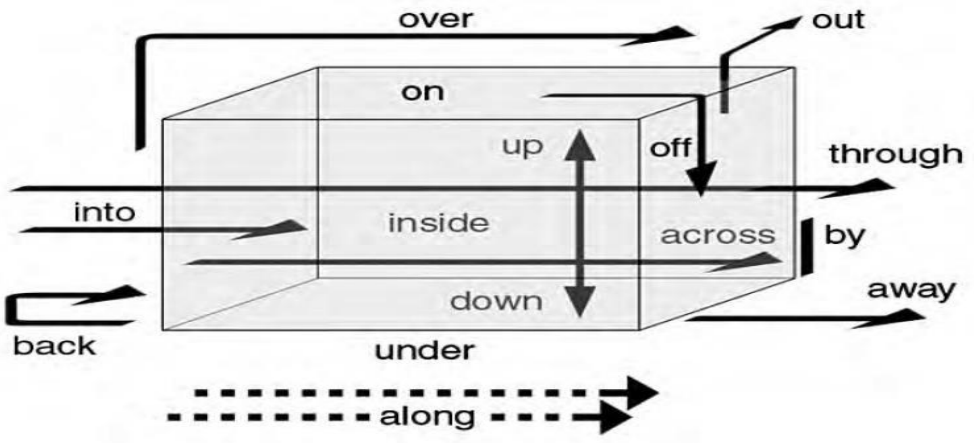

Fig. 6 Prototypical meanings of prepositions and particles

(Rudzka-Ostyn, 2003, p. 4)

[6] For the exemplifications of exetests from the book of Rudzka-Ostyn (2003), see Appendix A.

Rudzka-Ostyn (2003) also explicates the metaphorical meanings of each particle in semantic terms, divided up into several chapters in her book. For example, the particle 'off' is explained having the meaning of 'breaking contact' (Rudzka-Ostyn, 2003, p. 121). The metaphorical meanings of 'off' include loss of spatial contact or spatial separation (e.g. break off, come off) (Rudzka-Ostyn, 2003, p. 122), separation as loss of contact (e.g. beat off, carry off, see off, set off) (Rudzka-Ostyn, 2003, p. 123), separation as interruption of flow/supply (e.g. cut off, switch off, turn off) (Rudzka-Ostyn, 2003, p. 125), separation due to motion away from its former state, condition or point of reference (e.g. cool off, ease off, let off) (Rudzka-Ostyn, 2003, p. 126). Overall, what Rudzka-Ostyn (2003) tries to suggest is to provide metaphorical exemplifications of each particle constituting PVs to enhance learning PVs with a cognitive linguistic approach.

In terms of teaching of PVs, Pozdnyakova and Gunina (2011) examine the techniques and strategies for the teaching of PVs and illustrate a number of PVs being grouped together according to their topics (e.g. food: boil over, chop up; health: come round, pass away, throw up, wear off, travelling: break down, drive off/away, pull in/over, pull out, slow down, speed up; job: make up for, take on, turn down, set up) (p. 357). Pozdnyakova and Gunina (2011) provide different tasks for PVs to be taught by creating a positive environment for language teaching.

4.5.5 Nativeness criterion

In respect to the nativeness taxonomy, a distinction has been made between native and non-native speakers of English in terms of the use of PVs and their one-word equivalents through questionnaires (Siyanova & Schmitt, 2007). Siyanova and

Schmitt (2007) have indicated the existence of a difference between native and non-native speakers of English, drawing attention to the higher percentages of uses of one-word equivalent of PVs when non-natives results are compared to native ones.

4.5.6 Material design effects

The studies investigating PVs within several aspects lead to varying results (Blais & Gonnerman, 2013; Dagut & Laufer, 1985; Gonnerman & Hayes, 2005; Hulstijn & Marchena, 1989; Laufer & Eliasson, 1993; Liao & Fukuya, 2002; Siyanova & Schmitt, 2007). It leaves the question whether the methodological design of the studies affects their general findings or not unanswered as a part of the weaknesses in their studies. As a commonly shared value, the aforementioned studies used controlled material designs, evoking declarative knowledge of the participants by making them aware of the structures and bringing their controlled processing forward, and affected the overall outcomes of the experiments being conducted, such as elicitation tests (a multiple-choice test, a verb translation test, and a verb-memorization test) (Dagut & Laufer, 1985; Hulstijn & Marchena, 1989; Liao & Fukuya, 2002), semantic similarity judgment tasks (Blais & Gonnerman, 2013; Gonnerman & Hayes, 2005), questionnaires (Siyanova & Schmitt, 2007).

4.5.7 Psychotypology perspective

PVs are investigated in terms of psychotypology perspective (Dagut & Laufer, 1985; Hulstijn & Marchena, 1989; Sjoholm, 1995). These studies focus on the language distance situations related to the languages through which PVs are examined.

Dagut and Laufer (1985) have analyzed the avoidance of PVs with a contrastive analysis approach because the participants in the study were L1 Hebrew speakers of L2 English, the languages of whom are psychotypologically distant languages. The Hebrew language does not involve PV structures within the language itself; instead, it possesses one-word equivalents of those PVs although English highly involves a large spectrum of PV uses. Due to the nonexistence of a parallel structure of PVs in the Hebrew language, L1 Hebrew speakers of L2 English have demonstrated much more avoidance behavior in the use of PVs since those speakers do not have any chance to transfer the target structures, being asked, within the two languages since Hebrew and English do not commonly have PV structures. Thus, the majority of the participants in the study avoided using PVs, preferring the one-word equivalent verbs. This can also signalize an existence of psychotypological distance, displayed as a result of the outcome of the study by Dagut and Laufer (1985) in that the study shows the structural differences between L1 Hebrew and L2 English in terms of PVs as non-existent forms in the former and existent forms in the latter. The extent, to which L1 Dutch learners of L2 English avoid using PVs and use one-word equivalents of these PVs instead, outnumbers in terms of percentages in comparison to the extent to which avoidance behavior is indicated by psychotypological distinction L1 Dutch-L2 English speakers by Hulstijn and Marchena (1989).

Likewise, Sjoholm (1995) has also analyzed the cross-linguistic influence of PVs by examining the use of PV structures among L1-Finnish and L1-Swedish speakers of L2 English. He has compared the responses of Swedes and Finns in Finland to the uses of PVs and their one-part verb equivalents in English through the investigations being conducted with multiple-choice tests and questionnaires.

What Sjoholm (1995) has put forward about the psychotypology of the languages of L1 Finnish and L1 Swedish in relation to L2 English is that since there are literal equivalents of the target items in Swedish, but not in Finnish, the general underlying hypothesis by Sjoholm (1995) is to expect Finnish learners of English to make more errors in PV uses than those Swedish; moreover, this hypothesis is evidenced by Sjoholm (1995) through tests indicating that L1-Finnish speakers of English have made many more errors in comparison to L1-Swedish speakers of L2 English who has equal structure in their native language and can benefit from cross-linguistic effects from their L1.

As an overall approach, Sjoholm (1995) stresses the importance of language distance with respect to psychotypology criterion and states the following as a result of his study:

> [T]he data . . . seem to indicate very clearly that language distance plays a role in the acquisition of phrasal verbs. The error data showed that Finns had much greater problems with phrasal verbs (i.e. they made more errors) than Swedes. These differences could imply that phrasal verbs, because of structural differences between L1 and L2, constitute a specific learning problem for Finns. On the other hand, the differences could also partly be a result of "positive transfer" among Swedes. As a matter of fact, it has been found that lexical items in which there is similarity in form and meaning between L1 and L2 are easily transferred. (p. 168)

In the light of what Sjoholm (1995) has posed in the previous quotation, psychotypological closeness can be assumed to be an influential positive factor in the enhancement of correct uses of PVs. Nevertheless, in terms of transparency nature of PVs as an influence on the avoidance behavior of PVs when two languages are compared, interestingly enough, there are no significant results between two learners of English, in Finnish and Swedish because Finnish participants avoid using PVs and instead choose not PVs, but their one-word equivalents due to lack of representation of PVs in their language system, whereas more interestingly, Swedish participants

also avoid using PVs and instead choose not PVs, but their one-word equivalents because though Swedish learners in early stages still assume psychotypological closeness of their language in PV structures to the target language English and still choose PV options, the proficient Swedish learners have become sceptical about the transferability of PVs when they are more proficient learners and they choose equivalent one-part verbs, displaying "the U-shaped curve" (p. 185) according to Sjoholm (1995). Avoidance of PVs does not show any difference between L1 Finnish and L1 Swedish learners of L2 English when compared to the frequencies of the uses of their one-word equivalents.

Hulstijn and Marchena (1989) have investigated the avoidance of PVs among L1 Dutch learners of L2 English, to whom multiple choice, memorization and translation tests of PVs were administered, taking the study by Dagut and Laufer (1985) as their initiative for their own study. Hulstijn and Marchena (1989) indicate that not only structural differences between L1 and L2 being investigated but also structural similarities between L1 and L2 cause avoidance behavior in the use of PVs by Dutch learners of L2 English. Thus, as regards psychotypology of the languages, not only linguistic distance matters, but so does the closeness of the languages.

4.5.8 Corpus-based perspective

With regard to corpus-based perspective, there has been a recent interest among a few researchers aiming to illustrate the frequency of PVs in L1 English through corpus studies (Gardner & Davies, 2007; Liu, 2011).

In terms of the corpus-based perspective of PVs, Gardner and Davies (2007) have analyzed BNC corpus to point out the most frequent PVs in English. They have

identified 518,923 PV occurrences and determined the most frequent 100 PVs in the L1 English corpus. Gardner and Davies (2007) have provided a working definition of PV structures, describing PVs as "two-part words being composed of lexical verb and adverbial particle which is either adjacent to the verb or not". However, Gardner and Davies (2007) have also included VPCs in their corpus analysis, recognizing VPCs as part of PV structures, analyzing both structures in the same category in both written component of BNC (comprised of nearly 90 million words) and spoken component of BNC (comprised of nearly 10 million words).

Likewise, Liu (2011) has also analyzed the most frequent PVs that are used in two native corpora of L1 English. These corpora are BNC and Corpus of Contemporary American English (COCA). Similar to what Gardner and Davies (2007) have analyzed and illustrated in their corpus study on PVs, Liu (2011) has also identified the most frequent PVs in two native corpora in L1 English, providing a general descriptive depiction of the most frequent 150 PVs in British and American English.

More importantly, apart from the detailed perspectives of the aforementioned studies, up to now, there is a gap in the literature investigating PVs through corpus-based analysis within the scope of decomposability and transparency nature and presenting a comparative analysis of ENL corpus, BNC, and ELF corpus, VOICE. To this end, the current study tries to accomplish this mission by closing the gap of VOICE corpus-based analysis of PVs in its own right in order to form a basis for the prospective PV studies conducted in line with ELF framework, thereby providing explanatory adequacy to the PV phenomenon.

CHAPTER 5

DATA AND METHODOLOGY

This chapter presents the corpus data and methodology related to the current study. Research questions and hypotheses are proposed. The elaboration of participants and corpus data is provided. In addition, the procedures of the study are elucidated. The chapter ends by lucidly explicating the data analysis of the present study.

5.1 Research questions and hypotheses

The current study aims to explore the use of PVs and analyze the decomposability and transparency nature of PVs through two representative corpora in terms of native and non-native speakers of English, namely, native speakers of English in British National Corpus (BNC) and non-native speakers of English in Vienna Oxford International Corpus of English (VOICE), thereby delving into the usage-based patterns of PVs on a contrastive basis and contributing to the literature by closing an important gap in the decomposability and transparency of PV uses within ELF framework. It investigates the natural uses of PVs in authentic oral communication which denotes natural instances of PV uses, indicating the real language performance of the speakers. To this end, the study intends to evidence explanatory adequacy by resolving the following research questions:

1. Is there a similarity or difference in the frequency of the uses of PVs between the interactions of ENL and ELF speakers?
    a. What is the frequency of the use of PVs in the ENL corpus?
    b. What is the frequency of the use of PVs in the ELF corpus?

2. How can the constructional patterns of PVs be explained within the scope of decomposability nature?
    a. To what extent do PVs show decomposability in the ENL corpus?
    b. To what extent do PVs show decomposability in the ELF corpus?
3. How can the constructional patterns of PVs be explained within the scope of transparency nature?
    a. To what extent do PVs show transparency in the ENL corpus?
    b. To what extent do PVs show transparency in the ELF corpus?
4. What are the similarities or differences among the use of PV structures between the ENL interactions and the ELF interactions through two corpora?

In line with the research questions, the study ponders the hypotheses being formulated based on these questions. The first question asks about the frequency of PVs in ENL and ELF interactions through corpora. Concerning this, it was hypothesized that the frequency of PVs would be more or less similar in terms of percentages in ENL corpus and those in ELF corpus since ELF can be accepted as an emergent language through which ELF users use the PV structures for the effectiveness of mutual intelligibility as a sign of convergence and linguistic proficiency, inevitable as a consequence of the unprecedented language spread of English, by drawing on the previous ELF research, being explicated in the literature review. In order to find out the frequency of PVs in ENL and ELF interactions, two corpora analyses were conducted with respect to the first question.

The second research question aims to explore the constructional patterns of PVs in terms of decomposability. For the purpose of resolving this question, another corpora analysis was realized in both ENL and ELF corpus. Pertaining to the second research question, it was hypothesized that in ENL corpus, native speakers of

English would use PVs in a much more decomposable manner by exposing them to greater syntactic flexibility as a consequence of natural language processing. It was further hypothesized that ELF users would demonstrate the use of PV structures in a nondecomposable manner by involving PVs in their conversations as "unanalyzed chunks" (Alptekin, 2011, p. 157) and using them in a holistic manner.

Regarding the third question, it was hypothesized that the transparency nature would be a sign of difference between the interactions of ENL and ELF speakers. It was hypothesized that the ENL speakers would use PVs both transparent and opaque within the transparency nature, whereas the ELF users, who incorporate the "declaratively-governed" semantic knowledge representations (Alptekin, 2011, p. 159), would tend to use PVs in their transparent forms in terms of semantic analysis.

Finally, in terms of the fourth and last question, it was hypothesized that the structures of the PVs in ENL and ELF interactions would be different which might be illustrated with corpora analyses. This question was investigated based on the previous research questions, by taking the patterns of PV use as a baseline.

5.2 Participants

Participants were the speakers who participated in the corpus project of two corpora, namely BNC and VOICE, representing two speakers as ENL and ELF speakers. According to Hoffmann, Evert, Smith, Lee, and Prytz (2008), since spoken language constitutes the greater part of everyday language, a corpus consisting of a large proportion of natural spontaneous interactions can be expected to be a representative corpus of actual language production. In order for a representative spoken corpus to be obtained, the main sampling procedure in BNC was to choose "a number of

speakers who would carry a portable tape-recorder and record their conversations over a certain period of time" (Hoffmann et al., 2008, p. 32). A broad array of speakers was sampled, which demonstrated equal proportions of people reflecting the distribution of people in the country. They were called ‘respondents’.

According to the Users Reference Guide for the British National Corpus (Oxford Text Archive, last updated in 2015), 124 respondents participated in the spoken component of BNC. All the respondents were British English speakers being sampled in the United Kingdom. Concerning the age, gender and social group of the participants, the respondents were selected as the sample from the population by conducting personal interviews via random location sampling procedures (Oxford Text Archive, last updated in 2015). In the formation of the sampling group, Hoffmann et al. (2008) stated that the purpose was not to have equal numbers of people from each category, but to attain a mixture of people in the country as a whole. The participants were recruited in a balanced manner in terms of the amount of texts the participants recorded, as seen in Table 1. These respondents formed the demographically sampled component of the spoken BNC. Table 1 indicates the word units of the texts the respondents recorded and their percentages.

Table 1. Distributions of Spoken Texts and Word-Units Across Age Groups

Respondent Age Groups	Texts	W-Units	%
Respondent Age 0-14	26	267,005	6.30
Respondent Age 15-24	36	665,358	15.71
Respondent Age 25-34	29	853,832	20.16
Respondent Age 35-44	22	845,153	19.96
Respondent Age 45-59	20	963,483	22.75
Respondent Age 60+	20	639,124	15.09

Note: The table was adapted from the website Users Reference Guide for the British National Corpus (Oxford Text Archive, last updated in 2015).

The recruited respondents were requested to take down a record of the people they talked to and note the specific information about their interactants such as their age, gender and any other information about them. Overall, approximately 4.2 million words as part of actual spontaneous conversations were recorded to form the demographically sampled component of the corpus.

As well as the demographically sampled component of the spoken corpus, there was context-governed component of the BNC spoken having been compiled in particular settings and contexts for actual spontaneous conversations according to Hoffmann et al. (2008). The recordings of the respondents were noted down and formulated into different contextualized domains. Context-governed texts consisted of lectures, tutorials, meetings, etc. which displayed variations based on the settings.

As well as the respondents being recruited as the ENL speakers in the demographically-sampled component of the spoken BNC, there were the interactants of the respondents, forming a total of 5,394 speaker-IDs, involving in the spoken BNC though the characteristics of the speakers were not specifically noted down. However, all the spoken material in both demographically-sampled and context-governed part was categorized as male and female speech, shown in Figure 7.

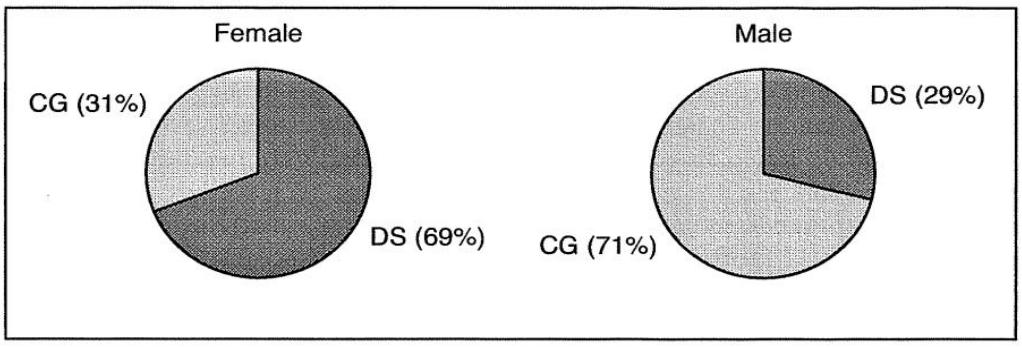

Fig. 7 Male and female speech in the spoken component of BNC
Note: CG and DS stand for context-governed and demographically-sampled materials, respectively, in the subsets of the male and female interactions in BNC (Hoffmann et al., 2008, p. 37).

With regard to the conversations of ELF speakers, the spoken corpus, VOICE, represents the actual language performance of non-native speakers of English. Seidlhofer et al. (2013) put forward that VOICE tries to maintain "the balance by providing a sizeable, computer-readable corpus of English as it is spoken by this non-native speaking majority of users in different contexts".

The participants in VOICE were experienced ELF speakers, the interactions of whom were recorded in the VOICE project. 1,260 ELF speakers were recruited in VOICE with approximately 50 different L1s (Corpus Description section, par. 2). Specifically, in terms of the ultimate aim of the VOICE project, Seidlhofer et al. (2013) stressed the importance of VOICE as "a corpus of spoken ELF interactions" involving actual spontaneous and non-scripted conversations of ELF speakers all around the world in order "to open the way for a large-scale and in-depth linguistic description of [the] most common contemporary use of English" (What is VOICE section, par. 4). The participants were categorized according to age in Table 2. What is eye-catching in Table 2 is that the highest number of speakers belongs to the ages between 17 and 24, teenagers and young adults.

Table 2. Distribution of Age Groups Based on Gender

Gender	Unknown	Ages 17-24	Ages 25-34	Ages 35-49	Ages 50+	Total
Unknown	2	0	0	0	0	2
Male	56	158	107	163	67	551
Female	35	358	115	143	56	707
Total	93	516	222	306	123	1,260

Note: The table was taken from Statistics VOICE 2.0 Online section in the VOICE Project website (Seidlhofer et al., 2013).

In addition, the recordings of the recruited speakers were paid attention to be categorized according to gender and divided into balanced groups in terms of their gender distribution in percentage and their total word counts, shown in Table 3. It is clear from Table 2 that the gender distribution of the participants and their word counts were arranged in a more balanced manner (cf. Figure 7).

Table 3. Gender Distribution and Word Counts Based on Gender

Gender	Word Counts	%
Unknown	8,997	0.88
Male	478,244	46.74
Female	533,024	52.09

Note: The table was taken from Statistics VOICE 2.0 Online section in the VOICE Project website (Seidlhofer et al., 2013).

Overall, the two corpora VOICE and BNC, having recruited ELF speakers and ENL speakers, respectively, reflect the actual spontaneous conversations in English for both sides. The number of the speakers in the spoken BNC (a total of 5,394 speakers) is nearly four times more than that in VOICE (a total of 1,260 speakers). Accordingly, the word units produced in BNC, nearly 4.2 million, are about four times more than those in VOICE, about 1 million, which shows a resemblance in the distribution of the speakers to the word units. However, for the distribution of gender this is not the case. While VOICE has nearly 47% of males and 52% of females in a more evenly distributed manner, BNC displays an uneven distribution as 71% of males and 31% of females in context-governed part and 29% of males and 69% of females in the demographically sampled part.

5.3 Corpus data

Two corpora have been in the limelight of the current study. These corpora are BNC and VOICE. BNC has two subcomponents, demographically-sampled and context-governed component in the spoken part. As shown in Table 4, context-governed component was divided into 'domains' while demographically-sampled one was categorized based on respondent characteristics according to Hoffmann et al. (2008).

Table 4. Spoken Component of BNC and Total Word Units Uttered

	Spoken Component	
Components	Context-Governed Component (CG)	Demographically-Sampled Component (DS)
Categorizations of the Components	Categorized by domain: • Educational and informative • Business • Institutional • Leisure	Sampled according to: • Respondent age • Respondent gender • Respondent social class • Geographical region
	6,175,896 words	4,233,962 words
Total Number of Words	10,409,858 words	

Note: The table was taken from the book *Corpus Linguistics with BNCweb* (Hoffmann et al., 2008, p. 134).

The focus of the current study is to analyze the context-governed part of the ENL spoken data, which represents the actual language performance in context. The domains of this component comprised of a wide array of speaking instances, each of which had distinctive features, leading to several different texts occurring in different contexts. Table 5 demonstrates such texts in the domains with their word counts. Table 5 also illustrates that the word counts in each domain are more or less the same amount, which enhances the representativeness of the component.

Table 5. Domains of the Context-Governed Component of BNC

Domain	Description	Word Count
Business	Company talks and interviews; trade union talks; sales demonstrations; business meetings; consultation	1,282,416
Educational and informative	Lectures, talks, educational demonstrations; news commentaries; classroom interaction	1,646,380
Leisure	Speeches; sports commentaries; talk to clubs; broadcast chat shows and phone-ins; club meetings	1,574,442
Public or institutional	Political speeches; sermons; public/government talks; council meetings; religious meetings; parliamentary proceedings; legal proceeding	1,672,658

Note: The table is from Hoffmann et al., 2008, p. 134.

In terms of VOICE, the ELF interactions being recorded contain a variety of "different speech events in terms of domain (professional, educational, leisure), function (exchanging information, enacting social relationships), and participant roles and relationships (acquainted vs. unacquainted, symmetrical vs. asymmetrical)" (Seidlhofer et al., 2013). Moreover, VOICE is also classified according to 'speech event types' such as meetings, panels, press conferences, workshop discussions, etc. Thus, the domain categorizations and speech event type classifications of VOICE bear a resemblance to those of BNC. However, there are no categorizations about the function and participant roles in BNC. Table 6 points out the domains of VOICE.

Table 6. Domains of VOICE

Domain	Speech Events	Words	%
ED (educational)	35	261,003	25.51
LE (leisure)	26	101,216	9.89
PB (professional business)	23	203,421	19.88
PO (professional organizational)	41	354,602	34.66
PR (professional research and science)	26	102,945	10.06

Note: The table is taken from Seidlhofer et al. (2013).

All in all, VOICE seeks to figure out spoken ELF interactions, having domains of these ELF interactions being categorized according to their functions, which illustrates spontaneous language performance of speakers from different L1 backgrounds all over the world, in which English is used "as the means of communication among [these speakers]" (Seidlhofer, 2005). Thus, what was intended in the corpus data compilation of the spoken component of BNC for the clear illustration of the interactions of ENL speakers similarly corresponds to what was intended in the corpus data compilation of the spoken ELF corpus, VOICE, for the depiction of the interactions of ELF speakers, which indicates that they complement each other as two sides of the English language medallion.

5.4 Procedure

The current study involves a contrastive corpus-based analysis, examining two spoken corpora representing ENL and ELF speakers to find out the use of PVs in terms of decomposability and transparency nature, as part of its procedure.

In the first phase, my motivation came from the immense development of ELF context and the emergence of VOICE as a spoken ELF corpus thanks to the unprecedented spread of English. Thus, a contrastive corpus-based analysis was necessary to draw a line between the conversations of ENL and ELF speakers by designating the use of linguistic structures, investigating two dichotomous corpora, to depict similarities and differences in this aspect. For the representation of the interactions of ELF speakers, VOICE was determined. Along with the illustration of ELF speakers, there was a need for a representative corpus of the interactions of ENL speakers for contrastive analysis. This called for the spoken component of BNC.

In the second phase, after the two corpora were decided, the "long-neglected" (McArthur, 1989) linguistic items in the field, namely phrasal verbs (PVs), were aimed for analysis. Since PVs have recently attracted an increasing attention in ENL research by several researchers, including Gardner and Davies (2007) who determined the most frequent PVs in the ENL use, PVs were targeted as the linguistic items for contrastive corpus analyses in this thesis so that they would not be 'neglected' any more but could be analyzed in depth.

In the third phase, since PVs were recognized in a variety of categorizations such as phrasal verbs, prepositional verbs, phrasal prepositional verbs, verb particle constructions, two-part verbs, multi-word verbs etc. by many researchers (Biber, Johansson, Leech, Conrad, & Finegan, 1999; Blais & Gonnerman, 2013; Darwin & Gray, 1999; Gardner & Davies, 2007), PV structures were to be restricted in terms of their classifications. Thus, a distinction was to be made between PVs and VPCs in this respect. Since VPCs do not possess any syntactic flexibility and they are not meaningful semantic units, the study eliminated VPCs from its analysis. The preposition in VPCs does not denote any semantic entity, but it functions as morphological complement of the verb in the sentence. Thus, it was decided to further PVs, whose particle can act as a semantic entity on its own and is exposed to syntactic flexibility complying with the syntactic nature of the lexical verb, thereby attaining new lexical meanings and different syntactic constraints rather than VPCs.

In the fourth phase, what types of PVs were going to be analyzed was determined. PVs were selected in the light of the study of Gardner and Davies (2007). The 20 most frequent lexical verbs and 16 adverbial particles were chosen following the footsteps of Gardner and Davies (2007). Although Gardner and Davies (2007) analyzed the whole BNC, including the spoken component (10% of BNC)

and the written component (90% of BNC), the current study only focused on the spoken BNC. The same PVs analyzed in the spoken BNC were also analyzed in VOICE to see the same phenomenon from the ELF counterpart.

In the fifth phase, the instruments of the analysis were determined for the examination of two corpora. For the corpus analysis of VOICE, the concordance program AntConc 3.4.4 was used to determine the frequency and decomposability nature of PVs in the ELF framework. For the corpus analysis of BNC, the full-fledged on-line concordance program of the corpus itself, BNCweb, was used to determine the frequency and decomposability nature of PVs in the ENL counterpart.

In the sixth phase, the domains of two corpora were restricted and analyzed accordingly. The spoken component of BNC and the spoken corpus VOICE were analyzed according to certain domains. In terms of counterbalancing, the same domains were chosen in both of the corpora[7].

In the seventh phase, the frequency of the most frequent 320 PVs and their decomposability nature were formulated in both corpora.

In the last phase, transparency nature of PVs was assessed by qualitative analysis of the PV occurrences in the corpus data.

All in all, the current study analyzed the use of the most frequent PVs by conducting an in-depth corpus-based analysis in order to compare the use of these target items in ENL and ELF corpora and examined the frequency, decomposability and transparency characteristics of PVs through the actual spontaneous language performance of ENL and ELF interactants, thereby providing explanatory adequacy to the usage patterns of the most frequent PVs with a socio-cognitive perspective.

[7] For further explications on the data analysis, see section 5.5.

5.5 Data analysis

The data analysis was realized through the corpus-based analyses of BNC and VOICE. The analyses were conducted through online concordance program BNCweb for the former and AntConc 3.4.4 concordance program for the latter. PVs were determined taking the most frequent PVs, which were formulated according to the study by Gardner and Davies (2007), as the baseline of the study.

Similar to what Gardner and Davies (2007) constructed, the most frequent PVs were determined via the juxtaposition of 20 lexical verbs with 16 particles. Overall, 320 PVs were formulated and analyzed in the BNC and VOICE corpora.

In line with the data analysis, to attain counterbalancing, the corpus domains were also restricted in two corpora. First of all, the written component of BNC was taken out from the analysis and focus of the study because the main objective was to elaborate on the spontaneous use of PVs in authentic oral communications of the speakers. VOICE constituted the ELF basis of the interactions, which naturally occurred among speakers actively using English language as the medium of their communications, within ELF framework as a spoken corpus. BNC constituted the ENL basis of the interactions via its spoken component. Secondly, the domains were chosen according to the level of natural language use as a possibility. In VOICE, the domains ED (education) and LE (leisure) were determined for the analysis of the use of PVs among ELFers. Moreover, for the purpose of counterbalancing, the spoken component of BNC was also reduced to context-governed component, and in that component ED and LE domains were selected. Therefore, the similar domains were selected in both corpora, which enabled a cross-reference comparison between the interactions of ENL and ELF speakers. Table 7 depicts the word counts of ED and LE domains in BNC spoken component and VOICE and the representative

percentages of these two domains altogether when compared to the entire corpora which include context-governed part of the spoken component for BNC and VOICE corpus as a whole.

Table 7. Target Domains of BNC and VOICE and Their Word Counts

Corpora	Domains	Word Counts	Total Word Counts	Representative % in the Corpus
BNC Spoken	ED	1,646,380	3,220, 822	30.9 %
	LE	1,574,442		
VOICE	ED	261,003	362,219	35.4 %
	LE	101,216		

Note: Total word counts and representative percentages of the two domains in each corpus were calculated taking word counts of ED and LE domains in Hoffmann et al. (2008) and VOICE Project (2013) into consideration.

As seen from Table 7, ED and LE domains were selected and calculated as 3,220,822 for these two domains in the whole spoken component of BNC as 10,409,858 words, yielding representation of ED+LE domains 30.9% of the spoken BNC. In terms of VOICE as a spoken ELF corpus, ED and LE domains were calculated as 362,219 for these two domains in VOICE. The whole corpus yielded 1,023,187 words in total. Thus, the representation of ED+LE domains in terms of percentage corresponded to 35.4% of the spoken ELF corpus, VOICE. Thus, the representative percentages of target domains which included ED and LE domains in both BNC spoken and VOICE corpora displayed similar percentages of the entire corpora, nearly one out of three as a whole. However, target domains (ED+LE) of BNC spoken were ten times bigger than target domains (ED+LE) of VOICE, being composed of 3,220,822 words for the former and 362,219 words for the latter.

As part of data analysis, the items were entered and searched via grammatical tagging into the concordance program of BNC, BNCweb. First, the frequency data analysis was conducted for PVs. The keyword structure being entered

to the concordance programs was ‘{set/V} * * * * up’. Adding the grammatical tagging of ‘/V’ represented the idea that only the verb form of 'set' would be analyzed in the concordance and not any other grammatical form, thereby eliminating the possibility of other form usages such as the noun form ‘setup’. Moreover, since the present simple, past simple and past participle forms of the verb ‘set’ do not change, this did not place any difficulty in analyzing these kinds of verbs whose internal grammatical form, their V2 and V3 forms, do not change.

However, for the analysis of lexical verbs (LVs) such as pick, go, give, etc., an additional keyword structure formation was to be applied in the analysis. This other keyword structure was ‘pick*_V * * * * up’ for the changes of the verb ‘pick’ such as ‘picks, picked, picking’, ‘go*_V * * * * over’ and ‘went_V * * * * over’ for the changes of the verb ‘go’ such as ‘goes, went, gone, going’, ‘g*ve*_V * * * * out’ for the changes of the verb ‘give’ such as ‘gives, gave, given, giving’, to illustrate both tense markings and gerund formations of the verbs in PV structures. Even so, the analysis by using this keyword structure yielded the results of both particles as adverbial particle of the verb among PVs and prepositions as prepositional phrases of the verb among VPCs. Thus, the last and more reliable keyword structure was formulated by adding a grammatical tagging also to AVP, thereby transforming the keyword formula into ‘g*ve*_V * * * * out_AVP’, which would lead to all adverbial particles of PVs tagged as AVP in the concordance accurately.

Still, an in-depth qualitative analysis was essential to determine which structures were really adverbial particles (AVPs) and which structures were prepositions. Since BNC tagged some prepositional phrases (PRPs) as AVPs and vice versa, leading to fuzziness in the formulation of PVs, a qualitative close re-examination was also administered to double-check the accurate use of AVPs.

The reason why four stars were situated between LV and AVP in the keyword structure of the entrance of each PV in the online concordance program BNCweb was to be able to determine the decomposability nature of the same PV structures. As what was emphasized by Gardner and Davies (2007), the maximum word limit which could be placed between LV and AVP was four because separating LV and AVP with the use of more than four words in the intermediary level might cause "false phrasal verbs beyond the four-word scenario" (Gardner & Davies, 2007, p. 345). Thus, four stars were placed between LV and AVP, leading to the analysis of PVs within six-word structures at the top-level design. This type of analysis would inevitably lead to many unrelated false positives, so another cross-check was also administered to get rid of false positives.

A similar analytical process was also administered to ELF counterpart of the analysis of the use of PVs in VOICE corpus. However, since VOICE 2.0 XML version which was transformed from VOICE 2.0 Online was downloaded without part-of-speech tagging, a closer corpus-based analysis was necessary for the PV structures to be accurately analyzed. The latest of the ELF corpus was downloaded and it was analyzed through AntConc 3.4.4 concordance program. Except for the use of grammatical tagging, the PV structures were analyzed by entering the following formula ‘LV + AVP’, ‘LV + * + AVP’, ‘LV + * * + AVP’, ‘LV + * * * + AVP’ and lastly ‘LV + * * * * + AVP’, and all the aforementioned formula was repeated for the tense-aspect marking changes and gerund forms of the same PV structures. Thus, the base form in simple present tense, the simple present tense form with the addition of 3rd person -s marking, simple past tense form, past participle form, gerund form of the same PVs were entered into AntConc and analyzed through 5 stages in non-decomposable adjacent position, one word in between, two words in between, three

words and four words in between. For instance, a PV structure was analyzed through five separate words-in-between searches and for each of these searches, another sublevel five-stage tense marking and gerund formation level of analysis for each PV. This analysis led to 25 combinatorial analyses for each PV structure for the purpose of frequency determination.

Within the scope of decomposability analysis, in order to define the basic conceptualizations of decomposability paradigm, an American-English native instructor teaching in the Foreign Language Education department at one of the universities in Turkey in which the medium of instruction is English was referred to for the distinctions of what kinds of PV structures could be decomposable, nondecomposable or both. In order to double-check, a part-time professor teaching in the same department at the same university, representing ELF counterpart of the issue, went through the most frequent PVs which were obtained through the juxtaposition of 20 LVs and 16 AVPs. The decomposability analysis paradigm was situated within a cline, having three subcomponents. Thus, the use of PVs in ENL and ELF interactions was analyzed according to whether PVs were used in a nondecomposable position in which AVP was adjacent to LV or PVs were used in a decomposable position in which LV and AVP were separated by one word to four words based on the syntactic constraints of the PV structures or as the third hypothesis, they were used in a non-decomposable position although they were also convenient for the use of decomposable position.

Within the scope of the analysis of transparency nature, PVs were categorized according to their semantic entities. Transparency was determined as an analysis factor in which the meaning of LV and AVP contributed to the overall meaning of PVs or not. If the overall semantic meaning could be interpreted by

understanding the meanings of LV and AVP individually, these PVs would be accepted as transparent. If the overall meaning was not overt by the individual meanings of LV and AVP, then these PVs would be accepted as opaque. However, an intermediary level was also formulated for the transparency nature of PVs. If the semantic entity of LV and AVP led to both transparent meaning and another opaque meaning, these items were regarded as semi-transparent PVs.

All in all, the data analysis was conducted through multi-directional corpus-based analyses to reach a contrastive analysis of the use of PVs in terms of their frequencies, decomposability and transparency nature between the spontaneous authentic language performance of ENL and ELF speakers via two spoken domains of two corpora, BNC for the former and VOICE for the latter.

5.6 An alternative approach to decomposability and transparency distinctions

Within the scope of the decomposability and transparency paradigm, the current study proposes an alternative approach to the decomposability and transparency distinctions. The thesis formulates an alternative approach to decomposability and transparency distinctions of PVs as Nondecomposable Lexical Annotation (NLA), which resulted from the data analysis in this thesis.

The studies in the literature take the idiom decomposition hypothesis as their baseline with respect to syntactic analyzability, which suggested that PVs can be decomposable and non-decomposable dependent on the particle movement. However, NLA proposes the existence of an inclination to use non-decomposable syntactic forms of the target items though these items can be perfectly used in a decomposable position, as opposed to what was suggested by Gibbs et al. (1989).

That is to say, even if there is an existence of ‘syntactic flexibility’ (the idiom decomposition hypothesis by Gibbs et al., 1989) in which decomposable items can be processed and used as analyzable items separately in the stream of conversation, NLA suggests that since there is a slight chance of making a ‘mistake’ as a risk factor in the decomposable model in syntactic terms, the non-decomposable position is much more secure for individuals to display their performance level of proficiency. In fact, this thesis is only concerned with the retrieval and performance levels of the users as the speakers of the language adopting a usage-based approach since it is a corpus-based study and all of the data are restricted to the collections of individuals' speaking, yet with an advantage of the fact that the conversations are the actual representations of unconscious processing. Thus, NLA puts forward non-decomposability as its syntactic approach, which has emerged out of the corpora analyses as an evidential finding.

With regard to the semantic representation, on the other side of the PV medallion, NLA displays resemblance with the standard pragmatic model and the production hypothesis in line with the storage of lexical items, PVs in this case, in terms of belonging to holistic units rather than separate items in the mental lexicon no matter what sort of internal syntactic features these constructions have as either decomposable or non-decomposable nature. NLA illustrates high congruence with the two hypotheses mentioned in this respect. Nevertheless, more importantly, the point at which NLA diverges from the standard pragmatic model and the production hypothesis is that NLA suggests that the lexical constructions, incorporating annotations related to their transparency nature, whether transparent or opaque, can be produced as the initial activation of opaque meaning, denoting non-decomposable, at the performance level as opposed to what was suggested in the standard pragmatic

model by stressing the importance of literal-first. Thus, lexical annotations contribute to the meaning of PVs in the context. Thus, NLA puts forward lexical annotation as its semantic approach, which has emerged out of the corpora analyses as an evidential finding.

NLA points out that decomposability is explicated in terms of the adjacency nature of AVP which either obligatorily used next to its LV or not, but with an intermediary position of 'both'. Therefore, decomposability shows a cline in three stages. Non-decomposability paradigm occurs when the PV has to be used as its LV and AVP always adjacent, thereby not accepting any object in between. Decomposability paradigm occurs when the PV can be flexibly produced (e.g. any objects such as pronoun, noun or phrase can intervene between LV and AVP). The 'both' paradigm refers to the stage which displays an intermediary structure, suitable in both decomposable and non-decomposable manner.

Similarly, NLA points out that transparency is explicated in terms of semantic representations of lexical items. Therefore, transparency occurs when PVs can be interpreted by understanding the meanings of their LV and AVP individually. Opaqueness occurs when individual meanings of LV and AVP do not contribute the overall figurative meaning of PVs. Moreover, semi-transparent PVs occur as an intermediary level, in which the semantic entities of LV and AVP contribute to the overall meaning of PVs to a degree, involving both transparent and opaque meaning.

CHAPTER 6

RESULTS

6.1 Frequency of the items

Two corpora were analyzed, namely the BNC spoken and the spoken ELF corpus, VOICE. The frequency of the PV use in the ED and LE domains of both corpora were calculated through corpus-based analyses via concordance programs, BNCweb and VOICE 2.0 XML being incorporated into AntConc 3.4.4.

The frequency results of PVs along with the percentages within 3,220,822 words in BNC spoken are illustrated in Table 8. The frequency results of PVs along with the percentages within 362,219 words in VOICE are given in Table 9.

Table 8. Frequency of PV Occurrences in BNC

20 LVs x 16 AVPs	TOTAL PV RESULTS	%
LV 'GO' x 16 AVPs	4,836	0.150
LV 'COME' x 16 AVPs	3,789	0.118
LV 'TAKE' x 16 AVPs	1,275	0.040
LV 'GET' x 16 AVPs	2,217	0.069
LV 'SET' x 16 AVPs	366	0.011
LV 'CARRY' x 16 AVPs	470	0.015
LV 'TURN' x 16 AVPs	532	0.017
LV 'BRING' x 16 AVPs	762	0.024
LV 'LOOK' x 16 AVPs	331	0.010
LV 'PUT' x 16 AVPs	1,273	0.040
LV 'PICK' x 16 AVPs	547	0.017
LV 'MAKE' x 16 AVPs	408	0.013
LV 'POINT' x 16 AVPs	99	0.003
LV 'SIT' x 16 AVPs	186	0.006
LV 'FIND' x 16 AVPs	444	0.014
LV 'GIVE' x 16 AVPs	297	0.009
LV 'WORK' x 16 AVPs	765	0.024
LV 'BREAK' x 16 AVPs	235	0.007
LV 'HOLD' x 16 AVPs	183	0.006
LV 'MOVE' x 16 AVPs	404	0.013
TOTAL	19,419	0.603

Table 8 outlines the results of PVs depicting the overall frequencies of PVs in BNC. The results indicate the occurrences of each of the most frequent 20 lexical verbs (LVs) being juxtaposed with the most frequent 16 adverbial particles (AVPs) in BNC. In total 19,419 PV occurrences are detected in the ED and LE domains of BNC. The overall percentage of these occurrences yields 0.603% in the data analysis of BNC. The most frequent five PVs consist of PVs with 'go' having 0.15%, PVs with 'come' having 0.12%, PVs with 'get' having 0.07%, PVs with 'take' having 0.04% and lastly PVs with 'put' having 0.04% with a very small difference to the percentage of 'take' PVs. The least frequent PV use includes PVs with 'point' having only 99 instances of ENL PV use within 3,220,822 words.

Table 9. Frequency of PV Occurrences in VOICE

20 LVs x 16 AVPs	TOTAL PV RESULTS	%
LV ‘GO’ x 16 AVPs	285	0.079
LV ‘COME’ x 16 AVPs	195	0.054
LV ‘TAKE’ x 16 AVPs	35	0.010
LV ‘GET’ x 16 AVPs	105	0.029
LV ‘SET’ x 16 AVPs	12	0.003
LV ‘CARRY’ x 16 AVPs	3	0.001
LV ‘TURN’ x 16 AVPs	17	0.005
LV ‘BRING’ x 16 AVPs	19	0.005
LV ‘LOOK’ x 16 AVPs	23	0.006
LV ‘PUT’ x 16 AVPs	146	0.040
LV ‘PICK’ x 16 AVPs	9	0.002
LV ‘MAKE’ x 16 AVPs	38	0.010
LV ‘POINT’ x 16 AVPs	17	0.005
LV ‘SIT’ x 16 AVPs	27	0.008
LV ‘FIND’ x 16 AVPs	44	0.012
LV ‘GIVE’ x 16 AVPs	22	0.006
LV ‘WORK’ x 16 AVPs	94	0.026
LV ‘BREAK’ x 16 AVPs	7	0.002
LV ‘HOLD’ x 16 AVPs	9	0.003
LV ‘MOVE’ x 16 AVPs	34	0.009
TOTAL	1,141	0.315

Table 9 summarizes the results of PVs depicting the overall frequencies of PVs in VOICE. The results indicate the occurrences of each of the most frequent 20 LVs being juxtaposed with the most frequent 16 AVPs in VOICE. The PV frequency analysis in the ED and LE domains of VOICE yields 1,141 PV occurrences in total. The overall percentage of these occurrences signalizes 0.315%. The most frequent five PVs consist of PVs with 'go' having 0.079%, PVs with 'come' having 0.054%, PVs with 'put' having 0.04%, PVs with 'get' having 0.029% and lastly PVs with 'work' having 0.026%. The least frequent PV use includes PVs with 'carry' having only 3 instances of ELF PV use within 362,219 words.

The combinations of each LV, involving the most frequent 20 LVs, and the most frequent 16 AVPs, according to Gardner and Davies (2007), yield 320 PVs via juxtaposition of each LV with each AVP. These combinations are also analyzed on an individual basis through two corpora. The frequency lists of PV occurrences are displayed accordingly. Besides, 16 AVPs are divided into two groups, forming 8 more frequent AVPs and 8 less frequent AVPs, leading to the formations of two tables for each corpus. Therefore, Table 10 outlines the frequency list of the most frequent 160 PVs, formulated with the combination of 20 LVs and 8 more frequent AVPs, analyzed within 3,220,822 words in ED and LE domains of BNC, along with the percentages of the PV use, being formed of 8 more frequent AVPs. In addition, Table 11 displays the frequency list of the other 160 PVs including 8 less frequent AVPs, analyzed within 3,220,822 words in the same domains of BNC, along with the percentages of the PV use, being formed of 8 less frequent AVPs.

Table 10. Frequency of 160 PVs with 8 More Frequent AVPs in BNC

	Out	Up	On	Back	Down	In	Off	Over	Total	%
GO	608	551	1,277	792	521	288	134	79	4,250	0.132
COME	620	596	275	695	355	648	54	77	3,320	0.103
TAKE	287	189	128	88	83	60	164	207	1,206	0.037
GET	585	268	342	334	166	211	92	33	2,031	0.063
SET	55	248	4	6	5	10	27	0	355	0.011
CARRY	165	5	260	3	6	5	1	4	449	0.014
TURN	128	131	27	25	37	11	38	31	428	0.013
BRING	72	276	13	98	59	183	2	12	715	0.022
LOOK	50	84	5	111	32	16	0	1	299	0.009
PUT	82	279	249	63	217	276	52	14	1,232	0.038
PICK	65	469	8	1	1	0	2	1	547	0.017
MAKE	123	242	5	6	8	7	10	1	402	0.013
POINT	95	1	0	0	2	0	0	0	98	0.003
SIT	5	27	8	7	119	7	0	2	175	0.005
FIND	424	4	1	1	0	9	0	0	439	0.014
GIVE	55	164	0	35	0	26	15	1	296	0.009
WORK	584	32	62	6	15	38	3	1	741	0.023
BREAK	36	71	0	1	109	6	3	1	227	0.007
HOLD	18	66	54	22	16	5	0	1	182	0.006
MOVE	75	29	137	26	19	45	5	10	346	0.011
TOTAL	4,132	3,732	2,855	2,320	1,770	1,851	602	476	17,738	0.551

Table 11. Frequency of 160 PVs with 8 Less Frequent AVPs in BNC

	Round	About	Through	Around	Along	Under	By	Across	Total	%
GO	199	32	144	46	151	2	2	10	586	0.018
COME	127	39	76	11	187	4	1	24	469	0.015
TAKE	9	43	3	3	10	0	0	1	69	0.002
GET	61	10	74	14	13	0	3	11	186	0.006
SET	0	11	0	0	0	0	0	0	11	0.0003
CARRY	3	4	6	7	1	0	0	0	21	0.0007
TURN	92	0	2	10	0	0	0	0	104	0.003
BRING	5	33	1	2	5	0	0	1	47	0.002
LOOK	9	1	10	10	1	1	0	0	32	0.0009
PUT	13	4	18	1	2	2	0	1	41	0.0013
PICK	0	0	0	0	0	0	0	0	0	0
MAKE	2	0	0	0	4	0	0	0	6	0.0002
POINT	1	0	0	0	0	0	0	0	1	0.00003
SIT	3	0	0	8	0	0	0	0	11	0.00034
FIND	0	0	2	3	0	0	0	0	5	0.0002
GIVE	0	0	0	0	1	0	0	0	1	0.00003
WORK	8	0	9	2	3	2	0	0	24	0.0007
BREAK	1	0	7	0	0	0	0	0	8	0.0002
HOLD	0	1	0	0	0	0	0	0	1	0.00003
MOVE	12	10	5	28	3	0	0	0	58	0.002
TOTAL	545	188	357	145	381	11	6	48	1,681	0.052

For the illustration of PV occurrences among the interactions of ELF users, VOICE analysis of the PV use leads to two tables designating the frequency of the same 320 PVs being analyzed through similar procedures in VOICE. In this respect, the frequency of the use of 160 more frequent PVs is provided in Table 12, along with the percentages of these 160 PVs within 362,219 words in ED and LE domains of VOICE. Moreover, the occurrences of the other 160 PVs with 8 less frequent AVPs are given in Table 13.

Table 12. Frequency of 160 PVs with 8 More Frequent AVPs in VOICE

	Out	Up	On	Back	Down	In	Off	Over	Total	%
GO	56	22	79	31	22	24	1	5	240	0.066
COME	16	52	36	41	1	27	0	3	176	0.049
TAKE	10	2	6	1	1	3	2	8	33	0.009
GET	23	13	14	9	0	36	0	2	97	0.027
SET	0	12	0	0	0	0	0	0	12	0.003
CARRY	1	0	1	1	0	0	0	0	3	0.0008
TURN	2	2	5	0	0	1	2	1	13	0.004
BRING	4	7	0	3	0	4	0	1	19	0.005
LOOK	3	4	2	3	1	5	0	1	19	0.005
PUT	2	6	43	2	12	78	0	1	144	0.04
PICK	1	8	0	0	0	0	0	0	9	0.003
MAKE	5	13	5	0	0	12	0	0	35	0.01
POINT	15	0	1	0	0	0	0	0	16	0.004
SIT	1	0	4	0	4	16	0	1	26	0.007
FIND	24	0	4	1	2	12	0	0	43	0.012
GIVE	3	9	3	3	0	3	0	0	21	0.006
WORK	7	1	42	0	0	41	0	0	91	0.025
BREAK	0	0	0	0	6	0	1	0	7	0.002
HOLD	0	0	9	0	0	0	0	0	9	0.0025
MOVE	2	0	22	2	0	4	0	0	30	0.008
TOTAL	175	151	276	97	49	266	6	23	1,043	0.288

Table 13. Frequency of 160 PVs with 8 Less Frequent AVPs in VOICE

	Round	About	Through	Around	Along	Under	By	Across	Total	%
GO	2	3	17	6	8	0	8	1	45	0.012
COME	1	1	3	1	8	1	2	2	19	0.005
TAKE	1	1	0	0	0	0	0	0	2	0.0006
GET	0	1	1	0	2	0	1	3	8	0.0022
SET	0	0	0	0	0	0	0	0	0	0
CARRY	0	0	0	0	0	0	0	0	0	0
TURN	2	0	0	2	0	0	0	0	4	0.0011
BRING	0	0	0	0	0	0	0	0	0	0
LOOK	0	0	1	2	0	1	0	0	4	0.0011
PUT	0	0	0	0	0	2	0	0	2	0.0006
PICK	0	0	0	0	0	0	0	0	0	0
MAKE	0	1	0	0	0	0	2	0	3	0.0008
POINT	0	0	1	0	0	0	0	0	1	0.0003
SIT	0	0	0	0	0	0	1	0	1	0.0003
FIND	0	1	0	0	0	0	0	0	1	0.0003
GIVE	0	1	0	0	0	0	0	0	1	0.0003
WORK	0	0	1	0	1	0	1	0	3	0.0008
BREAK	0	0	0	0	0	0	0	0	0	0
HOLD	0	0	0	0	0	0	0	0	0	0
MOVE	0	0	2	2	0	0	0	0	4	0.0011
TOTAL	6	9	26	13	19	4	15	6	98	0.027

6.2 Syntactic analysis results[8]

Within the scope of decomposability nature, PVs were analyzed in BNC and VOICE, thereby delving into a closer examination of the authentic use of PVs, which included two dichotomous ends formed of decomposable and non-decomposable. Decomposable PVs involve the structures whose LV and AVP could be separated by adding any other items between the two. Nondecomposable PVs, on the other hand, involve PV structures whose LV and AVP could only be used in an adjacent position and could not have any items between the two. The current study also hypothesizes

[8] The syntactic analyses tables are provided in Appendix B.

an intermediary dimension, namely 'both' condition in which PV structures could be used in both decomposable and non-decomposable positions.

Syntactic analysis results of PV structures[9] are calculated by means of the quantification of decomposable and nondecomposable PVs and are formulated through individual analyses of each syntactic form of PV items (320 PVs in total) in two corpora representing ENL and ELF side of the PV medallion. Syntactic forms include the base form of LV, the simple present tense form with the addition of third person marking '-s', simple past form, perfect form, and gerund form. Each syntactic form is analyzed for each PV with their non-decomposable position entrance into the concordance program and their entrances of decomposable positions including one word, two words, three words and four words in between LVs and AVPs. The unrelated section denotes false positives which include either VPCs or consist of other structures coincidentally uttered together which have no association with the target PVs being analyzed.

Overall, 320 PVs are analyzed one by one in the ED and LE domains of BNC spoken and VOICE by being transformed into separate columns either decomposable or nondecomposable, constituting total PV uses in another column with the sum of decomposable and nondecomposable PVs and depicting unrelated false positives which are not included into the frequency of total PVs. However, unrelated false positives are important to point out the error rate of each corpus.

The overall results of the syntactic analyses of PVs are illustrated in Appendix B, Table 14 for ENL PV uses in BNC and Appendix B, Table 15 for the ELF PV uses in VOICE. Within the scope of decomposability nature, the frequency of decomposable and non-decomposable forms of each PV use is measured and

[9] For the explications of the results and further deductions, see Chapter 7.

demonstrated in separate columns in both Table 14 in Appendix B displaying the results of BNC and Table 15 in Appendix B displaying the results of VOICE. The frequency lists of total PV occurrences are also placed in both tables. The frequency lists of the unrelated false positives are also situated at the right side of both tables. In addition, the percentages of the decomposable and nondecomposable PV use, the percentages of their accumulations yielding total PV occurrence ratios, and lastly the percentages of the unrelated false positives are given in Table 16 and Table 17 in Appendix B, displaying the overall relative percentages of 320 PVs in BNC within 3,220,822 words and in VOICE within 362,219 words, respectively.

6.3 Semantic analysis results[10]

Within the scope of transparency nature, PVs were analyzed in BNC and VOICE, drawing upon the semantic analyses of each PV item and assessing their frequency in the ED and LE domains of BNC and VOICE. PVs are categorized in terms of their transparency characteristics by pondering upon the overall meaning comprehension of PVs, to which the semantic entities of LVs and AVPs contribute or not, thereby yielding two dichotomous ends, transparency and opaqueness.

Transparent PVs consist of PV structures whose overall meanings can be interpreted by figuring out both of the semantic entities of their LV and AVP individually or either of them. Opaque PVs, on the contrary, consist of PVs whose overall meanings are not overt by making sense of the individual meanings of LV and AVP being denoted and in which the semantic entities, each of these items represents, possess contextual figurative meanings as a whole. However, in addition

[10] The semantic analyses tables are provided in Appendix C.

to these two dichotomous ends, the current study constructs an intermediary level for the transparency nature of PVs, in a similar manner to what it proposes as an intermediary level for the decomposability nature of PVs, thereby leading to the formulation of semi-transparent PVs in which the overall meaning of PVs can be understood by interpreting the semantic entities of LV and AVP to an extent due to their possession of both transparent and opaque meaning.

Overall, 320 PVs were categorized with regard to their transparency nature and analyzed accordingly. Within the scope of designating transparency nature of each PV structure, Oxford Dictionary of Phrasal Verbs (ODPV) (1993) was utilized for the overall meaning of PVs. Table 18 and Table 19 in Appendix C indicate the semantic analyses of each PV item being categorized based on their overall lexical meaning into three groups according to whether these PVs are transparent, semi-transparent or opaque. All of these are sound categories for PVs being analyzed through their frequency in the ED and LE domains of both BNC and VOICE[11].

In this respect, the semantic analysis of 160 PVs comprising the most frequent 20 LVs juxtaposed with 8 more frequent AVPs is given in Appendix C, Table 18. In Appendix C, Table 19, the semantic analysis of the other 160 PVs including the most frequent 20 LVs with 8 less frequent AVPs is indicated. The overall distributional results of the transparency categorizations of 320 PVs are also shown in Appendix C, Table 20.

In Appendix C, Table 18 and Table 19, in addition to transparency categorizations of each PV, there are also non-existent PV forms whose uses are also interestingly witnessed in both corpora, which are evidenced in the frequency of these non-existent PVs. Taking Table 18 and Table 19 (see Appendix C) into

[11] For further explanations of the semantic analyses of PVs and their cross-reference comparison between two corpora, see Chapter 7.

consideration, the transparency categorizations of the target PVs comprising 320 PVs in line with the semantic analysis yield 39 ‘transparent’ PVs shown as TPV, 73 ‘semi-transparent’ PVs shown as STPV, 89 ‘opaque’ PVs shown as OPV and 119 PVs of ‘non-existent’ forms which do not actually appear in ODPV (1993), shown as NE. Figure 8 depicted the overall distribution of transparency categorizations of 320 PVs. The percentages of these categorizations are indicated in Figure 8.

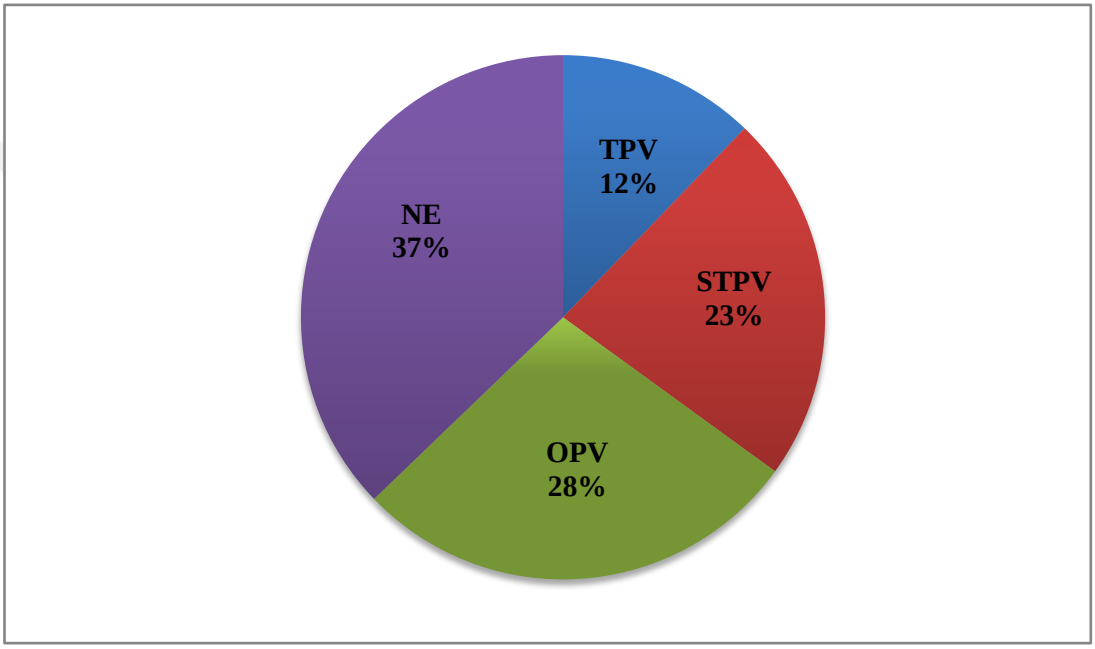

Fig. 8 Overall distribution of transparency categorizations of 320 PVs

CHAPTER 7

DISCUSSION

The current study has analyzed the use of PV structures within the scope of their frequency, decomposability and transparency nature through the spontaneous authentic language performance of ENL and ELF speakers, evidenced by the two representative corpora, BNC spoken component and VOICE as a spoken ELF corpus, by trying to delve into the usage patterns of 320 PVs, and hence providing an explanatory adequacy to the ENL and ELF use of PV structures.

The study has specifically focused on the use of English PVs on a contrastive basis to find out the answers of the research questions[12] about whether there is a similarity or difference in the frequency, decomposability and transparency characteristics of the PV use between the interactions of ENL and ELF speakers.

Corpus-based analyses results conducted through the ED and LE domains of BNC and the same domains of VOICE have indicated significant findings shedding some light on the ENL and ELF use of PVs which will be scrutinized in the current chapter through making cross-reference comparisons between the two corpora.

Taking the first research question into consideration, whether there was a similarity or difference in the frequency of the PV uses between the interactions of ENL and ELF speakers was questioned. Frequency of the PV uses in ENL and ELF corpora was intended to be figured out. In addition to what this first research question aimed to investigate, it was hypothesized that the frequency of PV uses between ENL and ELF speakers would be almost similar in percentages because ELF has become an emergent language being highly used all around the world by

[12] For the research questions and hypotheses of the study, see section 5.1.

numerous speakers with different language backgrounds to maintain mutual intelligibility, taking on a convergent position in relation to ENL use. However, the frequency results evidenced this situation vice versa.

The frequency results of the PV uses have yielded the overall percentages 0.603% for the frequency of ENL use of PVs in BNC and 0.315% for the frequency of ELF use of PVs in VOICE. These two overall percentage results have pointed out that the ENL use of PVs were twice as many as the ELF use of these target items. To this end, the study could propose the idea that within 1,000 words, ENL speakers could use six PVs in their stream of conversations with their interlocutors whereas within the same quantity of words, ELF speakers could use only half of the ENL PV use, namely three PVs in their stream of conversations.

The overall frequency of the PV uses by ENL and ELF speakers could elucidate the lucid presence of ENL PV use at a greater amount of prevalence rather than that of ELF PV use, which has, in fact, signified the existence of avoidance behavior in the frequency of PV use among ELF speakers in their spontaneous language performance in comparison to ENL speakers, evidenced by the quantification of the target items via the cross-reference checks between two corpora.

Nevertheless, with regard to what Gardner and Davies (2007) have proposed in their study, approximately one PV structure could occur out of every 400 words of English in the ENL use of PVs in both written and spoken components of BNC as a whole corpus. The study of Gardner and Davies (2007) has also interestingly assumed VPCs under the category of PVs, at which point displays divergence from the objective of the current study. Even so, the average native use of PVs in both spoken and written components in BNC has been indicated in the study of Gardner and Davies (2007) as roughly three PVs within 1,000 words (0.25 in percentage out

of 100). The average ENL PV use explained by Gardner and Davies (2007) equals to the overall percentage of ELF PV use determined as one of the frequency results of the current study through two corpus-based analyses. In addition, the frequency of the ENL PV use, namely six PVs within 1,000 words, was found to outnumber twice as many as the PV results of Gardner and Davies (2007) as the frequency of PVs for the ENL use. In fact, Gardner and Davies (2007) have analyzed all the components of BNC, which is different from the focus of this study, and Gardner and Davies (2007) have also included VPCs under PVs and found different results.

All in all, regarding the first research question, the current study could pose the fact that there was a difference in the frequency of PV use between ENL and ELF speakers. The percentage of ELF PV use at a low degree might as well suggest the existence of avoidance behavior of ELF speakers. Since it could be argued that ELF speakers may attach higher importance to making mistakes in their spontaneous use of language, avoidance behavior of ELF speakers can be quite acceptable in line with the frequency results of the thesis because studies in ELF literature have also indicated that the avoidance behavior displayed by ELF speakers is mostly claimed to have stemmed from the fear of making mistakes. This causes misunderstandings in the stream of their conversations since the main objective of using ELF in its own right is to be able to maintain mutual intelligibility. As a matter of fact, the avoidance behavior, which could be put forward taking the lesser amount of the frequency of the ELF PV use, can be assumed to have resulted from the reservations of the ELF users. ELF speakers might be avoiding using PVs not to produce "variations" (Prodromou, 2007) though they can actually be full-fledged in their ELF interactions. On the contrary, the full play of PV use and higher prevalence in the frequency of ENL PV use could make an impression as if ENL speakers were using novel

"creativities" (Prodromou, 2007). To this end, the avoidance behavior of ELF speakers might suggest that their avoidance as a tendency not to use the target items can be one of the essential factors to be considered in the frequency of PVs.

Within the scope of the second research question, the constructional patterns of PVs were inquired in terms of their decomposability nature. To what extent the PV uses denoted decomposability and non-decomposability between the interactions of ENL and ELF speakers was scrutinized through the analyses of the ED and LE domains of BNC and VOICE corpora. Regarding the second research question, it was hypothesized that the interactions of ENL speakers would lead to higher amounts of PV use in a much more decomposable position rather than those of ELF speakers on a contrastive basis. ENL speakers were assumed to take on greater syntactic flexibility through which they could allow the use of PVs full play in their language as part of their natural language processing, thereby transforming the process of PV use "within their procedural knowledge" (Alptekin, 2011). However, the interactions of ELF speakers were hypothesized to lead to the lesser amounts of the PV use in a decomposable position but a higher prevalence on the use of PVs in a much more non-decomposable position as ELF speakers might display avoidance behavior in the frequency of their PV use by trying to avoid making mistakes as a result of their syntactic combination process of PV rules "within their declarative knowledge" (Alptekin, 2011), thereby constructing the PV structures as "unanalyzed chunks" (Alptekin, 2011; Ellis, 1994) and using these items in a holistic manner.

Overall decomposability analyses, involving the closer examinations of 320 PVs based on their occurrences between BNC and VOICE, ended up with significant findings about the decomposability nature of PVs, characterizing the interesting cross-reference comparison results between the two corpora.

The decomposability results of the use of PVs yielded the percentages being valid for all the occurrences of 320 PVs in BNC as 0.1325% for the decomposable PVs and 0.4705% for the nondecomposable PVs within the interactions of ENL speakers, making up the overall the frequency of the ENL PV use as 0.603% within 3,220,822 words. Moreover, the decomposability results led to the percentages being valid for all the occurrences of 320 PVs in VOICE as 0.064% for the decomposable PVs and 0.251% for the nondecomposable PVs within the interactions of ELF speakers, constituting the overall frequency of the ELF PV use as 0.315% within 362,219 words. These results could be put forward to cause some significant conceptualizations for the stance of this thesis in terms of the decomposability nature of the PV use between ENL and ELF speakers.

Taking the results of frequency of the use of PVs between ENL and ELF into consideration, the non-decomposability nature of the PV use was interestingly found to be nearly at the exact similar percentage between the interactions of both ENL and ELF speakers in the two corpora, which was one of the most significant outcomes of the current study. The percentage of the nondecomposable PVs of ENL speakers (0.47%) to the whole percentage displaying all the PV structures of ENL speakers (0.60%) resulted in the rate of 78% as the non-decomposability percentage in BNC. Interestingly, similar to the rate of ENL non-decomposability percentage of the use of PVs, the percentage of the nondecomposable PVs of ELF speakers (0.25%) to the whole percentage displaying all the PV structures of ELF speakers (0.32%) resulted in the rate 78% as the non-decomposability percentage in VOICE.

To this end, within the scope of decomposability paradigm of the use of PVs, the non-decomposability nature of the PV use could be argued to be extraordinarily similar between ENL and ELF speakers. It was evidenced by their similar rate (78%)

of the proportions in the non-decomposability use of PVs through two representative corpora, thanks to which this finding of the thesis actually clashes with all the conceptualizations put forward in the literature about the groundbreaking differences between ENL and ELF speakers and the native and non-native distinctions, and it could pose a necessity in the redefinition of the ENL and ELF conceptualizations for the PV medallion.

Consequentially, regarding the second research question, the constructional patterns of the use of PVs with respect to the decomposability paradigm of the use of PVs were concluded to be interestingly similar between the interactions of ENL and ELF speakers, to the same extent showing a similarity in the non-decomposability nature and a huge resemblance in the decomposability nature in terms of percentage. This makes an essential breakthrough in the field of PV research on a contrastive basis. Therefore, ENL speakers could be discussed to have gone through the similar processes of the use of PVs, presumably not only in their involvement in the procedural knowledge but also in their involvement in the declarative knowledge (Ullman, 2004) based on the findings of the present study, to ELF speakers in terms of their tendency to involve the extent of syntactic flexibility for the PV use in their course of interactions within decomposability nature. This, as a matter of fact, paved the way for the need of more elaborations in other layers of further analyses of the use of PV structures between the interactions of ENL and ELF speakers within different contextual backgrounds, in addition to what was aimed in this study.

Within the scope of the third research question, the constructional patterns of PVs were investigated in terms of their transparency nature. To what extent the use of PVs displayed transparency, semi-transparency and opaqueness nature between the interactions of ENL and ELF speakers was scrutinized through closer

examinations on PVs which were conducted for the purpose of the determination of transparency categorizations realized among the target 320 PV structures, the percentages of which are demonstrated in Appendix C, Table 21 and Table 22 for the corpora BNC and VOICE respectively. Table 21 and Table 22 in Appendix C predicate the underlying transparency characteristics of PV structures drawing upon the analyses of the ED and LE domains of BNC and VOICE. Pertaining to the third research question, it was hypothesized that the interactions of ENL speakers in BNC would yield higher amounts of the PV use in a far more semi-transparent and opaque nature as well as transparent rather than the interactions of ELF speakers in VOICE on a contrastive basis because the general assumption puts forward that ENL speakers can involve highly complicated semantic representations of lexical items at full length. Thanks to this, ENL speakers may give greater prominence to the intricate use of linguistic structures that also has the likelihood to consist of the use of PVs in their language performance, via actively incorporating "both the automatic processes of the implicit grammatical knowledge in the procedural system and the controlled processes of the explicit lexical knowledge in the declarative system" (Alptekin, 2011), thereby stressing the existence of both declarative and procedural knowledge representations going hand in hand for ENL speakers in their transparency use of PVs. However, the interactions of ELF speakers were expected to yield the heavy reliance on the PV use in a more transparent nature and the presumably prevalent avoidance behavior on the use of PVs in a semi-transparent and opaque nature. Since ELF speakers were postulated to disclose avoidance behavior in their use of PVs, evidenced by the frequency of ELF PV use, by keeping to "word-level operations through lexical processing and connect[ing] each word to other conceptually congruent units of meaning through associative binding . . . [and

becoming] dependent on the use of semantic universals" (Alptekin, 2011) in their lexical constructions of semantic representations, this in the end provided the analytical processes to come in handy while ELFers' making sense of PVs in transparent nature by comprehending the meanings of LVs and AVPs individually.

The results of the transparency categorizations of 320 PVs are illustrated in Appendix C, Table 20. Based on the overall distribution results of the transparency categorizations, 320 PVs of which were examined through ODPV (1993) and were classified according to their transparency nature, Table 20 in Appendix C indicates that transparent PVs have the least amount of percentage, namely 12.19%. This is an essential factor for ENL and ELF PV use. However, opaque PVs have a relatively high percentage, namely 27.81% which is the second highest percentage in the overall distribution of PVs. In addition to transparent and opaque PVs, there are two other categorizations, semi-transparent PVs in the intermediary level between transparent and opaque PVs, and non-existent PV forms which do not actually exist in ODPV (1993). These categorizations yield significant results in terms of their percentage. Semi-transparent PVs have the overall percentage of 22.81% while non-existent PV forms have the most prevalent results in percentage, namely 37.19%, which also has significance because there are some instances of the occurrences of these non-existent PV forms, proclaimed through the analyses of BNC and VOICE to designate the cross-reference comparisons between ENL and ELF speakers.

Taking the overall distribution results of the transparency categorizations of PV structures illustrated in Appendix C, Table 20 into consideration, it could be discussed that the transparency categorizations of 320 PVs point out the transparency nature of the target PVs and yield the essential outcome that the prevalence of opaque PVs, 27.81%, is found to be a little higher than twice as much as that of

transparent PVs, 12.19%, which is clear from the overall distribution results of transparency categorization. It implies that the outnumbering distribution in proportion of opaque PVs would cause ELF speakers to involve them in their actual spontaneous language performance at a low degree because of their widely-presumed inclination to utilize "declaratively-governed" (Alptekin, 2011) processes in their semantic processing, which is interestingly found vice versa and even a little higher in percentage for their use of opaque PVs (see Figure 9).

In addition, there was an intriguing formulation of the categorization functioning in the intermediary level, namely semi-transparency. Semi-transparency PV use could be explicated in terms of their possession of both transparent and opaque meaning across the lexical meanings within semantic representations of PVs in the intermediary layer. The meaning interpretations would lead to the overall perception of the semantic entities of LVs and AVPs, thereby incorporating both of them within themselves, yielding the formulation of semi-transparent PVs, the preponderance of which among 320 PVs in the ENL and ELF use could be highly dependent on the extent of the context where these PVs were utilized in the actual language productions of ENL and ELF speakers. The percentage of semi-transparent PVs in the overall distribution of transparency categorizations constitute 22.81%, which could be interpreted as nearly one-fourth of all the targeted PVs, calling for a detailed examination in their own right.

Last but not least, it can be seen from Table 20 in Appendix C that there is a high prevalence of non-existent PV forms as part of the transparency categorizations, which is displayed as nearly one-third of the whole distribution of the transparency categorizations of PVs (in fact, a little higher than one-third), namely 37.19%. Thanks to this result, a few hypothetical evaluations can be enabled in tune with the

other categorizations because of the existing instances of the non-existent PV forms in ENL and ELF use, which would eventually lead to the formation of Table 21 and Table 22 in Appendix C for the purpose of the illustration of the semantic analyses of PVs within the scope of transparency nature of the PV use.

Overall transparency analyses, evidenced in percentages through BNC and VOICE as shown in Appendix C, Table 21 and Table 22, result in significant findings about the transparency nature of PVs, taking the discussion a step further by providing cross-reference comparisons between ENL and ELF use of PVs through the corpus data results of transparency nature.

With regard to the results of the transparency nature of PVs in BNC and VOICE, Table 21 and Table 22 in Appendix C summarize the findings of the transparency analyses by illustrating the percentages of these categorizations of PV lemmas in BNC and VOICE, respectively. In the light of these results, it can be argued that ENL speakers demonstrate an interestingly similar use of transparent PVs in comparison to ELF speakers, which is compatible with the expectation of the thesis based on the literature by furthering the high utilization of transparent PVs by both speakers because ENL speakers were expected to use transparent PVs without difficulty similarly to transparent PV use by ELF speakers due to the analytical process of making sense of PVs by ELF speakers.

Regarding the percentages of the transparency categorizations of PVs in Table 21 and Table 22 in Appendix C, the transparency nature of PVs is interestingly found to be similar in not only the frequency of transparent PVs but also that of opaque PVs between ENL and ELF use. This places a particular importance to the modus operandi in the semantic processing of PVs by both speakers, bringing about the elaboration of transparency nature of PVs which are dwelled on in Figure 9.

The overall percentages of each transparency categorization of PVs can be seen in Figure 9 which illustrates the percentages of transparency categories out of 100 by visualizing the variance in the proportions of PV use by ENL speakers in comparison to ELF speakers. This enables the detailed outline of the PV use under scrutiny within their transparency nature. As Figure 9 demonstrates, the overall percentages imply that ENL and ELF speakers go through similar interpretation processes in not only transparency but also opaqueness nature of PVs possessing equal manners in their semantic representations.

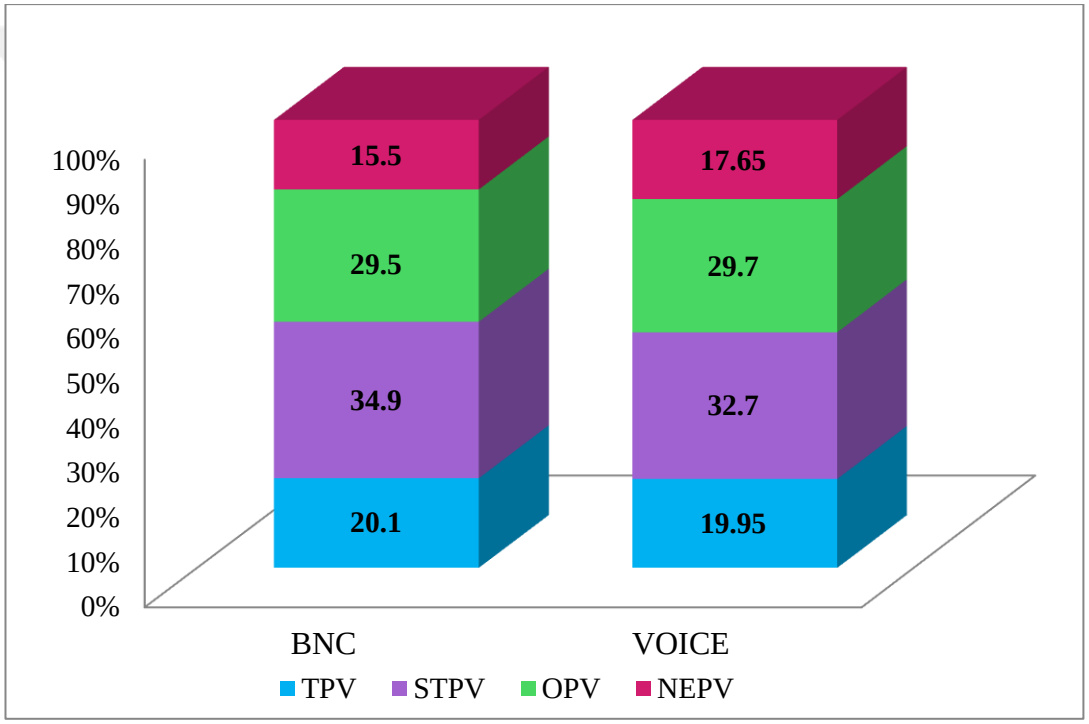

Fig. 9 Percentages of transparency categorizations in BNC and VOICE

In addition, according to what Figure 9 points out, it can be inferred that the overall semi-transparency percentage is a little higher in terms of the preponderance of the semi-transparent PV uses in the interactions of ENL speakers rather than those of ELF speakers though the relative difference remains 2.2% between the two. For the ELF counterpart, since the variance in the distributions between ENL and ELF use of opaque PVs is considerably similar, ELF speakers can be discussed to use the

same amount of opaque PVs with ENL speakers and they use even 0.2% higher than ENL use, which also rejects the third hypothesis about the language productions of ELF speakers in terms of their transparency nature. The initial hypothesis stated that the outcomes of the interactions by ELF speakers would lead to much more transparent PVs but far less opaque PVs. As a matter of fact, the opposite of what the hypothesis about the third research question set forth is uncovered by the corpus-based analyses of the spontaneous interactions of ENL and ELF speakers and this intriguing result is extrapolated for the use of PVs in their opaqueness nature by ELFers, which is also interpreted as a sign of proficient use among ELFers.

Moreover, the considerably high percentage of the non-existent PV forms (see Table 20, Appendix C) is not witnessed that much in the spontaneous language productions of both ENL and ELF speakers, only confined to low percentages in terms of the non-existent PV occurrences in BNC and VOICE, namely 15.5 for ENL use and 17.65 for ELF use, which is one of the interesting outcomes of the thesis. The current study elucidates the fact that since both ENL and ELF speakers are involved in their conversations by trying to maintain mutual intelligibility, both of the speakers do not attempt to use non-existent forms to a certain extent in spite of the existence of the non-existent PV forms in their interactions no matter what. That is because they could also creatively form new PVs to convey the intended meaning, which are labeled as non-existent since they do not exist in ODPV (1993).

As a consequence, what was found regarding the third research question was that the transparency nature of PVs was similar between the interactions of ENL and ELF speakers to a certain degree. Although there were also small variances in semi-transparent PVs and non-existent PV forms, the transparent and opaque PVs were involved to the very similar extent by the two parties. This may suggest that the

transparency characteristics of certain PVs are perceived similarly by both ENL and ELF speakers as opposed to the initial hypothesis. In fact, the transparency analysis outcomes have brought about the assumption towards a possibility of the semantic universals of opaque PVs for ELF speakers to an extent since they displayed the counterbalance behavior in their language performance in comparison to ENL speakers. This actually points out the necessity for a further analysis of each of the use of PV instances to find out the meaning variations in the corpus data. By any means, the transparency paradigm of the use of PVs can be discussed to display more or less similar constructional patterns as illustrated in Figure 9 between ENL and ELF speakers though the distributions of the percentages are not in accordance with each other, displaying some variances, which is worth investigating by incorporating other domains of interactions into analysis but without disrupting the counterbalance between the two corpora.

Last but not least, within the scope of fourth research question, whether there were similarities or differences in the use of PVs and what kinds of similarities and differences between the constructional patterns of PV use by ENL and ELF speakers existed were questioned through BNC and VOICE. Regarding the fourth research question, it was hypothesized that PVs used by ENL speakers would be different from those used by ELF speakers since both of them were expected to go through different processes in their use of PVs in both syntactic and semantic terms, leading to different outcomes in the frequency, decomposability and transparency nature.

In terms of the similarities and differences of the use of PVs, the order of lexical verbs based on their frequency of occurrences through BNC and VOICE as shown in Appendix D, Table 23, the order of PV lemmas[13] based on their frequency

[13] For the explications of the most and least frequent PV lemmas, see section 6.1.

of ENL and ELF use (see Table 8 and Table 9), and general outlines underlined through the corpus-based analyses to designate frequency, decomposability and transparency characteristics could be dwelled upon in line with the last research question.

In Appendix D, Table 23, the comparison of the frequency of LVs attracts attention as an influential factor on the frequency of ENL and ELF PV use, depicting a close resemblance in the order of the first five LVs as the most frequent LVs between ENL and ELF speakers, with a difference in the order of the first LVs as 'get' and 'go'.

In addition to this observable close resemblance in the most frequent five LVs between BNC and VOICE as can be seen from Table 23 in Appendix D, some other LVs in the list, apart from these most frequent five, can be discussed to display variances in the order, having only a few LVs, namely 'pick' and 'set' maintaining this resemblance in the order between BNC and VOICE, and a few more LVs, namely 'sit', 'move', 'bring', 'find' and 'give' keeping their resemblance to a certain extent by displaying only one order difference between BNC and VOICE, which can be perceived in cross-reference comparison. Moreover, there are the other LVs consisting of the rest of all the LVs in the LV list, in a completely different order between BNC and VOICE. These LVs showing so much variance in their order (more than two levels in BNC and VOICE) are also interesting to underline because one LV which is at the end of the list in one corpus as the least frequent LV is the one at the higher frequency of use in the other corpus as shown in Appendix D, Table 23. The case of the LVs 'point' and 'carry' can be given as an example. LV 'point' is the last verb in the list having the least amount of use in BNC, but the same LV is in the seventh rank in VOICE. Likewise, LV 'carry' is the last verb having the least

amount of use in VOICE, but this same LV is in the fourteenth rank. Overall, the results of the orders of LVs and their frequency can be discussed to display more similarities than differences in terms of their prevalence at higher rates between the interactions of ENL and ELF speakers, only with some variance in their order of LVs among the least frequent ones, which leads to a necessity for further investigations in the use of LVs between two corpora.

Regarding the order of PV lemmas based on their frequency of ENL and ELF use, PV lemmas can be discussed to disclose variances not only in their order but also in their frequency results between BNC and VOICE except for the similar order only for the first two PV lemmas as for PV lemma of ‘go’ in the first rank and PV lemma of ‘come’ in the second rank, which can be pondered taking the results in Table 8 and Table 9 into consideration. This implies that different PV structures are being utilized in the authentic interactions of ENL and ELF speakers for varied purposes to convey their intended meanings to their interlocutors, also thrusting a necessity to examine the contextual purposes of the ENL and ELF speakers and their actual utterances in pragmatic terms for further analysis to the forefront.

Finally, in tune with the last research question, the general outlines of the results of the frequency, decomposability, and transparency characteristics of the PV use accumulated through corpus-based analyses in BNC and VOICE can be expounded to display far more similarities rather than differences between the interactions of ENL and ELF speakers as opposed to the envisaged hypothesis at the initial phase of the thesis. This outcome can be also acknowledged to be significant in a way to lead to the essential assertion that ENL and ELF speakers are discussed to go through highly similar processes in their PV use in terms of their modus operandi in their linguistic systems, which could be functionally and socio-cognitive

linguistically encoded and influenced by similar representation systems in their language use of linguistic structures despite not genetically (cf. functional and genetic nativeness by Kachru [2005] on the issue). This distinguishing outcome has already been reflected onto the use of PVs by both ENL and ELF speakers evidenced as a result of their unconscious real life language use in the two corpora in a meaningful contextual environment, particularly without being exposed to any confounding factors. Therefore, the linguistic outcomes of ENL and ELF speakers are found to demonstrate utmost close resemblance in the PV use within the scope of decomposability and transparency paradigms whereas the frequency analyses of the PV use are found to display some differences to an extent, in which the relative frequency of the PV use by ENL speakers are found to outnumber twice as many as those of the PV use by ELF speakers[14] in terms of percentages, because of the presumably potential reservations of ELF users, whose linguistic system might be well equipped with ENL encoding at the same proficiency level with ENL users in the authentic language production, but their linguistic performances could be withdrawn and underestimated due to the common mentality of avoiding variations seen as a type of mistake for ELFers, as a convergence strategy, and to a degree due to the assumption for the predominance of ENL speakers for the standardization criterion in the language system.

[14] The interactions of ELF speakers yield higher in the percentage of ELF PV use, having 0.32% in VOICE, in comparison to the findings of ENL PV use in BNC by Gardner and Davies (2007). However, there are two main differences. Gardner and Davies (2007) analyzed both written and spoken BNC, not just ED and LE domains of the spoken BNC and accepted VPCs under PVs.

CHAPTER 8

CONCLUSION

The current study aimed to examine the use of PV structures within the scope of their frequency, decomposability and transparency nature through the spontaneous authentic language performance of ENL and ELF speakers. In order to demonstrate actual language productions of English among native and nonnative speakers of English, the spoken domains of two representative corpora were included in the corpus data. It was an exploratory corpus-based study and tried to analyze ENL and ELF use of PVs through BNC and VOICE. 320 PV structures were determined for the quantitative analyses of the current study. It mainly focused on the frequency, decomposability and transparency nature of these 320 PVs through the ED and LE domains of BNC and VOICE. In line with the corpus analyses of the target PVs, the following findings were obtained.

First, related to the frequency nature of PV use in BNC and VOICE, it was found that ENL speakers used PVs in their actual conversations twice as many as ELF speakers. The frequency results indicated that ELF speakers might be displaying an avoidance behavior since ELF speakers in VOICE are perceived as "experienced ELF speakers" (Seidlhofer et al., 2013). Thus, ELF speakers can be regarded to show a tendency to display avoidance behavior because they are proficient in their ELF interactions, in which they actually know that they might use PVs, but instead, ELF speakers may be considered to choose not to use these target items as a convergence strategy in the stream of their conversations. Therefore, the frequency distributions of the use of PV structures among ENL speakers outnumbered those among ELF speakers. It was assumed that since ELF speakers tried to maintain common ground

for an effective interaction, they might show avoidance behavior in their PV use possibly to prevent any misunderstandings in their interactions. It was concluded that PV structures needed to be incorporated into both ENL and ELF stream of conversations without hesitance to a higher degree and ELF speakers should take on more active role in their use of PVs displaying language creativities.

As for the decomposability nature of PVs, the current study resulted in a significant resemblance in the decomposable and nondecomposable use of PVs between ENL and ELF speakers. The proportion of non-decomposable PV percentages was strikingly the same for both speakers, which was an unexpected significant finding of the study. The decomposability nature of the use of PVs by both ENL and ELF speakers interestingly pointed to the existence of similar outcomes, evidenced through corpora analyses on a contrastive basis between BNC and VOICE. That is to say, the decomposable and non-decomposable characteristics of ENL and ELF PV uses were significantly found to be similar. Therefore, the thesis concludes that ENL speakers went through similar syntactic processing mechanisms when compared to ELF speakers in their PV use, which was evidenced by the data analysis results.

With regard to the corpora analyses results, the study formulated a new alternative approach and proposed Nondecomposable Lexical Annotation (NLA). In line with NLA, PVs are conceptualized to be processed as holistic units and used as nondecomposable though they were even decomposable in nature. NLA has further proposed that PVs display lexical chunks having annotations in the mental lexicon of speakers of English, so they cause opaque meaning interpretations based on these lexical annotations in the initial phase. In other words, in terms of what NLA has proposed, when PVs are processed and retrieved from the mental lexicon,

nondecomposable positions are used in syntactic terms and lexical annotations comprising opaque semantic representations are activated and produced in semantic terms, which is of vital importance through the corpora analyses results of the thesis. Thus, it was concluded that PVs were used based on NLA, also supported by evidential findings of the study.

Regarding the transparency nature, it was found that semi-transparent and opaque PV structures outnumbered transparent PVs. In line with this finding, it was concluded that NLA was involved in the spontaneous use of PVs in terms of their transparency characteristics, interestingly similar between ENL and ELF speakers. Therefore, the transparency nature of ENL and ELF PV use could be discussed to display similar characteristics in the interactions of both ENL and ELF speakers, drawing the inference that ENL and ELF speakers can go through similar semantic processing mechanisms in their modus operandi of PV use.

In conclusion, the findings of the current study paved the way for a reconceptualization of ENL and ELF language use indicated by the usage patterns of PV structures, which was evidenced by the corpora analyses on a contrastive basis within the scope of frequency, decomposability and transparency by providing an explanatory adequacy to the use of PVs. As opposed to what was indicated by other studies in the literature, this study has found an essential resemblance between the PV uses of ENL and ELF speakers by syntactic and semantic mechanisms. The thesis has resulted in the proposal of NLA.

8.1 Pedagogical implications

Within the scope of the corpus analyses conducted through ENL and ELF interactions with a cross-reference check, important findings were revealed about the use of PVs in terms of their frequency, decomposability and transparency. These findings can prevail in the tendency of NLA in the interactions of both ENL and ELF speakers, which necessitates raising awareness about the decomposability and transparency nature of PVs for both ENL and ELF speakers.

A higher prominence should be given to PV structures in English language learning and teaching. English course books can be intended to focus more elaborately on the PV use and their decomposable and transparent structures as well as non-decomposable and opaque structures, besides PVs in the intermediary position (i.e. semi-transparent PVs and both decomposable and nondecomposable PVs). Language course books can give a full play to the illustrative instances of PVs within the context to make the use of PVs permanent. Course book designers should take the decomposability and transparency nature of PVs into considerations and should give illustrative examples on the use of PV structures.

To be more specific, teaching material designers could dwell on the decomposable, nondecomposable and both decomposable and nondecomposable PV distinctions as well as transparent, semi-transparent and opaque PVs through authentic listening materials involving these structures and integrate the specific use of decomposable, nondecomposable, transparent and opaque PV uses and their representative differences into reading texts in the language course books. First, the inductive teaching of decomposability nature of PV items should be given prominence by exposing the learners of English within the classroom environment to the target items. Then, this should be supported by communicative language teaching

approach in a meaning-oriented classroom (Spada, 2007). L2 instruction activities should be provided to the learners of English in a meaningful context accordingly. Thus, the main aim within the actual classroom practice should be to direct L2 learners of English to meaningfully use the PV structures. In terms of the decomposability and transparency characteristics of PVs, the classroom activities can involve real life language production activities such as role-plays, information gap activities, and mock interviews, through which L2 speakers should be supported and moderated to use decomposable, nondecomposable, transparent, opaque, as well as both decomposable and nondecomposable, and semi-transparent PVs. The learners in the classroom should not hesitate to yield mistakes or produce ungrammatical versions of target items in their authentic interactions by the guidance of their instructors from an ELF perspective.

Teachers can also pay a close attention to the spontaneous use of PVs in students' language productions and should act as a moderator in the correct use of these structures. They should be ‘more tolerant to linguistic creativities’ of the language learners (Ersin, Abaylı, & Bayyurt, 2012), thereby adopting an alternative approach, NLA, to the language performance of the students.

Language instructors and in-service teachers can also effectively use corpora materials as an awareness-raising activity for L2 learners since the use of corpora display "samples of language through" concordances (Sinclair, 2004) by providing L2 learners with an opportunity to take in the intended usage patterns of PVs.

Moreover, the use of PVs should be highlighted not only in ELF contexts but also interestingly in ENL contexts as regards the findings of the study. Since the use of PVs denotes linguistic creativity, ENL speakers should also use PV structures actively in their language productions. The similar syntactic and semantic processing

mechanisms of ENL and ELF speakers imply the necessity to give PVs more credit in English encounters. This brings the significance of exposure to PVs in both ENL and ELF interactions at full length as a comprehensive pedagogical implication of the study.

The essential findings of the study in terms of frequency, decomposability and transparency characteristics of similar ENL and ELF use have encouraged me to actively involve PV distinctions into my English language teaching. I can propose that PVs should be integrated into authentic language materials within the classroom practice through meaningful language exchanges, incorporating into dramas, information gap activities and even games such as taboo, by exposing my students to the examples of real life PV uses. It is not meant any deductive language instruction should be avoided. On the contrary, what I can suggest in line with the current study is that first meaningful input should be provided for students to figure out the usage patterns of PVs in terms of decomposability and transparency characteristics, and then, language instructors should explain the internal mechanisms of PVs, by furthering decomposability and transparency of PVs. Language instructors can point out that both decomposable and nondecomposable and semi-transparent PVs as the intermediary position can be used interchangeably, and emphasize that language users can freely use these structures in either positions with either intended meaning.

8.2 Limitations

The current study has explored the decomposability and transparency nature of PV uses in ENL and ELF context through counterbalanced corpora designs and reached important conclusions that both parties have gone through similar processing systems

in their use of PVs. However, a few limitations should be taken into consideration with respect to the corpus design and analysis of the study.

Firstly, the ED and LE domains were chosen for the analyses in both BNC and VOICE to attain counterbalancing. However, if all the domains in the spoken component of BNC and VOICE corpora had been analyzed, a greater insight might have been provided in terms of decomposability and transparency characteristics of PV structures. Secondly, 320 PV structures were determined and they were analyzed in both ENL and ELF corpora to provide a contrastive perspective. However, the number of PVs could be a limitation in the study. Many more PVs could have led to a more reliable and generalizable insight.

8.3 For further studies

This study has shed light on the use of PV structures within the scope of frequency, decomposability and transparency in ENL and ELF interactions. It can be suggested that a more comprehensive corpus design can be conducted for further studies. A further step can be to analyze professional domains, namely business, organizational, research and science of VOICE and business, public and instructional domains of BNC to increase the extension of the ELF and ENL interactions so that a wider array of PV instances could be analyzed in terms of the frequency, decomposability and transparency. Although the current study has analyzed the most frequent PVs formulated with the juxtaposition of 20 most frequent lexical verbs and 16 adverbial particles, which is adopted from Gardner and Davies (2007), further studies may increase their number of PV structures to be analyzed and more PVs can bring in more elaborate perspective on the issue.

The current study has analyzed the frequency of PV structures, indicating that ELF speakers might display an avoidance behavior because they are more conscious about their language productions by trying to avoid the target PV items, to reach a common ground free of mistakes, as a convergence strategy to the ENL counterpart in line with the corpora analyses of the study through BNC and VOICE. For future research, avoidance of the frequency of PVs within the ELF framework should be further investigated in terms of holistic, contextualized or longitudinal research designs to understand the avoidance issue of ELF speakers in their authentic use of PVs at full length.

The study conducted analyses in BNC representing ENL speakers and VOICE representing ELF speakers. As a further recommendation, another representative ENL corpus can be incorporated in the corpus design and more valid outcomes may be provided. Corpus of Contemporary American English (COCA) can be investigated as the next step. In this way, further studies can also examine to what extent similar findings can be revealed in the decomposability and transparency use of PVs between the ENL corpora of British and American English and the ELF corpus. It is hoped that valuable insights can be gained with more comprehensive corpora and a greater number of PVs for further studies.

APPENDIX A

EXETEST ILLUSTRATIONS

Exetest 2

drill ... in – flocked in – leaking in – moved in – packed in – pop in – pulled in – rushed in – slip in/sneak in – smuggled in – stormed in – turn in

1	The rain was ... ... through a crack in the roof.	l .....
2	As the lock of the front door is broken, we will have to ... our way	dr .....
3	Every day lots of tourists ... ... to see the exhibition.	fl .....
4	The youngsters hoped to ... .../... ... without anyone noticing.	sl ..... /sn .....
5	Quite a lot of drugs are ... ... through our ports.	sm .....
6	Madame Butterfly always ... ... a large audience.	p .....
7	I'll go and see the new neighbours who have just	m .....
8	Do you mind if I ... ... for tea tomorrow?	p .....
9	It's already midnight, it's high time to	t .....
10	The new film ... ... huge crowds for weeks on end.	p .....
11	She ... ... with the good news that she had passed her exams.	r .....
12	With fire in his eyes, the leader ... ... and began to yell out orders.	st .....

***Glosses:* 1** *leaking in: liquid, gas entering steadily in small quantities; crack: a narrow space or opening* **– 2** *drill ... in(to): force in one's way with a drill or machine by destroying the lock* **– 3** *came in in big groups, great numbers (a flock of sheep, goats)* **– 5** *are brought in illegally* **– 6** *attracts large audiences* **– 8** *visit you briefly* **– 9** *(informal) turn in: go to bed* **– 10** *packed in: attracted; for weeks on end: continuously for many weeks*

Fig. 10 An exetest illustration of the adverbial particle 'in'

(Rudzka-Ostyn, 2003, p. 50)

Exetest 4: Replace the italicised expressions by phrasal verbs.

beef up – bolster up (US: bolster) – brought up – build up – catch up with – fix ... up – fouled up – heating up – keep ... up – loosen ... up – marked up – perk up – polish up – speed up – stir up – stood up for – tighten up on

1	I really need to ... ... my English style.	*improve (the style)*	p .....
2	They are wondering how to ... ... production.	*make faster, raise*	sp.....
3	Our leader gave a firy speech to ... ... our morale.	*strengthen*	b .....
4	Work hard and you will soon ... ... the best students.	*know as much as*	c.....
5	She has ... ... her children entirely on her own.	*raised/reared*	br .....
6	Would it be possible to ... my car ... by lunch time?	*repair*	f.....
7	This diet will allow you to ... ... your strength in no time.	*develop*	b .....
8	She is still very weak but it seems she will soon	*regain vigour*	p .....
9	The police had to ... ... their border patrol presence.	*increase*	b .....
10	The butcher ... ... his prices by 25% and sold more.	*raised*	m .....
11	The new law is likely to ... ... racial hatred.	*cause/lead to*	st.....
12	The leader boldly ... ... the rights of minority groups.	*defended*	st.....
13	He was tense and thought that a drink would ... him	*relax*	l .....

Fig. 11 An exetest illustration of the adverbial particle ‘up’

(Rudzka-Ostyn, 2003, p. 82)

Exetest 1

are away – called away – clean away – cut ... away – is ... away – shied away – snatched ... away – stay away from – swept away – turned away from – wiped away

1	When I was young, I was not a party goer, I ... ... from large crowds.	sh.....
2	I was ... ... from the meeting by an important telephone call.	c.....
3	The meat will taste better if you don't ... the fat .../*off* before roasting it.	c.....
4	The university ... only a quarter of an hour ... from here.	i .....
5	Our neighbours ... ... on holiday for the whole month.	a.....
6	A good piece of advice: ... ... that area at night.	st.....
7	Hundreds of fans had to be ... ... the stadium.	t .....
8	He ... ... his tears and tried to regain his composure.	w.....
9	Please ... ... the plates and bring us some coffee.	cl.....
10	A youngster on a motorcycle ... my sister's bag	sn.....
11	The hurricane ... ... many cars and caravans.	sw.....

***Glosses: 1** I did not go to parties because of shyness or of lack of confidence – **3** roast: cook (meat) in an oven or over a fire – **7** were refused entry – **8** regain his composure: get back or find a normal, calm state of mind – **9** remove, take away – **10** took away, grabbed suddenly and with violence – **11** hurricane: storm; moved, displaced (vehicles) with violence*

Fig. 12 An exetest illustration of the adverbial particle 'away'

(Rudzka-Ostyn, 2003, p. 140)

APPENDIX B

SYNTACTIC ANALYSES

Table 14. Frequency of Decomposable and Nondecomposable PVs in BNC

20 LVs x 16 AVPs	DPVs	NDPVs	Total PV Results	Unrelated False Positives
LV 'GO' x 16 AVPs	214	4,622	4,836	565
LV 'COME' x 16 AVPs	100	3,689	3,789	223
LV 'TAKE' x 16 AVPs	659	616	1,275	101
LV 'GET' x 16 AVPs	998	1,219	2,217	786
LV 'SET' x 16 AVPs	45	321	366	15
LV 'CARRY' x 16 AVPs	69	401	470	11
LV 'TURN' x 16 AVPs	140	392	532	35
LV 'BRING' x 16 AVPs	325	437	762	12
LV 'LOOK' x 16 AVPs	57	274	331	81
LV 'PUT' x 16 AVPs	845	428	1,273	43
LV 'PICK' x 16 AVPs	173	374	547	6
LV 'MAKE' x 16 AVPs	161	247	408	113
LV 'POINT' x 16 AVPs	10	89	99	2
LV 'SIT' x 16 AVPs	5	181	186	16
LV 'FIND' x 16 AVPs	30	414	444	66
LV 'GIVE' x 16 AVPs	98	199	297	61
LV 'WORK' x 16 AVPs	198	567	765	71
LV 'BREAK' x 16 AVPs	39	196	235	19
LV 'HOLD' x 16 AVPs	58	125	183	16
LV 'MOVE' x 16 AVPs	42	362	404	28
TOTAL	4,266	15,153	19,419	2,270

Note: The results indicate the frequency of decomposable and nondecomposable PVs in BNC. DPVs stand for decomposable PVs and NDPVs for nondecomposable PVs.

Table 15. Frequency of Decomposable and Nondecomposable PVs in VOICE

20 LVs x 16 AVPs	DPVs	NDPVs	Total PV Results	Unrelated False Positives
LV ‘GO’ x 16 AVPs	16	269	285	115
LV ‘COME’ x 16 AVPs	5	190	195	49
LV ‘TAKE’ x 16 AVPs	14	21	35	33
LV ‘GET’ x 16 AVPs	23	82	105	67
LV ‘SET’ x 16 AVPs	1	11	12	3
LV ‘CARRY’ x 16 AVPs	1	2	3	1
LV ‘TURN’ x 16 AVPs	3	14	17	6
LV ‘BRING’ x 16 AVPs	7	12	19	4
LV ‘LOOK’ x 16 AVPs	7	16	23	17
LV ‘PUT’ x 16 AVPs	96	50	146	25
LV ‘PICK’ x 16 AVPs	7	2	9	2
LV ‘MAKE’ x 16 AVPs	14	24	38	31
LV ‘POINT’ x 16 AVPs	4	13	17	46
LV ‘SIT’ x 16 AVPs	2	25	27	28
LV ‘FIND’ x 16 AVPs	13	31	44	22
LV ‘GIVE’ x 16 AVPs	8	14	22	19
LV ‘WORK’ x 16 AVPs	8	86	94	25
LV ‘BREAK’ x 16 AVPs	1	6	7	4
LV ‘HOLD’ x 16 AVPs	0	9	9	1
LV ‘MOVE’ x 16 AVPs	2	32	34	5
TOTAL	232	909	1,141	503

Note: The results indicate the frequency of decomposable and nondecomposable PVs in VOICE. DPVs stand for decomposable PVs. NDPVs stand for nondecomposable PVs.

Table 16. Percentages of Decomposable and Nondecomposable PVs in BNC

20 LVs x 16 AVPs	% DPVs	% NDPVs	% Total PV Results	% Unrelated False Positives
LV 'GO' x 16 AVPs	0.0066	0.1435	0.1501	0.0175
LV 'COME' x 16 AVPs	0.0031	0.1145	0.118	0.0069
LV 'TAKE' x 16 AVPs	0.0205	0.0191	0.0396	0.003
LV 'GET' x 16 AVPs	0.031	0.0378	0.0689	0.024
LV 'SET' x 16 AVPs	0.0014	0.01	0.0114	0.0005
LV 'CARRY' x 16 AVPs	0.0021	0.013	0.0146	0.0003
LV 'TURN' x 16 AVPs	0.0044	0.012	0.0165	0.001
LV 'BRING' x 16 AVPs	0.0101	0.014	0.0237	0.0004
LV 'LOOK' x 16 AVPs	0.0018	0.0086	0.0103	0.0025
LV 'PUT' x 16 AVPs	0.0262	0.0133	0.0395	0.0013
LV 'PICK' x 16 AVPs	0.0054	0.0116	0.017	0.0002
LV 'MAKE' x 16 AVPs	0.005	0.0077	0.0127	0.0035
LV 'POINT' x 16 AVPs	0.0003	0.0028	0.0031	0.00006
LV 'SIT' x 16 AVPs	0.00016	0.0056	0.0058	0.0005
LV 'FIND' x 16 AVPs	0.0009	0.0129	0.0138	0.002
LV 'GIVE' x 16 AVPs	0.003	0.0062	0.0092	0.0018
LV 'WORK' x 16 AVPs	0.0062	0.018	0.0238	0.0022
LV 'BREAK' x 16 AVPs	0.0012	0.006	0.007	0.00059
LV 'HOLD' x 16 AVPs	0.0018	0.0039	0.0057	0.0005
LV 'MOVE' x 16 AVPs	0.0013	0.011	0.0125	0.0008
TOTAL	0.133	0.47	0.603	0.071

Note: The results indicate the percentages of decomposable and nondecomposable PVs, the percentages of total PV occurrences and the percentages of unrelated items in BNC within 3,220,822 words. DPVs stand for decomposable PVs and NDPVs stand for nondecomposable PVs.

Table 17. Percentages of Decomposable and Nondecomposable PVs in VOICE

20 LVs x 16 AVPs	% DPVs	% NDPVs	% Total PV Results	% Unrelated False Positives
LV 'GO' x 16 AVPs	0.0044	0.0743	0.079	0.032
LV 'COME' x 16 AVPs	0.0014	0.0525	0.054	0.014
LV 'TAKE' x 16 AVPs	0.0039	0.0058	0.0097	0.009
LV 'GET' x 16 AVPs	0.006	0.023	0.029	0.019
LV 'SET' x 16 AVPs	0.0003	0.003	0.003	0.0008
LV 'CARRY' x 16 AVPs	0.0003	0.0006	0.0008	0.0003
LV 'TURN' x 16 AVPs	0.0008	0.004	0.005	0.0017
LV 'BRING' x 16 AVPs	0.0019	0.003	0.005	0.001
LV 'LOOK' x 16 AVPs	0.0019	0.004	0.006	0.005
LV 'PUT' x 16 AVPs	0.0265	0.014	0.04	0.007
LV 'PICK' x 16 AVPs	0.0019	0.0006	0.003	0.001
LV 'MAKE' x 16 AVPs	0.0039	0.007	0.011	0.01
LV 'POINT' x 16 AVPs	0.0011	0.0036	0.0047	0.013
LV 'SIT' x 16 AVPs	0.0006	0.007	0.0075	0.008
LV 'FIND' x 16 AVPs	0.004	0.0086	0.012	0.006
LV 'GIVE' x 16 AVPs	0.002	0.0039	0.006	0.0052
LV 'WORK' x 16 AVPs	0.002	0.024	0.026	0.007
LV 'BREAK' x 16 AVPs	0.0002	0.0017	0.002	0.0011
LV 'HOLD' x 16 AVPs	0	0.0025	0.003	0.0003
LV 'MOVE' x 16 AVPs	0.0005	0.009	0.01	0.0014
TOTAL	0.064	0.251	0.315	0.139

Note: The results indicate the percentages of decomposable and nondecomposable PVs, the percentages of total PV occurrences and the percentages of unrelated items in VOICE within 362,219 words. DPVs stand for decomposable PVs and NDPVs stand for nondecomposable PVs.

APPENDIX C

SEMANTIC ANALYSES

Table 18. Transparency Categorizations: 160 PVs with 8 More Frequent AVPs

	Out	Up	On	Back	Down	In	Off	Over
GO	STPV	TPV	STPV	TPV	TPV	TPV	OPV	OPV
COME	STPV	TPV	OPV	TPV	STPV	TPV	STPV	OPV
TAKE	STPV	STPV	OPV	STPV	TPV	STPV	STPV	OPV
GET	STPV	STPV	OPV	STPV	STPV	STPV	OPV	OPV
SET	OPV	STPV	OPV	OPV	OPV	OPV	OPV	NE
CARRY	OPV	NE	OPV	OPV	NE	NE	OPV	TPV
TURN	OPV	STPV	STPV	TPV	STPV	STPV	STPV	STPV
BRING	STPV	STPV	OPV	STPV	TPV	STPV	STPV	STPV
LOOK	OPV	STPV	STPV	STPV	TPV	OPV	NE	OPV
PUT	OPV	STPV	STPV	STPV	STPV	STPV	OPV	OPV
PICK	STPV	STPV	OPV	NE	NE	OPV	OPV	STPV
MAKE	OPV	OPV	OPV	NE	NE	NE	OPV	OPV
POINT	TPV	OPV	NE	NE	NE	NE	NE	NE
SIT	OPV	STPV	OPV	TPV	TPV	OPV	NE	NE
FIND	STPV	NE	NE	NE	NE	OPV	NE	NE
GIVE	STPV	OPV	OPV	TPV	NE	STPV	TPV	OPV
WORK	OPV	STPV	TPV	TPV	NE	STPV	STPV	OPV
BREAK	OPV	STPV	NE	OPV	OPV	TPV	TPV	OPV
HOLD	OPV	OPV	OPV	STPV	TPV	STPV	OPV	OPV
MOVE	STPV	TPV	STPV	NE	TPV	STPV	NE	OPV

Note: The results indicate the transparency nature of PVs. TPV stands for transparent PV and STPV stands for semi-transparent PV. OPV stands for opaque PV. NE stands for non-existent PV form.

Table 19. Transparency Categorizations: 160 PVs with 8 Less Frequent AVPs

	Round	About	Through	Around	Along	Under	By	Across
GO	TPV	STPV	STPV	TPV	OPV	STPV	STPV	TPV
COME	OPV	OPV	OPV	STPV	STPV	STPV	STPV	OPV
TAKE	TPV	NE	OPV	OPV	NE	OPV	OPV	TPV
GET	STPV	STPV	OPV	STPV	OPV	TPV	OPV	STPV
SET	NE	OPV	NE	NE	NE	NE	NE	NE
CARRY	NE	TPV	OPV	NE	OPV	NE	NE	NE
TURN	TPV	STPV	NE	TPV	NE	NE	NE	NE
BRING	STPV	OPV	OPV	STPV	STPV	OPV	NE	NE
LOOK	OPV	OPV	STPV	OPV	NE	NE	NE	NE
PUT	NE	OPV	OPV	NE	NE	NE	OPV	OPV
PICK	NE	NE	NE	NE	NE	NE	NE	NE
MAKE	NE	OPV	NE	NE	TPV	NE	NE	TPV
POINT	NE	NE	NE	NE	NE	NE	NE	NE
SIT	NE	STPV	OPV	STPV	NE	NE	OPV	NE
FIND	NE	NE	NE	NE	NE	NE	NE	NE
GIVE	NE	NE	NE	NE	NE	NE	NE	NE
WORK	OPV	NE	OPV	OPV	TPV	TPV	NE	TPV
BREAK	NE	NE	STPV	NE	NE	NE	NE	NE
HOLD	NE	NE	NE	NE	NE	TPV	NE	NE
MOVE	OPV	OPV	NE	OPV	STPV	NE	NE	NE

Note: The results indicate the transparency nature of PVs. TPV stands for transparent PV and STPV stands for semi-transparent PV. OPV stands for opaque PV. NE stands for non-existent PV form.

Table 20. The Overall Distribution of Transparency Categorizations of PVs

PV Lemmas 20 LVs x 16 AVPs	% TPVs	% STPVs	% OPVs	% NEPVs
LV 'GO' x 16 AVPs	43.75%	37.5%	18.75%	0%
LV 'COME' x 16 AVPs	18.75%	43.75%	37.5%	0%
LV 'TAKE' x 16 AVPs	18.75%	31.25%	37.5%	12.5%
LV 'GET' x 16 AVPs	6.25%	56.25%	37.5%	0%
LV 'SET' x 16 AVPs	0%	6.25%	43.75%	50%
LV 'CARRY' x 16 AVPs	12.5%	0%	37.5%	50%
LV 'TURN' x 16 AVPs	18.75%	43.75%	6.25%	31.25%
LV 'BRING' x 16 AVPs	6.25%	56.25%	25%	12.5%
LV 'LOOK' x 16 AVPs	6.25%	25%	37.5%	31.25%
LV 'PUT' x 16 AVPs	0%	31.25%	43.75%	25%
LV 'PICK' x 16 AVPs	0%	18.75%	18.75%	62.5%
LV 'MAKE' x 16 AVPs	12.5%	0%	37.5%	50%
LV 'POINT' x 16 AVPs	6.25%	0%	6.25%	87.5%
LV 'SIT' x 16 AVPs	12.5%	18.75%	31.25%	37.5%
LV 'FIND' x 16 AVPs	0%	6.25%	6.25%	87.5%
LV 'GIVE' x 16 AVPs	12.5%	12.5%	18.75%	56.25%
LV 'WORK' x 16 AVPs	31.25%	18.75%	31.25%	18.75%
LV 'BREAK' x 16 AVPs	12.5%	12.5%	25%	50%
LV 'HOLD' x 16 AVPs	12.5%	12.5%	31.25%	43.75%
LV 'MOVE' x 16 AVPs	12.5%	25%	25%	37.5%
TOTAL AVERAGE	12.19%	22.81%	27.81%	37.19%

Note: The results indicate the overall distributions of transparent, semi-transparent, opaque PV categorizations as well as non-existent PV forms out of 100. TPVs stand for the distribution of transparent PVs, STPVs stand for the distribution of semi-transparent PVs, OPVs stand for the distribution of opaque PVs, and NEPVs stand for the distribution of non-existent PVs.

Table 21. Percentages of Transparency Categorizations of PVs in BNC

PV Lemmas 20 LVs x 16 AVPs	% TPVs	% STPVs	% OPVs	% NEPVs	% Total PV Results
LV 'GO' x 16 AVPs	0.066	0.056	0.028	0	0.15
LV 'COME' x 16 AVPs	0.022	0.052	0.044	0	0.118
LV 'TAKE' x 16 AVPs	0.007	0.013	0.015	0.005	0.0396
LV 'GET' x 16 AVPs	0.004	0.039	0.026	0	0.069
LV 'SET' x 16 AVPs	0	0.0007	0.001	0.006	0.012
LV 'CARRY' x 16 AVPs	0.002	0	0.006	0.007	0.015
LV 'TURN' x 16 AVPs	0.003	0.0072	0.001	0.005	0.017
LV 'BRING' x 16 AVPs	0.002	0.0133	0.006	0.003	0.024
LV 'LOOK' x 16 AVPs	0.001	0.0026	0.004	0.003	0.01
LV 'PUT' x 16 AVPs	0	0.0124	0.017	0.01	0.039
LV 'PICK' x 16 AVPs	0	0.0032	0.003	0.011	0.017
LV 'MAKE' x 16 AVPs	0.0016	0	0.005	0.006	0.013
LV 'POINT' x 16 AVPs	0.0002	0	0.0002	0.003	0.003
LV 'SIT' x 16 AVPs	0.0007	0.0011	0.002	0.002	0.006
LV 'FIND' x 16 AVPs	0	0.0009	0.001	0.01	0.014
LV 'GIVE' x 16 AVPs	0.0012	0.00115	0.002	0.005	0.009
LV 'WORK' x 16 AVPs	0.007	0.0045	0.007	0.005	0.024
LV 'BREAK' x 16 AVPs	0.0009	0.0009	0.0018	0.004	0.007
LV 'HOLD' x 16 AVPs	0.0007	0.0007	0.002	0.003	0.006
LV 'MOVE' x 16 AVPs	0.0016	0.003	0.003	0.005	0.013
TOTAL	0.121	0.211	0.178	0.094	0.603

Note: The results indicate the percentages of transparent, semi-transparent, opaque PVs as well as non-existent PV forms and the percentages of total PV occurrences in BNC within 3,220,822 words. TPVs stand for transparent PVs, STPVs stand for semi-transparent PVs, OPVs stand for opaque PVs, and NEPVs stand for non-existent PV forms.

Table 22. Percentages of Transparency Categorizations of PVs in VOICE

PV Lemmas 20 LVs x 16 AVPs	% TPVs	% STPVs	% OPVs	%NE-PVs	% Total PV Results
LV 'GO' x 16 AVPs	0.034	0.0295	0.0148	0	0.079
LV 'COME' x 16 AVPs	0.01	0.0236	0.0202	0	0.054
LV 'TAKE' x 16 AVPs	0.002	0.003	0.0036	0.0012	0.0097
LV 'GET' x 16 AVPs	0.002	0.016	0.0109	0	0.029
LV 'SET' x 16 AVPs	0	0.0002	0.0015	0.0017	0.0033
LV 'CARRY' x 16 AVPs	0.0001	0	0.0003	0.0004	0.0008
LV 'TURN' x 16 AVPs	0.0009	0.002	0.0003	0.0015	0.0047
LV 'BRING' x 16 AVPs	0.0003	0.003	0.0013	0.0007	0.005
LV 'LOOK' x 16 AVPs	0.0004	0.0016	0.0024	0.002	0.006
LV 'PUT' x 16 AVPs	0	0.013	0.0176	0.01	0.04
LV 'PICK' x 16 AVPs	0	0.0005	0.0005	0.002	0.003
LV 'MAKE' x 16 AVPs	0.0013	0	0.004	0.005	0.011
LV 'POINT' x 16 AVPs	0.0003	0	0.0003	0.004	0.005
LV 'SIT' x 16 AVPs	0.0009	0.001	0.0023	0.0028	0.008
LV 'FIND' x 16 AVPs	0	0.0008	0.0008	0.011	0.012
LV 'GIVE' x 16 AVPs	0.0008	0.0008	0.0012	0.0034	0.006
LV 'WORK' x 16 AVPs	0.008	0.0049	0.0081	0.0049	0.026
LV 'BREAK' x 16 AVPs	0.0002	0.00024	0.0005	0.0009	0.0019
LV 'HOLD' x 16 AVPs	0.0003	0.0003	0.0008	0.0011	0.0025
LV 'MOVE' x 16 AVPs	0.0012	0.0024	0.0024	0.0035	0.0094
TOTAL	0.063	0.103	0.094	0.056	0.315

Note: The results indicate the percentages of transparent, semi-transparent, opaque PVs as well as non-existent PV forms and the percentages of total PV occurrences in VOICE within 362,219 words. TPVs stand for transparent PVs, STPVs stand for semi-transparent PVs, OPVs stand for opaque PVs, and NEPVs stand for non-existent PV forms.

APPENDIX D

LEXICAL VERB LIST

Table 23. Order and Frequency of Lexical Verbs in BNC and VOICE

	BNC		VOICE	
Rank order	Order of LVs	Frequency of LV Occurrences	Order of LVs	Frequency of LV Occurrences
1	get	21,846	go	1,544
2	go	18,292	get	1,082
3	come	8,956	come	672
4	make	6,398	make	620
5	take	6,246	take	518
6	look	4,911	work	422
7	put	4,533	point	364
8	give	4,332	look	345
9	work	3,947	give	339
10	find	3,065	put	324
11	bring	1,445	find	298
12	move	1,343	bring	99
13	turn	1,024	move	79
14	carry	828	sit	68
15	sit	783	turn	62
16	hold	725	break	62
17	pick	721	pick	45
18	set	714	set	43
19	break	555	hold	29
20	point	205	carry	11

Note: The results about the frequency of LVs are measured by the quantifications conducted through BNC and VOICE. These results depict the occurrences of only lexical verbs in the corpora and their order among themselves.

REFERENCES

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. Retrieved from http://www.languagevalue.uji.es/index.php/languagevalue/issue/view/18

Alptekin, C. (2011). Beyond ENL norms in ELF use: A cognitive perspective on ELF output. *Australian Review of Applied Linguistics, 34*, 148-165.

Anthony, L. (2015). AntConc (Version 3.4.4) [Computer Software]. Tokyo, Japan: Waseda University. Retrieved from http://www.laurenceanthony.net/

Azzaro, G. (2012). Phrasal verbs through data driven learning. *Journal of Theories and Research in Education, 7*(2), 1-24.

Balconi, M., & Amenta, S. (2010). From pragmatics to neuropragmatics. In M. Balconi (Ed.), *Neuropsychology of communication* (pp. 93-109). Milan, Italy: Springer.

Bhatt, R. M. (2001). World Englishes. *Annual Review of Anthropology, 30*, 527-550. doi:10.1146/annurev.anthro.30.1.527

Biber, D., & Conrad, S. (1999). Lexical bundles in conversation and academic prose. In H. Hasselgard & S. Oksefjell (Eds.), *Out of corpora: Studies in honour of Stig Johansson* (pp. 181–189). Amsterdam: Rodopi.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English.* London, England: Longman.

Björkman, B. (2011). Pragmatic strategies in English as an academic lingua franca: Ways of achieving communicative effectiveness? *Journal of Pragmatics, 43*(4), 950-964.

Blais, M., & Gonnerman, L. M. (2013). Explicit and implicit semantic processing of verb-particle constructions by French-English bilinguals. *Bilingualism: Language and Cognition, 16*(4), 829-846. doi:10.1017/S1366728912000673

Bobrow, S., & Bell, S. (1973). On catching on to idiomatic expressions. *Memory & Cognition, 1*(3), 343-346.

Bohrn, P. (2008). *English as a lingua franca: Negotiating common ground in business interaction* (MA Thesis). Retrieved from http://othes.univie.ac.at/3590 (AC07519824)

Breiteneder, A. (2005). The naturalness of English as a European lingua franca: The case of the 'third person -s'. *Vienna English Working Papers, 14*(2), 3-26. Retrieved from https://anglistik.univie.ac.at/fileadmin/user_upload/dep_anglist/weitere_Uploads/Views/Views0502ALL_new.pdf

Breiteneder, A. (2009). English as a lingua franca in Europe: an empirical perspective. *World Englishes, 28*(2), 256-269.

British National Corpus Online. (2007). *BNCweb.* (Corpus Query Processor edition). Oxford: Oxford University Computing Services on behalf of the BNC Consortium. Available from https://bncweb.lancs.ac.uk

Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language, 27*, 668-683.

Caillies, S., & Declercq, C. (2011). Kill the song - steal the show: What does distinguish predicative metaphors from decomposable idioms? *Journal of Psycholinguistic Research, 40*, 205–223. doi:10.1007/s10936-010-9165-8

Cappelle, B., Shtyrov, Y., & Pulvermüller, F. (2010). Heating up or cooling up the brain? MEG evidence that phrasal verbs are lexical units. *Brain & Language, 115*(3), 189-201. doi:10.1016/j.bandl.2010.09.004

Chen, J. (2007). On how to solve the problem of the avoidance of phrasal verbs in the Chinese context. *International Education Journal, 8*(2), 348-353.

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

Cieślicka, A. (2012). Do nonnative language speakers chew the fat and spill the beans with different brain hemispheres? Investigating idiom decomposability with the divided visual field paradigm. *Journal of Psycholinguistic Research, 42*(6), 475-503. doi:10.1007/s10936-012-9232-4

Cogo, A. (2008). English as a lingua franca: form follows function. *English Today 95, 24*(3), 58-61. doi:10.1017/S0266078408000308

Condon, N. (2008). How cognitive linguistic motivations influence the learning of phrasal verbs. In F. Boers & S. Lindstromberg (Eds.), *Cognitive Linguistic Approaches to Teaching Vocabulary and Phraseology* (pp. 133-158). Berlin: Mouton de Gruyter.

Cowie, A. P., & Mackin, R. (1993). *Oxford dictionary of phrasal verbs.* Oxford: Oxford University Press.

Crystal, D. (2003). *English as a global language* (2nd ed.). Cambridge, England: Cambridge University Press.

Cutting, J. C., & Bock, K. (1997). Thats the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends. *Memory & Cognition, 25*(1), 57–71.

Dagut, M., & Laufer, B. (1985). Avoidance of phrasal verbs - A case for contrastive analysis. *Studies in Second Language Acquisition, 7*(1), 73-79. doi:10.1017/s0272263100005167

Darwin, C. M., & Gray, L. S. (1999). Going after the phrasal verb: An alternative approach to classification. *TESOL Quarterly, 33*(1), 65-83.

Davis, D. R. (2011) Standardized English: The history of the earlier circles. In A. Kirkpatrick (Ed.) *The Routledge handbook of world Englishes* (pp. 17-36). Oxon: Routledge.

Decke-Cornill, H. (2002). We would have to invent the language we are supposed to teach: The issue of English as lingua franca in language education in Germany. *Language, Culture and Curriculum, 15*(3), 1-16.

Dewey, M. (2007). English as a lingua franca and globalization: an interconnected perspective. *International Journal of Applied Linguistics, 17*(3), 332-354.

Dirven, R. (2001). English phrasal verbs: Theory and didactic application. In M. Pütz, S. Niemeier & R. Dirven (Eds.), *Applied Cognitive Linguistics II: Language Pedagogy* (pp. 3-28). Berlin: Mouton de Gruyter.

Drew, P., & Holt, E. (1988). Complainable matters: The use of idiomatic expressions in making complaints. *Social Problems, 35*(4), 398-417.

Ellis, N. C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly, 42*(3), 375-396.

Ellis, R. (1994). Implicit/explicit knowledge and language pedagogy. *TESOL Quarterly, 28*(1), 166-172.

Ersin, P., Abaylı, Z. S. & Bayyurt, Y. (2012). English as a lingua franca in a Turkish academic context: The case of the third person "-s". In Y. Bayyurt & S. Akcan (Eds.), *Proceedings of the 5th International Conference of English as a Lingua Franca* (pp. 263-270). Istanbul: Bogaziçi University Press.

Firth, A. (1996). On the discursive accomplishment of normality: On 'lingua franca' English and conversation analysis. *Journal of Pragmatics, 26*(2), 237-259. doi:10.1016/0378-2166(96)00014-8

Fischer, I., & Keil, M. (1996). Parsing decomposable idioms. In *Proceedings of the 16th conference on Computational linguistics* (pp. 388-393). Association for Computational Linguistics.

Fraser, B. (1970). Idioms within a transformational grammar. *Foundations of language*, 22-42.

Gardner, D., & Davies, M. (2007). Pointing out frequent phrasal verbs: A corpus-based analysis. *Tesol Quarterly, 41*(2), 339-359.

Gelder, T. (1990). Compositionality: A connectionist variation on a classical theme. *Cognitive Science, 14*(3), 355-384.

Ghabanchi, Z., & Goudarzi, E. (2012). Avoidance of phrasal verbs in learner English: A study of Iranian students. *World Journal of English Language, 2*(2), 43-54.

Gibbs, R. W. (1992). What do idioms really mean? *Journal of Memory and Language, 31*(4), 485-506.

Gibbs, R. W., & Gonzales, G. P. (1985). Syntactic frozenness in processing and remembering idioms. *Cognition, 20*(3), 243-259.

Gibbs, R. W., Nayak, N. P., & Cutting, C. (1989). How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of Memory and Language, 28*(5), 576-593. doi:10.1016/0749-596X(89)90014-4

Giora, R. (2002). Literal vs. figurative language: Different or equal? *Journal of pragmatics, 34*(4), 487-506.

Giora, R., & Fein, O. (1999). On understanding familiar and less-familiar figurative language. *Journal of pragmatics, 31*(12), 1601-1618.

Glucksberg, S. (1989). Metaphors in conversation: How are they understood? Why are they used? *Metaphor and Symbol, 4*(3), 125-143.

Glucksberg, S., & Keysar, B. (1993). How metaphors work. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed.) (pp. 401-424). Cambridge: Cambridge University Press.

Gonnerman, L. M., & Hayes, C. R. (2005). The professor chewed the students…out: Effects of dependency, length, and adjacency on word order preferences in sentences with verb particle constructions. In *Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society* (pp. 785-790).

Görlach, M. (1990). *Studies in the history of the English language*. Heidelberg, Germany: Carl Winter.

Graddol, D. (1997). *The future of English?* London, England: The British Council. Retrieved from www.uniurb.it/lingue/matdid/rossi/2012-13/Lin-AZ_I_anno/The_Future_of_English_-_David_Graddol.pdf

Graddol, D. (2004). The future of language. *Science, 303*(5662), 1329-1331. doi: 10.1126/science.1096546

Graddol, D. (2006). *English next: Why global English may mean the end of 'English as a foreign language'*. London, England: The British Council. Retrieved from http://www.data.ulis.vnu.edu.vn/jspui/handle/123456789/2165

Grant, L., & Bauer, L. (2004). Criteria for re-defining Idioms: Are we Barking up the Wrong Tree? *Applied linguistics, 25*(1), 38-61.

Griva, E., & Chostelidou, D. (2011). English language teachers' conceptions and attitudes to multilingual development in Education. *Procedia - Social and Behavioral Sciences, 15*, 1780-1785. doi:10.1016/j.sbspro.2011.04.002

Hillert, D., & Swinney, D. (2001). The processing of fixed expressions during sentence comprehension. *Conceptual structure, discourse, and language, 4*, 107-121.

Hoffmann, S., Evert, S., Smith, N., Lee, D., & Prytz, Y. B. (2008). *Corpus linguistics with BNCweb.* Frankfurt am Main: Peter Lang.

House, J. (2003a). English as a lingua franca: A threat to multilingualism? *Journal of Sociolinguistics, 7*(4), 556–578.

House, J. (2003b). Teaching and learning pragmatic fluency in a foreign language: The case of English as a lingua franca. In A. M. Flor, E. U. Juan & A. F. Guerra (Eds.), *Pragmatic competence and foreign language teaching* (pp. 133-159). Castellón de la Plana: Universitat Jaume I.

Hulstijn, J. H., & Marchena, E. (1989). Avoidance. *Studies in second language acquisition, 11*(03), 241-255.

Hyland, K. (2012). Bundles in academic discourse. *Annual Review of Applied Linguistics, 32*, 150-169.

Jenkins, J. (2006). Points of view and blind spots: ELF and SLA. *International Journal of Applied Linguistics, 16*(2), 137-162. doi:10.1111/j.1473-4192.2006.00111.x

Jenkins, J. (2007). *English as a lingua franca: Attitude and identity*. New York: Oxford University Press.

Jenkins, J. (2008). *English as a lingua franca* [PowerPoint slides]. Lecture presented at JACET 47th Annual Convention, Waseda University. Retrieved from http://www.jacet.org/2008convention/JACET2008_keynote_jenkins.pdf

Jenkins, J. (2009). English as a lingua franca: Interpretations and attitudes. *World Englishes, 28*(2), 200–207.

Jenkins, J., Cogo, A., & Dewey, M. (2011). Review of development in research into English as a lingua franca. *Language Teaching, 44*(3), 281-315.

Kachru, B. B. (1992a). *The other tongue: English across cultures.* (2nd ed.). Urbana and Chicago: University of Illinois Press.

Kachru, B. B. (1992b). World Englishes: Approaches, issues and resources. *Language Teaching, 25*(1), 1-14. doi:10.1017/s0261444800006583

Kachru, B. B. (2005). *Asian Englishes: Beyond the canon.* Hong Kong: Hong Kong University Press.

Kamp, H., & Partee, B. (1995). Prototype theory and compositionality. *Cognition, 57*, 129-191.

Knapp, A. (2011). Using English as a lingua franca for (mis-)managing conflict in an international university context: An example from a course in engineering. *Journal of Pragmatics, 43*(4), 978-990. doi:10.1016/j.pragma.2010.08.008

Konopka, A., & Bock, K. (2009). Lexical or syntactic control of sentence formulation? Structural generalizations from idiom production. *Cognitive Psychology, 58*(1), 68-101. doi:10.1016/j.cogpsych.2008.05.002

Kovecses, Z., & Szabo, P. (1996). Idioms: A view from cognitive semantics. *Applied Linguistics, 17*(3), 326-355.

Kristiansen, G., & Dirven, R. (2008). Cognitive linguistics: Rationale, methods and scope. In G. Kristiansen & R. Dirven (Eds.), *Cognitive Sociolinguistics: Language Variation, Cultural Models, Social Systems* (pp. 1-20). Berlin: Mouton de Gruyter.

Kuo, I-C. V. (2006). Addressing the issue of teaching English as a lingua franca. *ELT Journal, 60*(3), 213-221. doi: 10.1093/elt/ccl001

Kurtyka, A. (2001). Teaching English phrasal verbs: a cognitive approach. In M. Pütz, S. Niemeier & R. Dirven (Eds.), *Applied Cognitive Linguistics II: Language Pedagogy* (pp. 29-54). Berlin: Mouton de Gruyter.

Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed.) (pp. 202-251). Cambridge: Cambridge University Press.

Langlotz, A. (2006). *Idiomatic creativity: a cognitive-linguistic model of idiom-representation and idiom-variation in English.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Laufer, B. (2000). Avoidance of idioms in a second language: The effect of L1-L2 degree of similarity. *Studia Linguistica, 54*(2), 186-196.

Laufer, B., & Eliasson, S. (1993). What causes avoidance in L2 learning. *Studies in second language acquisition, 15*(01), 35-48.

Lewis, M. (1993). *The lexical approach.* Hove, England: Language Teaching Publications.

Liao, Y. D., & Fukuya, Y. J. (2002). Avoidance of phrasal verbs: The case of Chinese learners of English. *Second Language Studies, 20*(2), 71-106.

Littlemore, J. (2001). Metaphoric intelligence and foreign language learning. *Humanising Language Teaching Magazine, 3*(2). Retrieved from http://byeways.net/webreadings/Internet-Research/RESEARCH/24.htm

Liu, D. (2011). The most frequently used English phrasal verbs in American and British English: A multicorpus examination. *TESOL Quarterly, 45*(4), 661-688.

Ljosland, R. (2011). English as an academic lingua franca: Language policies and multilingual practices in a Norwegian university. *Journal of Pragmatics, 43*(4), 991-1004. doi:10.1016/j.pragma.2010.08.007

Louhiala-Salminen, L., & Kankaanrata, A. (2012). Language as an issue in international internal communication: English or local language? If English, what English? *Public Relations Review, 38*(2), 262-269. doi:10.1016/j.pubrev.2011.12.021

MacArthur, F., & Littlemore, J. (2008). A discovery approach to figurative language learning with the use of corpora. In F. Boers & S. Lindstromberg (Eds.), *Cognitive Linguistic Approaches to Teaching Vocabulary and Phraseology* (pp. 159-188). Berlin: Mouton de Gruyter.

Matlock, T., & Heredia, R. R. (2002). Understanding phrasal verbs in monolinguals and bilinguals. In R.R. Heredia & J. Altarriba (Eds.), Bilingual sentence processing (pp. 251-274). North Holland, The Netherlands: Elsevier Science B.V.

Mauranen, A. (2012). *Exploring ELF: Academic English shaped by non-native speakers.* New York: Cambridge University Press.

Mauranen, A., Hynninen, N., & Ranta, E. (2010). English as an academic lingua franca: The ELFA project. *English for Specific Purposes, 29*(3), 183–190.

McArthur, T. (1987). The English languages. *English Today, 11*(3), 9-13.

McArthur, T. (1989). The long-neglected phrasal verb. *English Today, 5*(02), 38-44.

McCarthy, D., Keller, B., & Carroll, J. (2003). Detecting a continuum of compositionality in phrasal verbs. In *Proceedings of the ACL 2003 workshop on Multiword expressions: analysis, acquisition and treatment, Volume 18* (pp. 73-80). Association for Computational Linguistics.

McKay, S. L., & Bokhorst-Heng, W. D. (2008). *International English in its sociolinguistic contexts: Towards a socially sensitive EIL pedagogy.* New York: Routledge.

Mesthrie, R., & Bhatt, R. M. (2008). *World Englishes: The study of new linguistic varieties*. Cambridge, England: Cambridge University Press.

Moiron, M. B. V. (2005). Data-driven identification of fixed expressions and their modifiability. (Unpublished Doctoral Dissertation). Groningen University, Netherlands.

Molinaro, N., Canal, P., Vespignani, F., Pesciarelli, F., & Cacciari, C. (2013). Are complex function words processed as semantically empty strings? A reading time and ERP study of collocational complex prepositions. *Language and Cognitive Processes, 28*(6), 762-788. doi:10.1080/01690965.2012.665465

Morgan, P. S. (1997). Figuring out figure out. Metaphor and the semantics of the English verb-particle construction. *Cognitive Linguistics, 8*(4), 327-358.

Nassaji, H., & Tian, J. (2010). Collaborative and individual output tasks and their effects on learning English phrasal verbs. *Language Teaching Research, 14*(4), 397-419.

Nunberg, G., Sag, I., & Wasow, T. (1994). Idioms. *Language, 70*(3), 491-538.

Omoniyi, T., & Saxena, M. (2010). Introduction. In M. Saxena & T. Omoniyi (Eds.), *Contending with globalization in world Englishes* (pp. 19-40). Bristol: Multilingual Matters.

Oxford Text Archive (2015). Users Reference Guide for the British National Corpus. Retrieved from http://www.natcorp.ox.ac.uk/corpus/

Pakir, A. (2009). English as a lingua franca: Analyzing research frameworks in international English, world Englishes, and ELF. *World Englishes, 28*(2), 224-235. doi:10.1111/j.1467-971X.2009.01585.x

Pennycook, A. (2007). *Global Englishes and transcultural flows.* London, England: Routledge.

Phillipson, R. (2003). *English-Only Europe? Challenging language policy.* London: Routledge.

Pickering, L. (2006). Current research on intelligibility in English as a lingua franca. *Annual Review of Applied Linguistics, 26*, 219-233.

Pitzl, M. L. (2005). Non-understanding in English as a lingua franca: examples from a business context. *Vienna English Working Papers, 14*(2), 50-71.

Pitzl, M. L. (2009). 'We should not wake up any dogs': Idiom and metaphor in ELF. In A. Mauranen and E. Ranta (Eds.), *English as a lingua franca: Studies and findings* (pp. 299–322). Newcastle-upon-Tyne: Cambridge Scholars.

Pitzl, M. L. (2011). *Creativity in English as a lingua franca: Idiom and metaphor*. (Unpublished Doctoral Dissertation). University of Vienna, Vienna, Austria.

Pitzl, M. L. (2012). Creativity meets convention: idiom variation and remetaphorization in ELF. *Journal of English as a Lingua Franca, 1*(1), 27-55. doi:10.1515/jelf-2012-0003

Pozdnyakova, O. V., & Gunina, N. A. (2011). Teaching phrasal verbs. *Problems of Modern Science and Practice, 3*(34), 357-360.

Prodromou, L. (2007). Bumping into creative idiomaticity. *English Today, 23*, 14-25.

Prodromou, L. (2008). *English as a lingua franca: A corpus-based analysis.* London, England: Continuum.

Riehemann, S. (1997). *Idiomatic constructions in HPSG*. Paper presented at Cornell HPSG Workshop, MS, Ithaca.

Rudzka-Ostyn, B. (2003). *Word power: phrasal verbs and compounds*. Berlin: Mouton de Gruyter.

Rumelhart, D. E. (1993). Some problems with the notion of literal meanings. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed.) (pp. 71-82). Cambridge: Cambridge University Press.

Sag, I. A., Baldwin, T., Bond, F., Copestake, A., & Flickinger, D. (2002). *Multiword expressions: A pain in the neck for NLP*. Proceedings of Computational Linguistics and Intelligent Text Processing: Third International Conference: CICLing-2002, Lecture Notes in Computer Science (pp. 1-15). Springer, Berlin, Heidelberg. Retrieved from DR-NTU, Nanyang Technological University Library, Singapore: https://dr.ntu.edu.sg/handle/10220/6828

Saraceni, M. (2008). English as a lingua franca: Between form and function. *English Today, 24*(2), 20-26 doi: 10.1017/S0266078408000163

Saxena, M., & Omoniyi, T. (2010). *Contending with globalization in world Englishes*. Bristol: Multilingual Matters.

Seidlhofer, B. (2000). Mind the gap: English as a mother tongue vs English as a lingua franca. *Vienna English Working Papers, 9*(1), 51–68.

Seidlhofer, B. (2001). Closing a conceptual gap: The case for a description of English as a lingua franca. *International Journal of Applied Linguistics, 11*(2), 133-158.

Seidlhofer, B. (2002). The shape of things to come? Some basic questions about English as a lingua franca. In K. Knapp & C. Meierkord (Eds.), *Lingua franca communication* (pp. 269-302). Frankfurt/Main: Peter Lang.

Seidlhofer, B. (2004). Research perspectives on teaching English as a lingua franca. *Annual Review of Applied Linguistics, 24*, 209-239. doi:10.1017/S0267190504000145

Seidlhofer, B. (2005). English as a lingua franca. *ELT Journal, 59*(4), 339-341. doi:10.1093/elt/cci064

Seidlhofer, B. (2006). English as a Lingua Franca in the expanding circle: What it isn't. In R. Rubdy & M. Saraceni (Eds.), *English in the world: Global rules, global roles* (pp.40–50). London: Continuum.

Seidlhofer, B. (2009a). Accommodation and the idiom principle in English as a lingua franca. *Intercultural Pragmatics, 6*(2), 195-215. doi:10.1515/IPRG.2009.011

Seidlhofer, B. (2009b). Common ground and different realities: World Englishes and English as a lingua franca. *World Englishes, 28*(2), 236–245.

Seidlhofer, B., & Berns, M. (2009). Perspectives on English as a lingua franca: Introduction. *World Englishes, 28*(2), 190–191.

Seidlhofer, B., Breiteneder, A., & Pitzl, M. L. (2006). English as a lingua franca in Europe: Challenges for applied linguistics. *Annual Review of Applied Linguistics, 26*, 3–34. doi:10.1017/S026719050600002X

Seidlhofer, B., Breiteneder, A., Klimpfinger, T., Majewski, S., Osimk-Teasdale, R., Pitzl, M. L., & Radeka, M. (2013). The Vienna-Oxford International Corpus of English (VOICE) Project. Vienna, Austria.

Shin, D., & Nation, P. (2007). Beyond single words: the most frequent collocations in spoken English. *ELT Journal, 62*(4), 339-348. doi:10.1093/elt/ccm091

Sholte, J. A. (2000). *Globalization: A critical introduction*. Basingstoke/New York: Palgrave MacMillan.

Side, R. (1990). Phrasal verbs: Sorting them out. *English Language Teaching Journal, 44*(2), 144-152.

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

Sinclair, J. M. (2004). *How to use corpora in language teaching.* Amsterdam: John Benjamins Publishing Company.

Siyanova, A., & Schmitt, N. (2007). Native and nonnative use of multi-word vs. one-word verbs. *IRAL-International Review of Applied Linguistics in Language Teaching, 45*(2), 119–139. doi: 10.1515/IRAL.2007.005

Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research, 27*(2), 251–272. doi:10.1177/0267658310382068

Sjöholm, K. (1995). *The influence of crosslinguistic, semantic, and input factors on the acquisition of English phrasal verbs: A comparison between Finnish and Swedish learners at an intermediate and advanced level.* Åbo, Finland: Åbo Akademi University Press.

Smit, U. (2010). *English as a lingua franca in higher education.* Berlin: Mouton de Gruyter.

Smith, L. P. (1925). *Words and idioms.* Great Britain: The Glasgow University Press.

Spada, N. (2007). Communicative language teaching: Current status and future prospects. In J. Cummins & C. Davison (Eds.), *International handbook of English language teaching* (p. 271-288). New York: Springer.

Swinney, D., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior, 18*(5), 523-534. doi:10.1016/s0022-5371(79)90284-6

Tabossi, P., Fanari, R., & Wolf, K. (2009). Why are idioms recognized fast? *Memory & Cognition, 37*(4), 529-540. doi:10.3758/MC.37.4.529

Teubert, W. (2001). Corpus linguistics and lexicography. *International Journal of Corpus Linguistics, 6*(SI), 125-153.

Titone, D. A., & Connine, C. M. (1994). Descriptive norms for 171 idiomatic expressions: Familiarity, compositionality, predictability, and literality. *Metaphor and Symbolic Activity, 9*(4), 247-270.

Titone, D. A., & Connine, C. M. (1999). On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics, 31*(12), 1655-1674. doi:10.1016/s0378-2166(99)00008-9

Traxler, M. J. (2012). *Introduction to Psycholinguistics: Understanding Language Science.* Boston, MA: Wiley-Blackwell.

Tremblay, A., Derwing, B., Libben, G., & Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning, 61*(2), 569-613.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92*, 231-270. doi:10.1016/j.cognition.2003.10.008

Van Lancker Sidtis, D. (2010). Formulaic and novel expressions in mind and brain: Empirical studies and a dual process model of language competence. In J. Guendouzi, F. Loncke, & M. Williams (Eds.), *The handbook of psycholinguistic & cognitive processes: Perspectives in communication disorders* (pp. 247-272). London: Taylor & Francis.

Van Lancker Sidtis, D. (2012). Two-track mind: Formulaic and novel language support a dual process model. In M. Faust (Ed.), *The Handbook of the Neuropsychology of Language* (pp. 342-367). London: Blackwell Publishing Ltd.

Van Lancker, D., & Rallon, G. (2004). Tracking the incidence of formulaic expressions in everyday speech: methods for classification and verification. *Language and Communication, 24*, 207-240. doi:10.1016/j.langcom.2004.02.003

Vega-Moreno, R. E. (2003). Relevance Theory and the construction of idiom meaning. *UCL Working Papers in Linguistics, 15*, 83-104.

VOICE. (2013). *Corpus Description.* Retrieved from VOICE Project website: https://www.univie.ac.at/voice/page/corpus_description

VOICE. (2013). *Statistics VOICE 2.0 Online.* Retrieved from VOICE Project website: https://www.univie.ac.at/voice/stats/

VOICE. (2013). *The Vienna-Oxford International Corpus of English* (version 2.0 XML). Director: Barbara Seidlhofer; Researchers: Angelika Breiteneder, Theresa Klimpfinger, Stefan Majewski, Ruth Osimk-Teasdale, Marie-Luise Pitzl, Michael Radeka. Retrieved from http://voice.univie.ac.at

VOICE. (2013). *What is VOICE?* Retrieved from VOICE Project website: https://www.univie.ac.at/voice/page/what_is_voice

Wacker, E. (2011). English as a lingua franca and the third person –s (BA thesis). Retrieved from http://www.univie.ac.at/voice/page/abstracts/wacker_2011_full.pdf (a0700047)

Wood, D. (2010). *Perspectives on formulaic language: Acquisition and communication.* London: Continuum.

Wray, A. (2002). *Formulaic language and the lexicon.* Cambridge: Cambridge University Press.

Wray, A. (2012). What do we (think we) know about formulaic language? An evaluation of the current state of play. *Annual Review of Applied Linguistics, 32*(1), 231-254. doi:10.1017/S026719051200013X

Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: an integrated model. *Language & Communication, 20*(1), 1-28. doi:10.1016/S0271-5309(99)00015-4

Yano, Y. (2001). World Englishes in 2000 and beyond. *World Englishes, 20*(2), 119-132. doi: 10.1111/1467-971X.00204