

Running head: ASSESSING COMPREHENSION OF ENGLISH METAPHOR IN
ENGINEERING TEXTBOOKS

Atoms misbehaving: Assessing university students' comprehension
of English metaphor in engineering textbooks

by

Tina Beynen

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial
fulfillment of the requirements for the degree of

Master of Arts

in

Applied Linguistics and Discourse Studies

Carleton University

Ottawa, Ontario

© 2015

Tina Beynen

Abstract

This two-phase mixed methods sequential exploratory study adds to the scant literature on second language (L2) learners' metaphoric competence in the post-secondary context by investigating: 1) the prevalence of English metaphor in engineering textbooks; and 2) first-year students' metaphoric comprehension. The Phase 1 corpus analysis uncovered personification and family and relationship metaphors across all five engineering sub-disciplines, but greater concentrations in chemistry, less in engineering mechanics. In Phase 2, the responses to a metaphor comprehension task, administered to 42 newly admitted engineering students at a Canadian university as part of an overall diagnostic assessment, were rated on a 5-point scale. A bi-modal distribution differentiated high achievers (primarily English-speaking students) from low achievers (primarily foreign language students). The piloted scale and metaphor test were analyzed using item discrimination, Spearman's rho, simple linear regression and Bachman and Palmer's (2010) usefulness criteria. Implications are discussed for language teaching pedagogy and post-entry diagnostic assessment.

Acknowledgements

There have been a great number of people in my hometown of London and second home of Ottawa who have helped, encouraged and supported me in countless ways throughout this journey. I have been so incredibly fortunate to work under the guidance of my supervisor, Dr. Janna Fox. What a privilege it has been to learn from a world-renowned expert in the field of language testing and assessment! Besides being an incredibly mentor, Janna is also one of the kindest, most compassionate people I have ever met. She has been a trooper in guiding my thesis via email and Skype, and I will be forever indebted to her for not only keeping me (mostly) sane throughout the writing and editing process, but for going above and beyond in helping me get through the editing stage in record time. I am grateful to Janna for helping me to produce “the little test that could,” and a piece of research that I recognize as novice, but that I am incredibly proud of.

I would like to thank my two thesis examiners, Drs. Guillame Gentil (Carleton) and Scott Douglas (UBC, Okanagan) for their insightful comments and suggestions. I would also like to thank Drs. David Wood and Eva Kartchava for their wisdom and support, as well as their guidance through the courses and practicum that will now allow me to apply for TESL Ontario certification. And, I must thank Joan and Connie in the office for what seems like a million little (but not insignificant) things.

I have been fortunate back in London to work in an incredibly supportive environment, where every Dean, Associate Dean, and Graduate Chair I have worked with, as well as numerous staff, faculty and students have offered advice and

encouragement every step of the way. I was fortunate to meet two special women who have been mentoring me for years: Drs. Shelley Taylor and Tania Granadillo. I am also indebted to the two incredibly kind supervisors I worked with during this time: Linda Kulak and Karen Kueneman. They not only supported and encouraged me, but picked up the slack without a word of complaint while I was out of the office. I also must thank Dr. Kathryn Noel for her kind help and advice. I think our Bob would be pleased at the friendship we have developed.

Last but not least, I must thank my family and friends both in Ottawa and London. From Ottawa, I will fondly remember chats and homework sessions in the lounge, music and lively dinners, skating on the canal, biking, touring around to see tulips, and pints at Mike's – especially after a few tough stats classes! In addition, I am grateful to Brittany and Paul (honorary ALDSers) who helped me move a month into my program, and who have graciously hosted me, as did my skating partner Steph, during trips back to Ottawa. My friends in London have helped me move more times than I care to count, treated me to dinner when I was a “starving student,” and given me rides to and from the bus station. They've encouraged and cheered me on, and helped me through the rough times when I was homesick, or struggling to balance working full time while writing my thesis. I especially have to thank my angels, Joanna and Shirley, for more than I could ever express. And finally, I have to thank my sister Tammy and mom, Lynne, who are my biggest cheerleaders, and have more faith in me than I have in myself. I will be forever grateful for everything they did to help me achieve this dream.

Table of Contents

Abstract.....	ii
Acknowledgements	iii
Table of Contents	v
List of Tables	viii
List of Figures.....	ix
List of Appendices.....	x
Chapter 1: Introduction	1
<i>Researcher Background.....</i>	<i>3</i>
<i>Research Questions.....</i>	<i>4</i>
<i>Thesis Overview.....</i>	<i>4</i>
Chapter 2: Literature Review.....	7
<i>What is Metaphor?.....</i>	<i>7</i>
<i>Why Research Metaphor Comprehension?</i>	<i>10</i>
<i>Issues Affecting Reading Comprehension.....</i>	<i>14</i>
<i>The Significance of Reading for EAL Students.....</i>	<i>16</i>
<i>Metaphor Studies in the Domain of Reading.....</i>	<i>21</i>

<i>Metaphor Studies in Other Language Skill Domains</i>	22
<i>Post-entry Diagnostic Assessment and Self-Assessment for Engineering</i>	24
<i>Chapter Summary</i>	27
Chapter 3: Theoretical Framework	29
<i>Conceptual Metaphor Theory</i>	29
<i>Critiques of Conceptual Metaphor Theory</i>	34
<i>Why Conceptual Metaphor Theory?</i>	36
<i>Chapter Summary</i>	41
Chapter 4: Method	42
<i>Method overview</i>	42
<i>Participants</i>	44
<i>Instruments</i>	46
<i>Procedures</i>	48
Phase 1	48
Phase 2	49
<i>Analysis</i>	56
Analytic Approach – Phase 1	56
Phase 1 – The Corpus Analysis	58
Phase 2	62
<i>Chapter Summary</i>	65

Chapter 5: Findings and Discussion.....	67
<i>Phase 1</i>	67
<i>Phase 2</i>	70
<i>The Results Considered Through the Lens of Conceptual Metaphor Theory</i>	77
<i>The Metaphor Task as a Diagnostic Tool</i>	78
Bachman and Palmer’s Usefulness Criteria.....	82
<i>Chapter Summary</i>	86
Chapter 6: Conclusion.....	88
Thesis Summary.....	88
Implications of the Study	89
Limitations of the Study.....	93
Directions for Future Research	96
References.....	100
Appendix A.....	114
Appendix B.....	115

List of Tables

TABLE 1 *INTERPRETING LITERAL AND FIGURATIVE LANGUAGE, AND LINGUISTIC AND
CONCEPTUAL METAPHOR*..... 9

TABLE 2 *FOREIGN FIRST LANGUAGES SPOKEN BY ENGINEERING COHORT*..... 45

TABLE 3 *WORDS AND PHRASES TESTED IN THE METAPHOR COMPREHENSION TASK*..... 47

TABLE 4 *SELECT COH-METRIX INDICES FOR THE METAPHOR TEST PARAGRAPHS* 51

TABLE 5 *RATING SCALE FOR METAPHOR TEST TASK* 54

TABLE 6 *MOST FREQUENT WORD TYPES FOUND IN FIRST YEAR ENGINEERING TEXTBOOKS*. 59

TABLE 7 *FINAL COLLOCATES SAMPLED FOR METAPHORIC CONTENT* 61

TABLE 8 *EXAMPLES OF METAPHORIC CONTENT IN THE ENGINEERING CORPUS* 68

TABLE 9 *STUDENT RESPONSES AND RATINGS* 72

TABLE 10 *ITEM FACILITY AND ITEM DISCRIMINATION FOR THE 10 TEST WORDS AND PHRASES*
..... 73

TABLE 11 *SIMPLE LINEAR REGRESSION: METAPHOR PREDICTING VARIANCE ON OTHER SAFE
TASKS*..... 81

List of Figures

<i>FIGURE 1</i> SEQUENTIAL EXPLORATORY MIXED METHODS DESIGN.....	43
<i>FIGURE 2</i> FOREIGN LANGUAGES SPOKEN AND ENGLISH-MEDIUM HIGH SCHOOL EDUCATION.....	46
<i>FIGURE 3</i> STUDENT RESPONSES BY TEST ITEM.	63
<i>FIGURE 4</i> DISTRIBUTION OF STUDENT METAPHOR TEST SCORES.....	71
<i>FIGURE 5</i> METAPHOR TEST SCORE DISTRIBUTION BY L1.....	75
<i>FIGURE 6</i> BOX PLOT OF METAPHOR TEST SCORES BY L1.	75
<i>FIGURE 7</i> COMPARISON OF METAPHOR TEST RESULTS WITH SAFE RESULTS.	79

List of Appendices

METAPHOR COMPREHENSION TASK.....	114
STOP WORD LIST	115

Chapter 1: Introduction

Internationalization may be a global trend but the majority of international students attend English speaking universities (Altbach, 2004). Data from the Institute of International Education (IIE) (2013b) tell us that the top countries of origin for international students are China, India, South Korea and Saudi Arabia, and the most popular fields of study tend to be business and engineering (IIE, 2013a). Current estimates are that the first language of up to half of the students studying in Canadian institutions is something other than one of the country's two official languages, English and French (Fox, 2015). The dramatic increase in international students and overall diversity in Canadian classrooms has resulted in the need to better understand the academic preparedness of this group of students. This is the broad context within which this study takes place.

The small amount of existing literature suggests that metaphor comprehension for those learning English as an additional language (EAL) is difficult; whether it affects students' academic preparedness is further explored in this investigation. Although the concept of metaphor is not unique to the English language, metaphoric words and phrases are ubiquitous in English. A typical native English speaker invents around 3000 metaphors over the course of a week (Pollio, Barlow, Fine & Pollio, 1977, as cited in Danesi, 1995). A corpus study by Steen, Dorst, Herrmann, Kaal and Krennmayr (2010) that analyzed a sampling of lexical units¹ from the British National Corpus also provides insight into the prevalence of metaphor in the English language. Steen et al. found

¹ A term used by the Pragglejazz Group, 2007, to mean mostly words, but also polywords, idioms, phrasal verbs, and compounds, as cited in Steen et al., 2010.

metaphoric content in every seven to eight lexical units. These instances of metaphor were largely non-signalled, or not prefaced by ‘like’ or ‘as,’ which signal simile. A genre comparison in Steen et al.’s study revealed that amongst the four registers examined (academic, news, fiction and conversation), the highest concentration of metaphors were found in the academic text, at just under 20% . Dorst (2015) further examined linguistic metaphors in British novels, news texts, academic discourse, and face-to-face speech, expecting to find the highest concentration of metaphors in the literary genre. However, she again found that the academic discourse text contained the most metaphors, while the literary genre placed a distant third.

These results, as well as the title of this thesis foreshadow some of the findings of this study. In the exploratory phase of the corpus analysis, I was finding random instances of metaphor but no real patterns were emerging, until I suddenly realized in a moment that constituted my first breakthrough that there was a great deal of personification. I was investigating concordances of the word ‘atom’ and began highlighting all the instances of metaphor I had noted. I began to see that atoms were very interesting characters in these texts! They interacted; linked; combined to achieve; formed networks, and chemical bonds; got excited; had neighbouring atoms; and even sometimes, did not behave. It seemed only fitting to draw from these findings and have some fun with the thesis title!

On a more serious note: although the study which is the focus of this thesis suggests that EAL students may have difficulties recognizing and interpreting English metaphors, metaphor comprehension has not been viewed as a core competency that should be taught or tested in EAL classrooms (Littlemore & Low, 2006b). Given the increase in international student enrollment in North American institutions, many of

whom do not speak English as their first language, as well as the time and expense invested in recruiting and retaining these students, it is important to investigate factors that may affect their chances of academic success. In this study, students' comprehension of English metaphor is examined in the receptive language domain of reading because as Littlemore and Low (2006a) state, "foreign language learners probably need to understand metaphor more often than they need to produce it" (p. 46).

Researcher Background

My interest in metaphors began after reading Lakoff and Johnson's (1980) *Metaphors We Live By* in a Language and Power course I completed in my undergraduate studies. For my final paper in the course, I looked at the Test of English as a Foreign Language (TOEFL) preparation materials for reading (assumed representative of actual test material) to investigate the volume and types of conceptual metaphors. I hypothesised that TOEFL reading passages contained metaphors that would prove difficult for non-native speakers due to a lack of shared cultural background, increasing the gate-keeping nature of the test. I did find a good number of conceptual metaphors in the preparation material reading passages, but equally as significant, was the lack of reference to metaphor comprehension in the test questions. Roessingh and Kover (2003) explain that TOEFL does not test "concept understandings or metaphoric competence" (p. 17). In addition, Littlemore and Low (2006b) summarize that "there are few commercial second-language courses which teach metaphor as anything other than the basis of colourful idiomatic phrases" (p. 268). This led to questions about if and how EAL students learn about metaphor in English, the extent to which they do or do not

understand English metaphors, and whether any lack of understanding impedes academic success in English-speaking, Western, post-secondary institutions.

Research Questions

The original intent for this study was to make a case for the inclusion of metaphor comprehension in English proficiency testing. However, the limited number of studies that document the difficulties EAL students have in understanding and producing English metaphor, especially in the post-secondary context, meant that more evidence was needed in this realm. The goals of this study are to shed light on the prevalence of metaphors in English, including the genre of academic writing, which is not an area where heavy metaphor content would be expected; and to provide further evidence of EAL students' difficulties with English metaphor that may impede comprehension of their university reading materials.

There are three questions guiding this research:

1. To what extent are metaphors prevalent in textbooks used in first year engineering courses?
2. Is metaphor a factor in reading comprehension of engineering textbooks?
3. Is a student's first language a factor in comprehending English metaphors?

Thesis Overview

In the following chapter, metaphor and related terms will be defined, explanation provided as to why it is important for EAL learners to develop competence in understanding and using English metaphor, and the difficulties they often have in doing

so. Factors affecting reading comprehension are introduced before moving into a review of the existing literature in studies investigating metaphor comprehension. It will be revealed that although there has been some research into metaphor comprehension in the various language skill domains, more is needed, particularly in the post-secondary context. The literature will be reviewed not just on metaphor comprehension in reading since that is the focus of this study, but also in the domains of listening, writing and speaking. These studies contribute to our understanding of how learners recognize and process English metaphor.

The Theoretical Framework, the lens through which this study and its research questions will be viewed, is described in Chapter 3. Although ground breaking in its day, Conceptual Metaphor Theory (CMT) is not without its critiques, so these will be examined. The chapter concludes with an explanation of how CMT contributes to the current study, despite its flaws.

In Chapter 4, the methods are explained. The research questions were addressed with a two-phase mixed methods sequential exploratory design (Plano Clark & Creswell, 2008). This chapter provides the details of the research design, participants, instruments, and analyses from both phases of the study.

The findings of the corpus analysis, which were used along with readability factors to design the metaphor comprehension task in Phase 2, are presented in Chapter 5, followed by the results of the metaphor test. In addition, item facility and item discrimination, Spearman's rho, and simple linear regression analyses of the test task itself are presented. These findings are discussed in relation to the research questions that guided the study, existing literature, and the study's theoretical lens. In addition, since the

metaphor test task was piloted in a larger diagnostic assessment, its effectiveness and usefulness as a diagnostic tool are examined.

The thesis concludes in Chapter 6 with a discussion of implications for EAL and English for Academic Purpose (EAP) teaching pedagogy, post-admission diagnostic assessment, and post-secondary supports to facilitate the academic success and retention of (international) students. The study's limitations are outlined and suggestions made for future directions for further investigation.

Chapter 2: Literature Review

Having provided initial background to the study which is the focus of this thesis and introducing the research questions guiding the study in Chapter 1, this chapter begins with a definition of metaphor and why it is important for EAL learners. Then, some of the general issues in reading comprehension are examined, followed by a more specific discussion of why the skill of reading is important for EAL students in the post-secondary context. Next, existing studies into metaphor comprehension in the domain of reading will be reviewed. However, few studies exist in this area, particularly in the post-secondary context (e.g. Boers, 2000; Gunderson, Slade & Rosenke, 1988). As such, a discussion of metaphor studies in other language skill domains (i.e. listening, writing and speaking) are also reviewed to provide a more comprehensive picture of how non-native speakers of English cope with metaphoric words and phrases. The chapter concludes with an introduction to post-admission diagnostic assessment and the Self-Assessment for Engineering (SAFE) assessment tool.

What is Metaphor?

In the broadest sense, metaphor represents a type of figurative language, or that which is not literal. Kövecses (2010) explains that traditional thinking about metaphor can be summarized by a number of properties, including being used for artistic and rhetorical purposes, being consciously and deliberately used by those with the creative talent to do so, and being optional or unnecessary in the sense that it is used only for special effects, and not in everyday communication. On the other hand, Lakoff and

Johnson (1980) define metaphor as “understanding and experiencing one kind of thing in terms of another” (p. 5). They suggest that although metaphor is typically thought of as a way to embellish language in poetry and rhetorical writing, it is much more than that: “metaphor is pervasive in everyday life, not just in language but in thought and action. Our ordinary conceptual system, in terms of which we both think and act, is fundamentally metaphorical in nature” (p. 3). One example of how metaphor guides our thoughts and perceptions provided by Lakoff and Johnson is the TIME IS MONEY conceptual metaphor. We use language like “spending time,” “investing time,” and “saving time” because Western culture values time as if it was a physical thing, and further, as something that has monetary value. Lakoff and Johnson reason that this is not a necessary way for humans to conceptualize time and in fact, there are cultures that do not. This last point demonstrates a key property of metaphors: their culture-specific nature. This will be further discussed in the following chapter.

There are some distinctions to be made regarding metaphor, beginning with linguistic and conceptual metaphor. Linguistic metaphor “takes account of the connotations of particular words and the morphological, syntactic, and collocational characteristics of the expression” (Littlemore & Low, 2006b, p. 270). Shutova, Teufel, and Korhonen (2013) explain that these most commonly occur at the lexical (word) level, but can also be found in multi-word expressions, and even or larger sections of discourse, at which time they are referred to as “extended metaphor” (p. 304). In contrast, they explain conceptual metaphor is the abstract, underlying relationship between two words or concepts. Lakoff and Johnson suggest that linguistic metaphor results from deliberate word choice, whereas conceptual metaphor results from the more subtle, unconscious

realization of our culture-specific thought processes. Table 1 provides examples to demonstrate how the expression “to spend time” could be analyzed literally and figuratively, as well as from the perspectives of linguistic and conceptual metaphor.

Table 1 *Interpreting Literal and Figurative language, and Linguistic and Conceptual Metaphor*

Expression: to spend time			
Literal Interpretation	Figurative Interpretation	Linguistic metaphor Interpretation	Conceptual Metaphor Interpretation
This is not possible; time is not a physical thing, like currency, that can be spent.	To pass time.	Considers the word choice of ‘spend’ versus ‘pass’ time. Both constructions are metaphorical because time is not a physical thing.	Considers the underlying relationship between ‘spend’ and ‘time;’ Western culture conceives time as not only as physical, but valuable.

Another distinction is novel and conventional metaphors. Shutova et al. explain the novel metaphors as those that are made up for creative effect, and the conventional metaphors as having become entrenched in everyday speech. Nunberg (1987) suggests that all metaphors begin as novel, but become normalized or idiomatic with the passing of time. When they are no longer psychologically distinct from literal language, he classifies them as “dead” or “frozen” metaphors (p. 198). Littlemore and Low classify idioms as dead metaphors, because the non-metaphoric sense of the words has been lost. In my anecdotal experience, the terms ‘idiom’ and ‘metaphor’ are sometimes used synonymously. It depends on how broad or narrow of a view you take on the meaning of ‘idiom,’ but Kövecses (2010) summarizes the traditional meaning of idiom as two or

more words for which “the overall meaning of these words cannot be predicted from the meanings of the constituent words” (pg. 231). Examples of idioms include ‘every cloud has a silver lining,’ and ‘actions speak louder than words.’

Finally, a few other related terms require clarification. Similes are direct comparisons preceded by ‘like’ or ‘as:’ she ran as fast as lightning. Metonymy is typically (but not exclusively) a referential device where one entity is used to represent another that is related to it (Lakoff & Johnson, 1980). An example of this is the statement “She likes to read Shakespeare,” where “Shakespeare” refers to the author’s writings. Lakoff and Johnson differentiate metaphor as conceiving one thing in terms of another, from metonymy, where something is used or referenced to stand for another, allowing focus on certain aspects of what is being referred to. Synecdoche is a type of metonymy where “the part stands for the whole” (Lakoff & Johnson, p. 36). In the example ‘The student is the heart of the university,’ both ‘student’ and ‘university’ represent many students and universities, as opposed to one specific person and institution.

Why Research Metaphor Comprehension?

Previous research demonstrates the importance of EAL students acquiring the ability to understand and use English metaphors, as well as the difficulty they tend to have in doing so. One way that EAL students often stand out is by their “textbook literalness” (Danesi, 1995, p. 4). This may be explained by a lack of time in EAL classrooms dedicated to developing metaphoric competence, or as Danesi explains, “the opportunity to access the metaphorical structures inherent in the target language and culture” (p. 4). Littlemore (2001a) defines metaphoric competence as “the ability to

acquire, produce, and interpret metaphors in the target language” (p. 459). Low (1998) argues for the inclusion of metaphor in the EAL curriculum so that students can learn to create, manipulate and understand conventional and novel metaphors that are context and socially appropriate. Littlemore (2001b) argues for metaphoric intelligence to be added as a ninth form to Gardner’s 1983 model that outlines eight other forms of intelligence: visual, verbal, mathematical, kinaesthetic, interpersonal, intrapersonal, naturalistic, and rhythmic. Her rationale is that it not only makes it easier to understand metaphors, but improves overall communicative competence by facilitating use of communication strategies such as lexical innovation and paraphrasing.

Expanding the notion of metaphor comprehension’s usefulness in overall communicative competence, Littlemore and Low (2006b) demonstrate how metaphor knowledge manifests in each of the five components of Bachman’s (1990) communicative competence model: illocutionary, textual, grammatical, strategic and sociolinguistic competence. Arguably the most significant is the relationship between metaphor comprehension and sociolinguistic competence. Littlemore and Low include in this category “sensitivity to dialect, register, naturalness and the ability to interpret cultural references and figures of speech” (p. 274). They explain that errors could come from cultural transfer from students’ home language and culture, or misunderstanding or misinterpretation of the connotations of words and expressions. These culture-based errors could cause embarrassment or offence.

Littlemore and Low also demonstrate how metaphor is used in the rest of Bachman’s competencies. In doing so, they provide convincing evidence for the pervasiveness of metaphor in English and the importance of metaphoric competence.

They explain illocutionary competence as the understanding of the messages conveyed with words, including the use of ideational, manipulative, heuristic and imaginative functions. For example, metaphors can be used persuasively in descriptions or arguments. The ability for learners to identify and objectively or critically assess these constructions is important (Littlemore & Low), particularly in academic study. They explain textual competence as the “ability to understand and produce well organized and cohesive text in both written and spoken contexts” (p. 282), noting that metaphors can be used to summarize and end conversation or text, change topics, and structure arguments.

Littlemore and Low describe grammatical competence as a learner’s ability to accurately use the second language grammatical system and point out that some content treated as grammar is actually metaphoric in nature. An example is provided where the preposition ‘beyond’ is shown in its literal sense, followed by three instances of metaphorical use: ‘beyond our means’, ‘beyond my understanding’, and ‘beyond me’ (p. 285). Finally, Littlemore and Low demonstrate the role of metaphoric thinking in strategic competence for both the compensation strategies and interactional strategies used by interlocutors to make up for gaps in second language knowledge.

Various studies identify the role of metaphor in supporting other forms of knowledge. MacLennan (1994) suggests that because of certain properties of metaphor, namely its central role in language structure, the way in which it facilitates concept development, and the manner in which it cognitively links physical objects and abstract concepts, it can help in learning grammar and vocabulary. In a paper reporting her research into metaphor in academic lectures, Littlemore (2001c) suggests that metaphors are used in an evaluative capacity, citing examples from Carter (1997) and Moon (1998).

She also suggests that metaphors are used to identify and discuss new concepts at the post-secondary level because this type of vocabulary allows the speaker to “stretch the established boundaries of word meaning” (Clark, 1981, 1982, as cited in Littlemore, 2001, p. 225). More recently, MacArthur (2010) points to a number of studies that demonstrate metaphor comprehension’s role in learning phrasal verbs, idioms and vocabulary.

In addition to the numerous studies providing evidence for the value of metaphor comprehension for its own sake, as well as assisting with other forms of learning, there are others that demonstrate the difficulty that many EAL students have in distinguishing figurative and literal language and interpreting English metaphors. MacArthur (2010) sheds light on the domains of speaking and writing, noting that while some students attempted to produce written and spoken metaphors, the use was not always natural. She explains that students do not always have an understanding of the motivation for metaphoric words and expressions, their origins or referents, and as a result the ways in which particular words can be used. Kathpalia and Carmel (2011) also investigated the domain of writing, and found that although students tried to incorporate metaphor, it often proved to be unidiomatic or unnatural (not native-like). In Littlemore’s 2001 (c) listening comprehension study previously mentioned, she discovered that EAL students had difficulty recognizing figurative language, as well as correctly interpreting the meaning of some English metaphors. In a follow-up study with colleagues, she found that of the words that caused problems for students, approximately 40% were metaphoric, and students were largely unaware of their difficulties in interpreting metaphoric content (Littlemore, Trautman Chen, Koester, & Barnden, 2011).

In summary, despite all of the research demonstrating the benefits of metaphoric competence and the difficulties EAL learners have achieving it, it is not consistently or systematically taught in EAL classrooms (Kathpalia & Carmel, 2011). Metaphor comprehension is still not seen as a core competency that should be taught or tested (Littlemore & Low, 2006). Inspired by the previous work investigating metaphoric competence, and in particular, Littlemore's 2001(c) study and the follow up study with colleagues in 2011, this thesis aims to provide further evidence for the need to include metaphoric competence in the EAL curriculum and fill a gap in the literature investigating metaphor in university-level reading. The focus of this literature review now shifts to a discussion of issues affecting reading comprehension and the importance of reading as a skill for EAL academic students.

Issues Affecting Reading Comprehension

Throughout the research available on reading both in first and additional languages, the one thing that is clear is how difficult and complex it is to define exactly what we do when we read. One factor that affects this is the context or purpose for which we read. Grabe (2009b) explains that more formal reading, such as in workplace or academic contexts, may require synthesizing, interpreting, evaluating, and selectively using information from texts, especially considering the many, sometimes conflicting viewpoints we may encounter. Grabe defines six academic purposes for reading: to search for information (scanning and skimming); for quick understanding (skimming); to learn; to integrate information; to evaluate, critique and use information; and for general comprehension (often, reading for interest or entertainment) (p. 8). Jang (2009) in her

overview of reading comprehension as a diagnostic test construct explains that it has gone through a somewhat controversial evolution from being considered as a single construct or entity, to a split construct of vocabulary and general reading comprehension, to multifaceted views that include a number of sub-skills within the reading domain. She summarizes that the field is still split on how best to define the construct of reading comprehension.

Parallel development has been occurring in the area of readability. Harrison (1980) differentiates readability, or “aspects of a text which make it easy for a reader to understand” from comprehension, which concerns reader attributes (p. 33). He notes the close connection between these two concepts but recognizes the difficulty in defining the concept of reading comprehension. There have been a number of readability formulas developed to more systematically and reliably calculate the ease or difficulty of a text; Harrison provides a good overview of what existed at the time of his 1980 book. However, technological advances in the field of computational linguistics has allowed for deeper consideration of what constitutes ‘readability’ (Graesser, McNamara, Louwerse, & Cai, 2004).

Graesser et al.’s Coh-Metrix software measures over 100 indices of readability, including traditional measures such as the Flesch Reading Ease and Flesch-Kincaid Grade Level, as well as a multitude of other measures of text coherence or cohesion. Similar to the distinction Harrison makes between readability and comprehension, Graesser et al. also differentiate cohesion, which happens at the level of language and text, and coherence, which is a reader’s mental representation of a text. The more in-depth Coh-Metrix measures, which delve beyond traditional surface level measures, may

provide a more reliable picture of readability (Crossley, Allen & McNamara, 2011). For the purpose of this study there are indices that measure word concreteness, or the extent to which words are concrete or abstract, and polysemy, or different word senses.

However, the technology and research is not so far advanced as to measure the impact of figurative language on readability.

The Significance of Reading for EAL Students

A number of studies shed light on the significance of academic reading as a factor of success in post-secondary study. A U.S. national survey was recently undertaken, focusing on the five most popular majors for international students (biology, business, computer science, engineering and psychology), and targeting universities with high concentrations of international students (Anderson, Evans & Hartshorn, 2014). Faculty in these universities identified the perceived importance of various expectations and challenges of a number of academic skills and tasks, and rated listening as the most important skill, closely followed by reading. The Anderson et al. study also revealed that on average, engineering students are expected to read anywhere from approximately 15-70 pages of text per week. In a localized study at another American university, student needs were analyzed again from the faculty members' perspectives. The majority of respondents (63%) reported the belief that international students have difficulties with reading comprehension (Webster, 2014).

Students' perspectives are more revealing. In a third American study from the international student perspective, almost 40% of students from programs in science, engineering, business, and social sciences reported that they regularly did not complete

their assigned reading (Keogh, 2014). The results specifically for engineering students were higher at just over 50% of respondents. There were various reasons provided: 32% stated there was too much to complete, 14% complained of a lack of background information and another 11% stated the reading was too difficult. Similarly, Cheng and Fox (2008) report comments from science and engineering students about avoidance strategies in reading. One second-year engineering student focused on PowerPoint slides because of the short sentences, and diagrams and tables in textbooks. He reasoned that they mainly work with numbers so do not need to do a lot of reading. These studies begin to paint a picture of how important academic reading is and how (international) students may or may not be coping.

In Keogh's study above, some students indicated that a lack of background knowledge impeded their ability to complete course reading. Issues of background and cultural knowledge have come to the forefront with the advancement of research in second language reading. Schema theory predates much of this research, but has found new applications in cross-cultural views of reading comprehension. As Steffensen, Joag-Dev and Anderson (1979) explain, Sir Frederic Bartlett's original 1932 hypothesis was that a second language reader will understand a text differently than a native speaker, and likely in ways deemed as incorrect because of different schemata or background knowledge that provides "the framework for understanding the setting, the mood, the characters, and the chain of events" (p. 11). They found this to be true in their study comparing American and Indian participants' reading and recall of a text (letter) about an American wedding, summarizing that "schemata embodying background knowledge

about the content of a discourse exert a profound influence on how well the discourse will be comprehended, learned, and remembered” (p. 19).

Other studies have investigated the effects of cultural and background knowledge on reading comprehension. Floyd and Carrell (1987) also used letters to test reading comprehension. The results showed that background knowledge, and not syntactic complexity, affected student understanding and recall. In an earlier study, Johnson (1982) investigated the effect of building background knowledge on reading a text about Hallowe'en and found that between prior cultural knowledge and providing information about selected unknown vocabulary in advance, only the former facilitated reading comprehension. Learned, Stockdill and Birr Moje (2011) make a case for building students' disciplinary knowledge in the form of “word, world and domain knowledge” (p. 159).

This leads to another facet of reading comprehension: vocabulary knowledge. Not all metaphoric content is lexical in nature, but a great deal of it can be analyzed at the word level. The importance and significance of vocabulary knowledge not only in reading but in other language skills, and in both first and additional languages, is undisputed. Research on first language (L1) reading tells us that poor readers who cannot process words automatically spend more cognitive resources on this task. This slows down reading and forces more energy and attention to word recognition instead of comprehension (Rasinski & Samuels, 2011). Willis (2008) points out that readers must connect word meaning in context to knowledge about language structure. In discussing qualities of strong and weak readers, she cites Piekarz's (1956) study that revealed “poor readers focused on a literal interpretation of text whereas good readers varied their level

of attention by relying on both literal and more sophisticated interpretations” (p. 203).

Although this is in relation to L1 reading, it has direct relevance to reading in an additional language since we have seen from the previous research (e.g. Danesi, 1995) that EAL learners are overly dependent on literal interpretations.

Roessingh (2005) provides another perspective on the links between vocabulary, background knowledge and comprehension. She explores the issue of meaning making for students, which takes on a whole new dimension for students who do not share the same background and cultural understandings. She notes, referencing Delpit’s (1988) “culture of power,” that “those who are not of the dominant group are marginalized and disadvantaged by not having access to the privileged, taken for granted, ways of knowing that structure power relationships – perhaps most especially in the classroom” (p. 121). In the context of this study, this takes on significance when considering the difficulty EAL learners may have in connecting metaphoric vocabulary with their cultural referents, something that native speakers unconsciously understand and take for granted. Roessingh stresses the importance of overtly teaching culture, as well as thinking about “metaphors used by a culture to make concrete the many abstract ideas that lie beneath the surface” since they are “highly systematic, cohesive, and coherent...and are selected for their potential to help make meaning of highly abstract conceptual wholes (gestalts)” (p. 122). She goes on to provide examples she has seen from students where metaphor comprehension was impeded by confusion caused from incongruent aspects of culture.

While the links between metaphor and vocabulary comprehension are an area for further exploration, the fact remains that both play a role in reading comprehension. This study hopes to shed light on how metaphor affects academic reading comprehension,

whereas the effects of vocabulary in this domain and context are more clearly established. Grabe (2009a) points out that native-speaking English high school graduates have mastered, on average, around 40,000 words (a figure that ranges widely in the literature); however, EAL readers do not typically reach this level of word recognition. Horst (2009) tells us that extensive reading can aid in vocabulary development; however, this creates a troublesome situation for learners. They need to build their vocabulary knowledge in order to read, but are supposed to read to build vocabulary! Word lists can provide guidance for students and teachers in focusing study on the most frequently found words, but as Douglas (2013) points out in response to Horst's (2013) assertion that language teachers should focus on the 2000 most frequent English word families, this alone will not sufficiently prepare students for academic study. Using an example from a former student who was trying to read George Orwell's *1984* in order to complete a written assignment, Douglas demonstrates that even if students can comprehend 80% of the vocabulary based on the 2000 most frequent word families, the remaining 20% is more complex and poses greater difficulties. He explains this in relation to the "Cognitive Academic Language Proficiency" (CALP) stages of Cummins' (1981) model of the four stages of language acquisition. According to the model, the initial stages of language learning involve students developing "Basic Interpersonal Communication Skills" (BICS) which is comprised of language that is cognitively more simple and embedded in context. For students to move beyond these initial stages to the CALP stages, they must deal with language that is less common, more cognitively demanding, and requires them to infer context and background. As such, Douglas' student struggled to understand the story because of the complexity of a small percentage of the vocabulary, and despite the fact

that he described her English communication skills as “particularly well developed” (p. 190).

Metaphor Studies in the Domain of Reading

Having examined literature that demonstrates the importance of metaphor comprehension and the difficulties it poses for EAL learners, as well as issues affecting reading comprehension and the importance of reading in academic success, existing research into metaphor comprehension in the domains of reading, listening, writing and speaking will now be discussed.

There has been some investigation into metaphor comprehension in the domain of reading, but little in the post-secondary context. Boers’ (2000) study tested the effectiveness of enhanced metaphor awareness in economic news reports. Specifically, he investigated metaphoric expressions from *The Economist*, since a course goal for students in the study was to be able to read and critique articles from this publication. He found that some students’ reading comprehension and retention of figurative vocabulary benefited from explicit attention to the source domain or origin of figurative expressions. Students were able to make the jump from literal to figurative when given a glossary of the literal sense of the test metaphoric phrases and were a little better at retaining the meanings a few days after the initial reading comprehension test. In another study, Gunderson, Slade and Rosenke (1988) compared English native-speaking university students’ and adult EAL learners’ comprehension of literal and figurative language in newspaper articles. They found that the native speakers outperformed the second

language learners overall and understood the figurative content better than the literal, while the reverse was true for the EAL participants.

Two studies examine metaphors in textbooks. Skorczynska Sznajder's (2010) study investigated which metaphors to focus on in EAL teaching by comparing metaphors in an English business textbook to those in a corpus of business periodicals and journals, but found very little overlap. Meissner (2010) explored Norwegian grade 8 students' understanding of metaphors in English textbooks and concluded that comprehension issues arose when English metaphors were different than Norwegian equivalents or were complicated in structure. Although this latter study is not in the post-secondary context, it reveals some of the strengths and challenges EAL students face when learning English metaphors.

Metaphor Studies in Other Language Skill Domains

These studies were previously referred to above; however, further commentary about them is added here. Littlemore's 2001 (c) study is perhaps the most relevant to this context, as it investigated international students' understanding of metaphors in university lectures. It revealed that some students did not understand some of the metaphoric words and phrases, and others misconstrued meanings. Besides revealing important findings that help us understand the nature of metaphor comprehension in the receptive language skills, this study is relevant because of parallels with reading comprehension. Littlemore borrowed from Anderson and Lynch (1988, p. 13) who suggested that listening comprehension depends on three main sources of knowledge: "schematic knowledge (background knowledge and knowledge of how ideas are usually

presented), contextual knowledge (knowledge of the situation and of what has already been said) and systemic knowledge (knowledge of the language)” (p. 334). The role of background and contextual knowledge was discussed above in relation to reading comprehension.

Irregular use of figurative language seems to be a common finding in the studies examining productive skills. MacArthur (2010) investigated fourth-year Spanish university students enrolled in an English language/literature program. In addition to previous comments about students producing metaphors that were sometimes unnatural, she notes that instructors should be aware of metaphors used in metalanguage used to provide verbal feedback. Kathpalia and Carmel (2011) found similar results to MacArthur. They looked at a corpus of writing samples from first-year Science & Arts university students in Singapore to analyze types of figurative language problems in the English as a lingua franca context. Specifically, they judged metaphoric competence in terms of grammatical, textual, illocutionary and sociolinguistic competencies. Problem areas included issues with nouns, determiners, and prepositional phrases; incorrect idioms and clichés; and awkward and incorrect use of metaphors as signalling devices. There were few examples of metaphors in illocutionary functions, but the metaphors used were done so to convey information and evaluate. Extended metaphors used to persuade, explain and entertain largely resulted in clichés, misused metaphors and awkward humour. In the area of sociolinguistic competence, they found inaccurate direct translation of metaphor from students’ L1 or inappropriate register use (i.e. level of formality).

Post-entry Diagnostic Assessment and Self-Assessment for Engineering

The literature review now shifts gear in the last section to introduce post-entry diagnostic assessment and the tool used at the Canadian university where this study took place. Diagnostic assessment is defined by Bachman and Palmer (2010) as identifying specific areas of strength and weakness of student abilities from test results, ideally to inform decisions made by both instructors and students on what can facilitate learning. This latter point is crucial, as noted by Fox, Haggerty and Artemeva who argue “a diagnostic assessment procedure cannot truly be diagnostic unless it is linked to feedback, intervention and support” (forthcoming). The potential of diagnostic assessment in identifying at-risk students, as well as the potential for its use in designing context-specific and personalized interventions, was recognized by researchers at the University of Melbourne and the University of Auckland, who developed the Diagnostic English Language Assessment (DELA) and the Diagnostic English Language Needs Assessment (DELNA) respectively (Fox, von Randow & Volkov, in press). Post-Entry Language Assessment (PELA) is now a trend that has garnered global attention (Read, 2008, 2013, as cited in Fox, von Randow & Volkov).

There has also been increased recognition of the value of formative assessment, and the role that diagnostic assessment can play in informing curriculum and directed, individual student development. Rea-Dickins (2001) notes that with formative assessment, stakeholders (e.g. institutions, instructors, students) take the results of assessments and act upon them in some way. In other words, these types of assessment are a beginning point. This is in contrast with summative measures that happen at the end of a learning period, as a final snapshot of student achievement. Diagnostic assessment

shares formative's goal to act upon test results. In the case of diagnostics, the specific aim is to use identified learner strengths and weaknesses for targeted learning and development. Both Green and Weir (2004) and Fox (2009) completed studies that resulted in recommending the benefits of diagnostic assessment in providing institutions and instructors detailed information for targeted instruction. Green and Weir investigated "low-cost placement instruments" that some universities turn to as a cost-saving measure. However, these were found to provide weak, surface-level information about students that proved ineffective in terms of diagnosis. Similarly, Fox's study was undertaken in response to top-down measures that saw English proficiency test scores used for placement purposes. They were also ineffective; however, diagnostic measures that were implemented did have some positive effects. Most of the instructors who participated in the study used the diagnostic information to work individually with students, which was a departure from traditional teaching methods. Fox concluded that "diagnostic assessment leading to targeted instruction that focuses increased attention on individual student needs and strengths is both efficient and effective in supporting language development, and works well in tandem with traditional EAP approaches" (p. 38).

Post-entry diagnostic assessment is increasingly used as a tool to boost student success and retention, in part, due to increasing economic constraints and diversity in post-secondary classrooms (Fox, von Randow & Volkov, in press). While it is important to help all new students increase their chances of success and improve their academic experience, this process could arguably be more crucial for international students whose first language is not English. These students leave their home cultures and support systems to travel great distances for their studies, and the international tuition rates they

pay are significantly higher than domestic rates. The Canadian Federation of Students' web site reports that as of 2012, international undergraduate students paid on average three times more than Canadian citizens (Canadian Federation of Students, 2015). An added challenge for institutions recruiting these students is the proficiency test scores that non-native English speaking students submit in support of their applications. English proficiency tests are typically designed to test general language competence, and are not necessarily indicative of academic language skills:

Defining English proficiency in terms of the bands, grades or scores that students obtain on entry, or with reference to the broad general descriptors compiled by the instrument developers, is clearly inadequate if there is no institutional process to link the measures that universities accept to the lived experience of the tertiary classroom... (Dunworth, 2010, p. 6).

Considering the limited resources universities have in assessing qualifications for large numbers of applicants and providing customized resources to assist students once they have begun their studies, early diagnosis of academic problems and timely, targeted intervention could go a long way in improving students' academic experience and success.

At the Canadian university where this study took place, Fox approached several departments to discuss diagnostic assessment in 2008. In 2009, the Faculty of Engineering began working with her in part, because the generic, universal support programs available to students were having little impact on course failure and retention rates in their faculty (Fox & Haggerty, 2014; Fox, Haggerty & Artemeva, forthcoming). Self-Assessment for Engineering or "SAFE" was piloted in 2010 (Fox, Haggerty &

Artemeva). The first iteration of SAFE included three tasks borrowed from DELNA, designed to test academic literacy in English or “the ability of incoming undergraduate students to cope with the demands of their degree programme” (Read, 2008, p. 184). In subsequent years this content was adapted to fit the localized, engineering context at this Canadian University. A mathematics task was added, the generic DELNA writing task was replaced with one specific to engineering, and the test went from being administered on-line, before the start of students’ programs, to being embedded into a mandatory class for all new, first-year students (Fox, Haggerty & Artemeva).

The current iteration of the assessment now includes an integrated reading/writing/listening task based on a class lecture and background reading material, DELNA-based Academic Word List and reading (cloze elide) tasks that survived from the original version, a mathematics task that was created by the Faculty of Engineering, and a multiple choice reading comprehension task based on a chemistry text (Fox, Haggerty & Artemeva). Students also provide biographic and demographic information, and quantitative responses to ten questions relating to academic preparedness, dealing with stress, and competition in a questionnaire that accompanies the assessment.

Chapter Summary

In this chapter the concept and different types of metaphor were defined, and the related concepts of idiom, simile, metonymy, and synecdoche were clarified. Research was reviewed that demonstrates the importance of developing metaphorical competence for EAL learners, as well as the struggles they have in identifying, understanding and producing English metaphors. Factors affecting reading comprehension were reviewed,

including readability measures and the importance of (cultural) background. The significance of reading skills for success in academic study was also discussed. Next, research on metaphor comprehension in the domain of reading was reviewed, followed by studies in other language skill domains. While these studies all contribute to the body of knowledge in this area, it becomes clear that we have much to learn about non-native speaker comprehension and use of metaphors in the various skill domains, especially in the post-secondary context. The chapter concluded with an introduction of the basics of diagnostic assessment and an overview of the historical development of SAFE.

In the following chapter, the facets of theoretical framework through which this study is framed, Lakoff and Johnson's (1980) Conceptual Metaphor Theory, are described. Although this theory was quite influential when introduced, it has received its fair share of criticism. The major critiques are outlined and discussed. The chapter concludes with an explanation of why this theory was chosen, despite its flaws.

Chapter 3: Theoretical Framework

This chapter provides an overview of Lakoff and Johnson's Conceptual Metaphor Theory, which is a cognitive-based theory about how humans process metaphoric content in language. Although ground-breaking in its day, it has garnered a great deal of attention over the decades and drawn criticism. Various researchers have expressed concerns over various facets of the theory, and these critiques are discussed, as well as the evolution of CMT. The chapter concludes with a discussion of what this theory contributes as the lens through which this study and its research questions (the prevalence of metaphors in engineering texts, whether metaphor is a factor in reading comprehension, and whether first language affects metaphor comprehension) are examined.

Conceptual Metaphor Theory

The theoretical lens that this study and its guiding research questions will be examined through is Conceptual Metaphor Theory (CMT). Lakoff and Johnson (1980) argue that our conceptual system, guiding our thoughts and day-to-day functioning, is not something we think of consciously. As such, one way we can find evidence as to how this works is through the language that we use, "since communication is based on the same conceptual system that we use in thinking and acting" (p. 3). They suggest that this system is metaphoric in nature and provide examples from common speech to demonstrate not only how metaphor manifests in our thought processes, but how pervasive it is in English. Littlemore and Low (2006b) provide an easy-to-understand explanation of how conceptual metaphors are expressed through an "A is B" structure,

where ‘A’ represents a source domain and ‘B’ a target domain (p. 270). For example, in the THEORIES ARE STRUCTURES metaphor, ‘theories’ is the target domain’ and ‘structures’ the source domain, and the metaphor manifests in statements like “That is hardly a strong *foundation* for your theory” and “The theory needs more *support*” (p. 270). In the first statement, the concept of theory, the target domain, is conceptualized as a structure with a foundation, like a building, and this is the source domain or how the target domain is realized metaphorically. Similarly in the second statement, theory is envisioned as a structure needing to be strengthened.

There are three main categorizations of conceptual metaphors presented by Lakoff and Johnson: structural, orientational and ontological, the latter which further subdivides into entity and substance metaphors, and container metaphors. Personification, where objects or concepts are given human characteristics, allowing us “to comprehend a wide variety of experiences with nonhuman entities in terms of human motivations, characteristics and activities” (p. 33), also falls under the category of ontological metaphors. Each of these categorizations will be explained and examples provided to demonstrate how they are used in everyday speech.

Structural metaphors are broad in scope, demonstrating the way that we form and perceive communications. Two common structural metaphors are ARGUMENT IS WAR and TIME IS MONEY (Lakoff & Johnson, 1980). Lakoff and Johnson suggest that native English speakers with Western cultural background use patterned and consistent language when communicating about arguments and time. In the various expressions that the ARGUMENT IS WAR metaphor appears, we see that “many of the things we do in arguing are partially structured by the concept of war” (p. 4). We can, win, lose, defend,

shoot down, and destroy arguments, claims and strategies. “We see the person we are arguing with as an opponent. We attack his position and we defend our own” (p. 4).

Similarly, in the TIME IS MONEY metaphor, the ways in which this metaphor manifests demonstrates the extent to which we see time as not only a physical thing, but one of great value that can be saved, spent, wasted, invested, put aside, spared, borrowed, lost, and so on.

Oriental metaphors as the name suggests, concern directions or orientations. They are typically expressed as binaries or opposites, such as up-down, happy-sad, in-out and front-back (Lakoff & Johnson). Lakoff and Johnson describe these types of metaphors as spatial orientations based on our physical bodies and how they function in our physical environment. A simple example is the HAPPY IS UP; SAD IS DOWN binary: She is in high spirits; he fell into a depression. Lakoff and Johnson suggest that these metaphors are embodied to the point that they manifest physically; our posture slumps when we are sad, and perks up when we are in a positive or happy mood.

Whereas the spatial orientations of orientational metaphors “provide an extraordinarily rich basis for understanding concepts in orientational terms” (Lakoff & Johnson, p. 25), ontological metaphors allow us to relate experiences and concepts to physical objects with dimensions. “Once we can identify our experiences as entities or substances, we can refer to them, categorize them, group them, and quantify them – and by the means, reason about them” (p. 25). Lakoff and Johnson suggest that ontological metaphors are based on human nature: “human purposes typically require us to impose artificial boundaries that make physical phenomena discrete just as we are: entities

bounded by a surface” (p. 25). They assert that ontological metaphors are necessary for us to rationally deal with our experiences.

There are three main types of ontological metaphors. Entity and substance metaphors allow us to discuss an abstract concept like the mind as if it was a physical object. As such, we say things like “the wheels are turning,” “I am a bit rusty today,” and “she is running out of steam” in comparing the mind with machinery (THE MIND IS A MACHINE) (p. 27). Similarly, container metaphors allow us to conceptualize things with physical dimensions or boundaries in a number of ways. Lakoff and Johnson point out that imposing boundaries or marking territory is a basic human instinct, so although we may not see a natural boundary to a physical area, we can see ourselves as *in* or *out* of a clearing or the woods, for example. (p. 29). We also tend to quantify our visual field and this manifests in statements like “I have him in sight” and “the ship is coming into view” (p. 30), which represent the VISUAL FIELDS ARE CONTAINERS metaphor. Events, actions, activities and states are also conceptualized through ontological metaphors. As such, in talking about the abstract concept of an event like a race, we say things like “did you *see* the race,” “are you *going to* the race” and “he is *out of* the race,” which position the race as a physical object (p. 31). Examples of emotional and physical states as objects include “being *in* love,” “*out of* trouble” and “getting *into* shape” (p. 32). The examples provided above also demonstrate the extent to which ontological metaphors manifest in prepositions, a notoriously difficult part of speech for EAL learners to master (Swan, M., 1988; Mukundan & Roslim, 2009).

Finally, the last type of ontological metaphor common in English is personification, which “allows us to comprehend a wide variety of experiences with non-

human entities in terms of human motivations, characteristics, and activities” (p. 33). As such, we often give concepts or objects human traits or place them in the position of doing human actions: the law *tells us* we cannot steal; debt *caught up* with him; their ideas *gave birth* to a new line of inquiry. In each of these examples, a non-human entity is put in the subject position and paired with a verb denoting human activity, none of which are literally possible. The law cannot speak, debt cannot catch and ideas cannot give birth.

There are two key characteristics of CMT that Lakoff and Johnson identify that are particularly relevant to this study: the systematicity or consistency of metaphors, and their cultural specificity. Systematicity and consistency are evident in structural metaphors. In the ARGUMENT IS WAR example, Lakoff and Johnson explain that there are typical things we do and do not do in arguing and the language used is not by chance; this speech is patterned and consistent. Similarly, the ways in which we talk about time is consistent because of the ways in which Western culture conceive of the concept of time. This idea is explored further below. We do see examples of systematicity and consistency in other types of metaphors. In the orientational metaphors HAPPY IS UP; SAD IS DOWN, we do not have expressions like ‘feeling low’ to mean that we are happy, optimistic or positive, and vice versa in an opposite example. Lakoff and Johnson do not provide specific examples of systematicity in ontological metaphors; however, we can imply it from their discussion of these metaphors as human nature. We are creatures of habit, and Lakoff and Johnson explain that ontological metaphors are necessary for us to rationally deal with our experiences. This, however, is an area that needs to be empirically explored.

Another key characteristic of CMT that Lakoff and Johnson discuss is how they are culture-specific. They assert that “The most fundamental values in a culture will be coherent with the metaphorical structure of the most fundamental concepts in the culture” (p. 22). They explain how the TIME IS MONEY metaphor stems from the development of the concept of work in Western society and our customs of paying for work and other things in increments of time: wages per hour, hotel rooms by night, telephone charges by the minute, and so on (p. 8). Of importance to note is that “These practices are relatively new in the history of the human race, and by no means do they exist in all cultures” (p. 8). The ways in which we conceptualize direction and express orientational metaphors also demonstrate their relation to Western culture: “For example, in some cultures the future is in front of us, whereas in others it is in back” (p. 14). Another example is the way in which Western culture conceives of things in the up-down opposing orientations, whereas other cultures favour balance or a middle-of-the-road approach. Lakoff and Johnson conclude that the values demonstrated in metaphoric expressions are consistent with and deeply embedded in our culture.

Critiques of Conceptual Metaphor Theory

Lakoff and Johnson’s CTM had significant impact when it was introduced, as it boldly took the field in a new theoretical direction by suggesting the cognitive connection in the way that humans understand and produce metaphors (Cameron, 2003). However, as with any theory, it is not perfect and has drawn its fair share of criticism over the decades. Lakoff and Johnson are forthcoming about some of these flaws in *Metaphors We Live By*. First, they admit that the systematicity that demonstrates our conceptual

thinking necessarily hides other inconsistent aspects of a metaphor. For example, in ARGUMENT IS WAR, there are co-operative aspects of argument that are overlooked, such as the time (a valuable commodity in Western culture) an opponent gives in engaging in the discussion (p. 10). They also concede that within cultures, not all members hold the same beliefs. There are conflicts in values, especially in sub-cultural groups: “The various subcultures of a mainstream culture share basic values but give them different priorities” (p. 23). For example, there are metaphors relating to spending habits such as BIGGER IS BETTER, THERE WILL BE MORE IN THE FUTURE, and SAVING RESOURCES IS VIRTUOUS. One can easily see conflict between these depending on the socio-economic status of those making the decisions, and state of the economy. Finally, Lakoff and Johnson discuss the coherence and consistency of metaphors, noting that there can be different senses of a broad metaphor that do not evoke a consistent image, but that are coherent in other ways. They provide an example with the idea of a journey, which can happen by various means of transport, all with their own metaphors and imagery, which can be a problem when expressing other metaphors involving ‘journey,’ such as in LOVE IS A JOURNEY: Look how far we have come; they are at a crossroads; we cannot turn back now; it has been a bumpy road; their marriage is on the rocks; and so on (p. 45). Lakoff and Johnson reason that although these metaphors evoke visions of different forms of transportation, they are all forms of a journey. This particular critique is one commonly identified by other researchers and discussed further, below.

Kövecses (2008) summarizes the five most common criticisms of CMT. The first is based on categorizing metaphors by concepts instead of words, and the more subjective

way in which they are identified. Second, the direction of analysis, or more specifically, the top-down approach of this theory is also critiqued because the focus is on “the conceptual metaphor itself as a (hypothetical) higher-level cognitive structure” instead of “language and linguistic metaphors, as well as their behaviour in specific contexts” (p. 170). The third most common argument is based on the level at which conceptual metaphors should be formulated because each element of a source domain cannot necessarily be used to talk about the same target. This was discussed above in terms of metaphor coherence and consistency. Next, embodiment, or the ways in which the human body plays a role in the formation of meaning, is criticized in CMT. Kövecses explains that this is due to the way in which image schemas, or how we use our physical experiences to make sense of the world, try to account for both universality and cultural specificity at the same time. The final argument is based on a critique of the relationship between metaphor and culture, relating back to the previous critique about universality and cultural specificity. This critique includes comments about cultural inconsistencies, which Lakoff and Johnson address in their book.

Why Conceptual Metaphor Theory?

The question that remains is: Why use CMT as the theoretical framework through which to examine this research, knowing it is flawed? There are several reasons why it remains a valid choice. First, no theory is perfect when critically scrutinized. The researcher must determine whether in a particular setting a theory is relevant and informative, recognizing its imperfections, or whether these weaknesses detrimentally affect the way in which findings may be interpreted. This decision depends, in part, on

the study, research questions, and how the theory is being applied. In this instance, the study is not designed to refute or support CMT. Rather, the theory contains elements that this researcher believes will prove insightful when used to examine the data.

Kövecses (2008) addresses each of the five most frequent critiques of CMT. In some cases he sees the differences of opinion as misinterpretation of the theory, or not interpreting it beyond the surface level. For example, in the critique about the level of analysis, the argument is that one conceptual metaphor cannot necessarily explain all of the expressions tied to it in the same manner. In the examples “add fuel to the fire” and “flare up” (p. 172) it is reasoned that the conceptual metaphor ARGUMENT IS FIRE cannot entirely explain the difference in the meaning of these two phrases. Kövecses counters this argument by pointing out that while ARGUMENT IS FIRE sets up the global conceptual metaphor, this is only the first step. There are finer levels that this larger metaphor can be broken down into and it is at these sub-levels that these expressions can be explained.

Legitimate shortcomings of CMT are discussed by Kövecses and he makes suggestions on how to strengthen the theory. For example, Kövecses indicates that establishing the appropriate level of schematicity, or the level at which metaphors are formulated, is a necessary step. This will assist in determining “which elements of the source domain are mapped onto the target, and which ones are not” (p. 175). For example, the LOVE IS A JOURNEY metaphor previously discussed can occur at the level of vehicle (love can drive someone to despair), but not the finer level of specific types of vehicles (e.g. train, boat) (p. 175).

Kövecses' own research contributes to modernizing CMT with the “differential experiential focus” concept he developed. It allows the idea of embodiment to be broken down into smaller parts that can be considered, emphasized and conceptualized individually and differently by various cultures (p. 178). In his book *Metaphor in culture: Universality and variation* (2005) investigating how universality and culture specificity in metaphor conceptualization can be simultaneously addressed, he discusses how both embodiment and context, the former leading to metaphors at the universal level and the latter determined by local culture, must be considered. He reasons that our conceptualization of metaphors is not automatically at the universal level, but is shaped by (local) context and other more personal factors. He concludes that expanding or reshaping our traditional, more narrow view of CMT will “open up new possibilities of understanding of linguistic and cultural creativity within the framework of CMT” (p. 181).

Regarding the argument that those who share a language and culture will vary in their beliefs, I would add the following. I believe that while choosing to use metaphoric expressions that represent different subcultural values, those sharing the same language and overarching culture would still understand the nuances that certain metaphors capture. In the example previously discussed relating to spending habits such as BIGGER IS BETTER, THERE WILL BE MORE IN THE FUTURE, and SAVING RESOURCES IS VIRTUOUS, I would argue that although members of different subcultures may have different priorities relating to spending and saving, they would still understand the intent of utterance that reference these metaphors. However, this argument is a hypothetical example based on native-speaker intuition that needs to be empirically tested.

An opponent to CMT, Cameron takes CMT as her inspiration, but “rejects its formulation of metaphor in terms of highly generalized and abstract conceptual domains that pre-exist actual use of metaphors in language” (2010, p. 77). She introduces an alternative, discourse-based framework that draws on cognitive psychology, Vygotskian sociocultural theory and discourse analysis, and is “designed to apply to metaphor in language use in social interaction” (p. 77). While an interdisciplinary approach to metaphor is useful, Cameron makes confusing claims about CMT that form the basis of her framework. Specifically, she suggests that CMT “artificially separates” language, thought, and culture and that it “downplays language in order to focus on the cognitive, and disputing the role of culture” (p. 78), even citing Lakoff and Johnson amongst others in making the claim. However, it is difficult to conceive how language, thought and culture are separated when Lakoff and Johnson insist that the conceptual system that guides our thoughts, that stems from our deeply embedded cultural values, manifests in our language. Cameron also criticizes the top-down application of CMT and the “highly generalized labels that are used,” since this results in losing information. However, the discourse dynamics framework is interested in “patterns of metaphor vehicles used across a discourse event, or particular bounded episode of discourse” (p. 80). Her argument is valid with regards to top-down application versus using naturally occurring discourse, but the distinction between labelling conceptual metaphors and searching for “patterns of metaphor vehicles” is murky at best.

To Cameron’s credit, her development of the concept of “metaphor shifting” (p. 88) and “systematic metaphors” (p. 91) could contribute to CMT to address metaphoric coherence and consistency. Ultimately, Cameron’s framework is clearly based on spoken

discourse, which is not surprising since she is famous for her work in this area. However, its weaknesses do not make it the best choice in this context of examining written discourse.

Kövecses is optimistic about advances in cognitive theories of metaphor. He describes the “Neural Theory of Metaphor,” derived from the “Neural Theory of Language” proposed by George Lakoff and Jerry Feldman, which carries the extension of metaphor from the language (linguistic metaphors), to the mind (conceptual metaphors) to the body (physical basis of metaphor), and finally to the brain (p. 87). A detailed description of this advancement is beyond the scope of this thesis (see Kövecses, 2010 and Lakoff, 2008 for increasingly detailed discussions), but this view explores how metaphors stimulate neurons in our brains. The contributions pointed out by Kövecses include that some conceptual metaphors are more easily learned and an explanation is provided of why the processing of conventional metaphors does not take longer than processing of nonmetaphorical language.

When it comes to theorizing about how humans interpret metaphor, researchers typically embrace one of two sides and never the two paths shall cross! As Littlemore and Low (2006b) explain, “the field of metaphor in particular is currently characterized by warring territorial groups who take few prisoners” (p. xv). One side favours the linguistic perspective and the other, cognitive-based approaches such as CMT, or offshoots. I, however, believe that these two theoretical perspectives need not be opposites, but complementary. Both approaches have shortcomings that the other begins to address. Further research is required in order to better understand how humans, with all their linguistic and cultural diversity, understand metaphor and other forms of figurative

language. For the purposes of this research, CMT has the potential to shed light on how students with different linguistic and cultural backgrounds understand metaphors they encounter in their engineering studies in large part, because of the systematic and consistent nature that metaphors are used, and their culture-specific nature.

Chapter Summary

An overview of this study's theoretical framework, Conceptual Metaphor Theory, was provided. Weaknesses of the theory recognized by Lakoff and Johnson, as well as critiques rendered by other researchers were explored. Advancements in CMT and cognitive-based theories of metaphors were touched upon and justification provided of why CMT was chosen as the theoretical lens through which to view this investigation and the guiding research questions addressing the prevalence of metaphors in engineering textbooks, whether metaphor comprehension factors in reading comprehension, and whether first language affects metaphor comprehension.

In the chapter that follows, the method of this two-part study is detailed, beginning with how a corpus of first-year engineering reading materials was examined for metaphoric content. These results drove the second phase of the study, in which a short metaphor comprehension test was developed and administered as part of an existing diagnostic assessment that all first year engineering students complete at this Canadian university.

Chapter 4: Method

Thus far, metaphor and related terms have been introduced and explanation provided as to why metaphor comprehension is important for EAL learners. Relevant literature was reviewed relating to reading comprehension and the importance of reading in EAL students' academic success. An overview was provided of existing metaphor comprehension studies in the language skill domains of reading, listening, writing and speaking. And post-admission diagnostic assessment and SAFE were introduced. The theoretical framework was detailed and its critiques discussed before providing a justification as to why it was chosen and what it can contribute to this study. Now the focus shifts to the method, findings, discussion and implications of this study.

Method overview

In this investigation, a two-phase mixed methods sequential exploratory research design (Plano Clark & Creswell, 2008) was applied in order to address the research questions. Plano Clark and Creswell explain that in a sequential design, the data is analyzed independently for each of the qualitative and quantitative phases. Putting together the sequential and exploratory parts of the design, they further explain that the first phase typically consists of collecting and analyzing qualitative data, followed by a second phase in which quantitative data is collected and analyzed. Sometimes, they note, there is a third phase to qualitatively explore the quantitative findings from Phase 2. The phases are integrated at the interpretation stage of analysis. Figure 1 below provides a visual representation of the study design. Plano Clark and Creswell note that this type of

study is used to explore a phenomenon; in this case, the phenomenon of interest is the use of metaphors in engineering reading materials and, whether (and how) they affect reading comprehension.

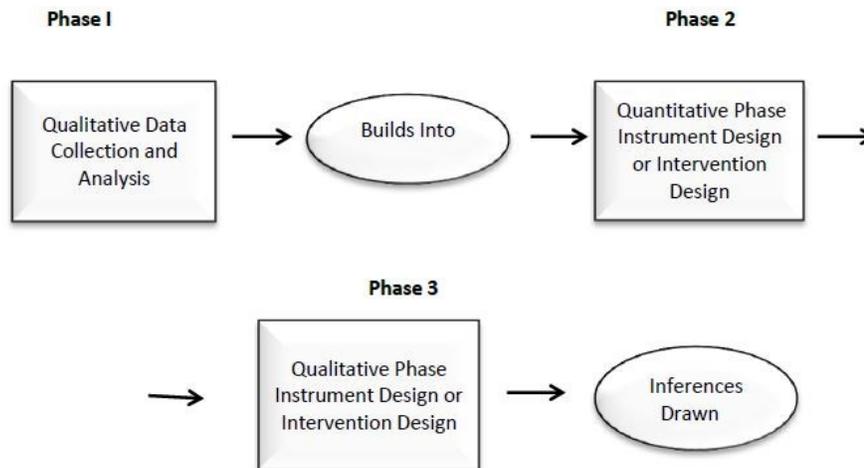


Figure 1 Sequential Exploratory Mixed Methods Design.
(Creswell, 2015, p. 41, as cited in Turner & Fox, 2015).

Phase 1 addressed the first research question, while questions 2 and 3 were investigated in Phase 2. Although the hope was to add a Phase 3 to interview some of the test takers about their understanding of English metaphor, it was beyond the scope of this Master's thesis. However, this would be a logical next step to pursue (see the discussion of future research in the conclusion, Chapter 6). In the following sections the participants, instruments, procedures and analysis of both phases are explained. This includes details about the students who completed the metaphor comprehension test, the corpus and the metaphor test task, the qualitative text analysis, and the quantitative analysis of students' responses to the metaphor comprehension test.

Participants

In January, 2015, a cohort of newly admitted engineering students completed the SAFE diagnostic assessment, part of which was the metaphor comprehension task. Although the initial hope was to collect data from the September intake, the metaphor comprehension task (the instrument used in Phase 2, discussed below) was not ready to be administered when this cohort was assessed. The winter cohort represented a second, mid-year intake of students that is much smaller than the one that commenced in the fall term. The 42 students that comprised this cohort were in some ways diverse, and in other ways, not. They ranged in age from 16 to 31, with the majority being 18 or 19. There were only four female students. As many of the students had just finished an EAP program and satisfied the English language requirement of their conditional admission², this group was ethnically and linguistically diverse – more so than the larger, fall cohort. Less than half of the winter group were Canadian citizens and most of the international students had been in Canada one year or less. All but three students (native English speakers) were multilingual, speaking at least two, if not three or four languages. Table 2 (below) summarizes the various languages spoken, and supports what IIE data (2013b) tells us about the most popular countries of origin of international students studying at Western, English speaking institutions.

² The exact number of students who entered the engineering program after satisfying an English language admission condition was not available from the biographic and demographic information collected with the administration of SAFE.

Table 2 *Foreign First Languages Spoken by Engineering Cohort*

Language	Number of Speakers
Arabic	11
Chinese	11
Farsi	2
Bangla	1
Bengali	1
Konkani	1
Korean	1
Tamil	1
Thai	1

Also of interest is data collected regarding the languages that students spoke at home, and whether they completed their secondary education in an English-medium institution. Of the nine native English speakers, three spoke another language at home. Both of the native French speakers also spoke French at home. Of the students whose mother tongue was a foreign language, seven spoke English at home. Not surprisingly, all of the native English speakers completed their secondary education in English. One of the French speakers completed English high school, as did 20 of the foreign language speakers, which may be a more surprising result. One additional non-native English speaker completed one year of English high school. Nine of the EAL students completed their high school in another language, but two of these were amongst those who spoke English at home. Figure 2 represents the foreign languages spoken and whether students completed their high school education (“HSE”) in an English-medium institution (the yes

and no sides of the graph).

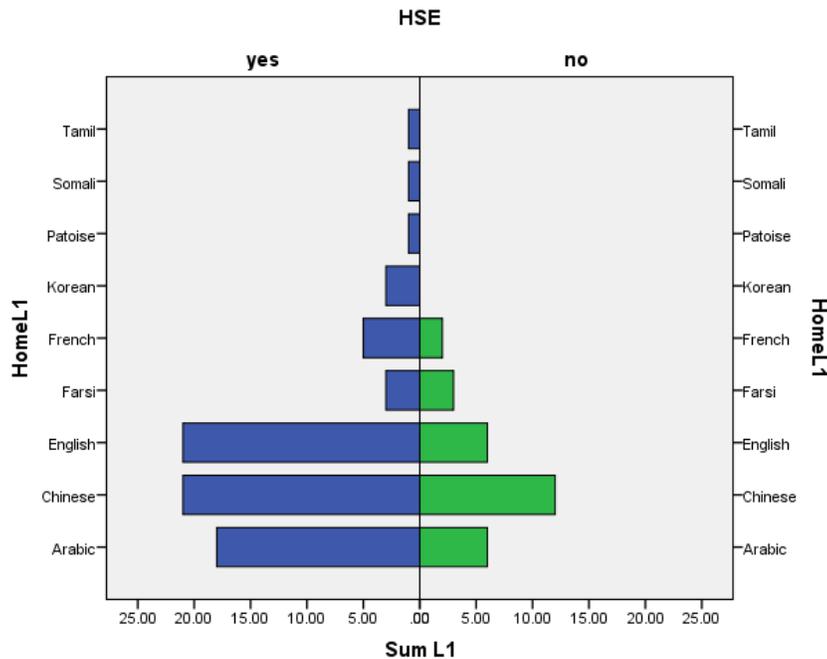


Figure 2 Foreign Languages Spoken and English-medium High School Education.

Instruments

There were two instruments used to complete this study. In Phase 1, an existing corpus of first-year engineering texts was used (Wood & Appel, 2012). The corpus consisted of reading material from five engineering sub-disciplines. Physics was proportionately the largest, with 242 505 words, followed by chemistry with 222 931 words, calculus with 180 831 words, problem solving and computers (hereafter referred to as computers) with 92 973 words, and engineering mechanics (hereafter referred to as mechanics) with 65 499 words. The entire corpus contained 804 739 words. The texts represented both required and recommended readings from first year engineering classes at the time of data collection, minus “instructional language (i.e. end of chapter problem

sets, instructional exercises/activities etc.)” that was removed from the corpus (Wood & Appel, 2014, p. 4).

The instrument used in Phase 2 of the study was the metaphor comprehension task. The final iteration consisted of three small paragraphs of text with ten underlined metaphoric words or phrases that constituted novel and conventional metaphors, idioms, and linguistic and conceptual metaphors. The test in its entirety can be viewed in Appendix A, which provides the context for the ten metaphoric words and phrases tested summarized in Table 3 (below). Each of the underlined words/phrases was attached to a comment box on the right-hand side of the page that reiterated the word or expression that the students were asked to explain in their own words, along with space to provide a written response. Although the initial aspiration was to include metaphoric content from all five sub-disciplines, the final text chosen came exclusively from the chemistry part of the corpus. The process of how the task was developed is described below, in Phase 2 of Procedures.

Table 3 *Words and Phrases Tested in the Metaphor Comprehension Task*

Test Item Number	Test Word/Phrase
1	introductory level
2	chemical connection
3	experience gained
4	employ
5	carrying out
6	led to a dead end
7	body of knowledge
8	are closely tied to
9	harnessing
10	relationship

Procedures

Phase 1

The study began in Phase 1 with a qualitative textual analysis of an existing corpus of engineering texts (Wood & Appel, 2012). This initial phase of analysis was important not only in determining whether metaphor was prevalent in the corpus, but also in providing authentic material for the metaphor comprehension test in Phase 2. Studies using corpus analysis are generally considered quantitative by nature. As Deignan (2005) explains, a software program is typically used to search the text for concordances, or instances of words or word forms. Had this been the primary focus and where the analysis ended in Phase 1, it would rightfully have been classified as quantitative. The goal of this phase of the analysis was not to document and define every single instance of metaphor, but to gain an impressionistic view of the prevalence and types of metaphors. As such, the corpus tools were used in the initial stages of exploration in order to examine the large quantity of text. Determining the most frequent words was the beginning point, and the concordance and collocate tools assisted to narrow down the final text that was examined for metaphoric content. The goal and the crux of the analysis in this phase was the manual, qualitative analysis of the collocations examined. The process of how this occurred is described in the analysis for Phase 1, below.

A further distinction that is often made with corpus studies is the classification of “corpus-based” versus “corpus-driven” research (Tognini-Bonelli, 2001, as cited in Deignan, 2005, p. 88). Deignan explains the former as starting with an existing paradigm and studying it via the corpus, and the latter as beginning with no assumptions and investigating what emerges from the corpus. This study is classified as corpus-driven as

no assumptions were made about the volume and types of metaphors that would be found. However, I must make a clarification in making this claim. Although these texts represent formal, academic writing, I expected to find *some* degree of metaphor due to its prevalence in the English language, as well as the results of the Steen et al. (2010) and Dorst (2015) studies mentioned in the introduction of this thesis. However, I did not make specific assumptions with regard to the volume or types of metaphors that may have been found.

Phase 2

In Phase 2 of the study, the results of the corpus analysis in Phase 1 were used to design the metaphor comprehension test instrument described previously. It proved challenging to develop a task that stayed true to the test construct (students' comprehension of English metaphors), and that facilitated statistical analysis. An initial attempt was made to develop a multiple choice test; however, it was not possible to do this and maintain the test construct. It was therefore decided to ask students to write definitions in their own words and develop a rating scale to code responses. This was a more interactive activity according to Bachman and Palmer's (2010) test usefulness criteria, and allowed for richer, more meaningful data that was statistically analyzable.

It also proved challenging to provide sufficient and readable context for the test words/phrases since this was also selected directly from the corpus. The initial iteration included ten sentences with metaphoric words or phrases representing all of the engineering sub-disciplines in the corpus. However, independent review of this version revealed that the task was too difficult due to insufficient context. This is consistent with

the findings of Henning's 1991 investigation into TOEFL vocabulary test items, that embedding the words in a reading passage was more effective than presenting them in sentences. Determining which part(s) of the corpus the test paragraphs would be taken from was a process of elimination that began by considering which sub-disciplines contained the largest concentration of metaphors: chemistry and physics. Within these two sub-disciplines, I searched for passages that not only contained simply-structured metaphors (i.e. those that were not split between different parts of a sentence or were not multi-word extended metaphors), but were also simple and straight forward in terms of readability. In the end, the physics passages were determined to be too complicated. The final iteration of the metaphor comprehension test task included three small paragraphs of text taken directly from the chemistry portion of the corpus.

Since the determination of readability was made intuitively by the researcher as a native English speaker not familiar with the subject content, the three test paragraphs were analyzed using Coh-Metrix version 3.0. Coh-Metrix is a software program that analyzes text for over 100 readability indices, including a number of more in-depth cohesion and coherence measures that go beyond the surface level qualities of traditional readability indices (Graesser, McNamara, Louwerse, & Cai, 2004). Selected results of the analysis of the three metaphor test paragraphs are summarized in Table 4 below.

While there is no doubt that cohesion measures add value and richness to the measurement of text readability, the numerous indices are complex. In addition, while there are resources available to help with interpretation (e.g. the Coh-Metrix web site <http://cohmetrix.com/>, Graesser et al., 2004; McCarthy, Lewis, Dufty, & McNamara, 2006), very few of the explanations put the numeric results into context. As such, the

results may allow for direct comparisons with other texts, but it is not always clear what the numbers mean more generally.

Table 4 *Select Coh-Metrix Indices for the Metaphor Test Paragraphs*

Paragraph 1		Paragraph 2		Paragraph 3	
Sentence count	7 ³	Sentence count	6	Sentence count	6
Word count	125	Word count	173	Word count	91
Sentence length – words mean	12.5	Sentence length – words mean	28.83	Sentence length – words mean	15.17
Word length – syllables mean	1.72	Word length – syllables mean	1.75	Word length – syllables mean	1.80
Word type-token ratio	.74	Word type-token ratio	.72	Word type-token ratio	.73
Word frequency (all words, mean)	2.88	Word frequency (all words, mean)	2.90	Word frequency (all words, mean)	2.64
Polysemy mean	3.33	Polysemy mean	4.87	Polysemy mean	3.74
Flesch Reading Ease	49.31	Flesch Reading Ease	29.44	Flesch Reading Ease	38.99
Flesch-Kinkaid Grade Level	9.49	Flesch-Kinkaid Grade Level	16.32	Flesch-Kinkaid Grade Level	11.59
Coh-Metrix L2 Readability	23.02	Coh-Metrix L2 Readability	8.57	Coh-Metrix L2 Readability	3.29

One such example that is particularly relevant for this study is the Coh-Metrix L2 (second language) Readability index. Crossley, Allen and McNamara (2011) convince us that this is a more reliable measure than traditional indices like the Flesch Reading Ease and Flesch-Kinkaid Grade Level formulas. However, the latter two measures are very clear on how to interpret the results: “The output of the Flesch Reading Ease formula is a number from 0 to 100, with a higher score indicating easier reading. The average

³ The researcher noticed that Coh-Metrix calculated the number of sentences in paragraph 1 as 10; however, upon inspection, the accurate number is 7. This inaccuracy may affect other Coh-Metrix results reported.

document has a Flesch Reading Ease score between 6 and 70” (McNamara, Louwse, Cai, & Graesser, 2005, Section 11 - Readability, para 2). Similarly, we are instructed that the index that turns the Flesch Reading into a U.S. grade level, the Flesch-Kinkaid Grade Level formula, ranges from 0 to 12, and the higher the level, the more difficult the text. However, the explanation for the Coh-Metrix L2 Reading index is simply “This is the second language readability score” (McNamara et al., Section 11 - Readability, final para). Crossley et al. do explain that this measure is comprised of a combination of three indices: “a word overlap index (related to text cohesion and meaning construction), a word frequency index (related to decoding), and an index of syntactic similarity (related to parsing)” (p. 88). Some detective work reveals that more overlap facilitates reading (a higher results is better); higher frequency words are processed by readers faster and better understood (again, a higher number is more readable); and the more uniformity and consistency that exists in parallel syntactic structures, the easier a text is to process (consistent with the other two descriptors, higher is more readable) (Crossley et al., McNamara, Graesser, McCarthy, & Cai, 2014). Therefore, a high L2 Reading score equates to higher readability, but what exactly is a high score? There is no scale range provided in order to interpret the results in context.

As such, the inferences to be drawn from the results in Table 4 above are limited. Paragraph 2 is the longest judging by word count. The sentence length means are less reliable as the standard deviations ranged from 5.35 to 14.77. Paragraph 3 seems to have the longest words according to the syllable count. Word length and word frequency figures are included since Harrison (1980) tells us that vocabulary (and especially word length and word frequency) is a key predictor of text difficulty. There are subtle

differences in the results for the three texts, but the second text appears slightly more difficult. Paragraph 2 also has the highest polysemy mean, the measure that comes the closest to measuring figurative language. The last three measures indicated in the table are consistent in pointing to the second paragraph as the least readable, particularly with the grade assignment of “16.” However, the Flesch Reading Ease and the Flesch-Kincaid Grade Level are more reliable with texts over 200 words (McNamara, Louwerse, Cai, & Graesser, 2005); all of these texts are under 200 words. Although we may summarize that the second paragraph may be the least readable of the three, we do not have a real sense of the general readability of these three paragraphs.

The metaphor comprehension test was administered to the 42 students that began their engineering studies in January 2015 in the second week of classes. It was embedded in Part 3 of the SAFE diagnostic assessment. Part 1 was a computer administered and graded Academic Word List (Coxhead, 2000) vocabulary task and reading – cloze elide task. In Part 2, students completed five math questions by hand. The 3 tasks that comprised Part 3 of the assessment were: the integrated listening, reading, and writing task; the reading – multiple choice task; and the metaphor comprehension task, which was positioned as “Reading for detail.” Students had 30 minutes to complete Part 3.

Parts 2 and 3 of SAFE were graded by trained raters and determinations made about at-risk students. The results of the metaphor comprehension task were separated out for further testing and analysis since it was being piloted; the results of the metaphor test did not factor into decisions about at-risk students. The relevant biographic and demographic data from a questionnaire that accompanies SAFE, the test scores of the metaphor task, and test scores from other SAFE tasks were statistically analyzed using

SPSS version 22 software. In addition, analysis was conducted in order to determine the value of the metaphor test as a diagnostic tool.

In order to quantitatively analyze the metaphor test responses, a scale was developed to rate the responses according to degree of correctness or incorrectness. Keeping in mind the goals of diagnostic assessment, this was a criterion-referenced scale, meant to measure and make inferences on individual strengths and weaknesses, irrespective of how others perform (Bachman & Palmer, 2010). This proved to be a thorny task, but after several iterations, a 5-point scale was determined (see Table 5, below). The end scale descriptors were straightforward; completely correct and completely incorrect (unrelated response). The mid-range points were more difficult to define, and even more difficult to implement.

Table 5 *Rating Scale for Metaphor Test Task*

Code	Scale Descriptor	Alternative Description
1	Incorrect response – unrelated	Clearly no understanding (no doubt)
2	Mostly incorrect, but some part slightly related (a stretch)	Very little understanding demonstrated (major doubt)
3	Somewhat correct/seems like they are on the right track, but something is off. Difficult to judge.	Some understanding demonstrated but clearly has some trouble with the concept and/or difficult to judge whether or not concept is understood; inconclusive.
4	Mostly correct, something minor a bit off	Understanding for the most part (minor doubt)
5	Correct response	Clear understanding (no doubt)

For each of the 10 metaphoric words/phrases, a description of what was considered correct was documented, sometimes referring to a dictionary to ensure accuracy. The scale required tweaking as rating began, however. For example, a correct response for item 9 (harnessing) was anything along the lines of “collecting, harvesting, gathering or acquiring” until I found some responses that included mention of ‘using’ as well. After consulting a dictionary, I realized that there are two components to ‘harnessing:’ both ‘gathering,’ and ‘using.’ The documentation and rating was corrected to reflect this. Several rounds of rating occurred before settling on the final judgements of students’ responses.

The ratings of student responses from the metaphor test task were coded into an Excel spreadsheet. Each student was assigned a case number to maintain anonymity but responses could be attributed to a specific student. A second rater also assessed a sample of 30 responses and inter-rater reliability was calculated. We were in exact agreement 40% of the time and within one scale descriptor – and more importantly, on the same side of the scale in terms of rating the response as correct or incorrect, another 50% of the time. Therefore, we were in agreement approximately 90% of the time.

Despite criticisms of intuitively-built scales and evidence for the reliability of data-driven scales (see for example Upshur & Turner, 1995 and Fulcher, 1996), there was no alternative in this instance due to the novel nature of this test. As such, reliability statistics were calculated using SPSS version 22 software. Internal consistency was calculated on rated responses for all cases where students had attempted at least five of the ten questions (N=23). The Cronbach’s alpha coefficient was acceptable at 0.79. Statistics were also calculated to determine the effects of each of the ten test items if they

were removed from the metaphor test. The Cronbach's alpha coefficients ranged from 0.74 if item 9 (harnessing) was deleted to 0.81 if item 3 (experience gained) was deleted. The latter was the only item for which the Cronbach's alpha exceeded 0.80 if the item was deleted.

Analysis

Analytic Approach – Phase 1

A two-pronged strategy was used for the text analysis in Phase 1 of the study. The first decision related to using a corpus and corpus tools. The corpus had conveniently been developed for another study predating the current one; however, the choice to use it was due to certain benefits found in the literature. There is the obvious point that it allows a researcher to analyze a large amount of data that would be difficult or impossible to do manually (Deignan, 2005). It allows for a more in-depth and efficient (both in time and expense) means of analysis (Shutova, Teufel, & Korhonen, 2013). Related to the issue of efficiency, genre-specific corpora facilitate the ability to search the most frequent uses of language forms in specific contexts and more generally, corpus investigation provides a more objective picture of language in use rather than relying on supposition or intuition (Berber Sardinha, 2007). This latter point takes on more relevance when recalling critiques of CMT regarding its top-down approach and subjective means for metaphor identification.

For the second part of the analytic strategy two methods of analysis were used to identify metaphors at both the conceptual and linguistic level: Conceptual Metaphor Theory (CMT) and the Pragglejazz Group's (2007) Metaphor Identification Procedure

(MIP). Although CMT provides the theoretical lens for this study, it was important to take an inclusive approach to identifying metaphors in the corpus since Littlemore and Low (2006b) note that EAL learners can benefit from comprehension of both linguistic and conceptual metaphors. Other sources incorporated to assist with identifying metaphors were Lakoff's Conceptual Metaphor web site (<http://www.lang.osaka-u.ac.jp/~sugimoto/MasterMetaphorList/MetaphorHome.html>), and Renton's *Metaphorically Speaking* dictionary of idiomatic expressions.

“Metaphor Identification Procedure” or MIP is outlined as follows (Pragglejazz Group, 2007, p. 3):

1. Read the entire text–discourse to establish a general understanding of the meaning.
2. Determine the lexical units in the text–discourse.
3. (a) For each lexical unit in the text, establish its meaning in context, that is, how it applies to an entity, relation, or attribute in the situation evoked by the text (contextual meaning). Take into account what comes before and after the lexical unit.
(b) For each lexical unit, determine if it has a more basic contemporary meaning in other contexts than the one in the given context. For our purposes, basic meanings tend to be
—More concrete [what they evoke is easier to imagine, see, hear, feel, smell, and taste];

- Related to bodily action;
- More precise (as opposed to vague);
- Historically older;

Basic meanings are not necessarily the most frequent meanings of the lexical unit.

(c) If the lexical unit has a more basic current–contemporary meaning in other contexts than the given context, decide whether the contextual meaning contrasts with the basic meaning but can be understood in comparison with it.

4. If yes, mark the lexical unit as metaphorical.

An adapted version of MIP was used as the starting point. The entire text (corpus) was not read, but the target collocations were examined in context to determine whether or not the use was metaphorical.

Phase 1 – The Corpus Analysis

The text analysis to determine the prevalence and types of metaphors in engineering texts (research question 1) was a multi-level process. It began by exploring word frequencies of the corpus as a whole in the software program AntConc version 3.4.3w, which provides word list, concordance and collocate tools (Anthony, 2014). A stop word list, or list of words not included in the word frequency count was used that consisted of 40 articles, conjunctions and pronouns (see Appendix B). The word frequency results revealed a total of 13 189 types of words. Word types are differentiated from word tokens. If a corpus contained ten instances of the word ‘cat’ and five instances of the word ‘dog’, it would be made up of two types (the words ‘cat’ and ‘dog’) and

fifteen tokens (total occurrences of all words). The first four word types of the engineering corpus represented 7.67% of the entire corpus, while the top ten represented just over 11%. The bottom of the word list, or the least frequently found word types, represented nearly three quarters of the corpus. Appearing ten times or less in the whole corpus were 9830 word types, while 4554 word types appeared only once, representing approximately 35% of the word types in the corpus. This demonstrated the specialized vocabulary used in these disciplines and make sense considering Roessingh's and Douglas' (2012) comments that advanced reading materials are "characterized by a linguistic burden because of the sheer density of low-frequency words in a given text" (p. 85). The top ten most frequent words of the entire corpus, as well as each of the five sub-disciplines are detailed in Table 6.

Table 6 *Most Frequent Word Types Found in First Year Engineering Textbooks*

Frequency	Whole Corpus	Calculus	Chemistry	Computers	Mechanics	Physics
1	of	of	of	to	of	of
2	is	is	in	of	is	is
3	in	x	is	is	to	in
4	to	in	to	in	in	to
5	x	to	are	program	force	as
6	for	f	as	for	be	for
7	as	y	for	as	fig	are
8	are	t	by	be	for	with
9	by	as	with	are	by	by
10	be	by	be	on	on	on

Since the aim was to get a sense of the prevalence and types of metaphors, an initial decision was made to focus on content words. Lam (1984) explains the dichotomy of categories that parts of speech are typically divided into as listeners or readers

construct meaning in language. Content words are generally understood as nouns, verbs, adjectives, adverbs, pronouns and non-article determiners (Brown, 1973, as cited in Lam, 1984) and function words include prepositions, conjunctions, articles, determiners and verb auxiliaries (Kimball, 1973). Duffy (2009) adds to our understanding of these categories, explaining that content words “describe concepts we can picture in our minds,” whereas function words “are words used to signal grammatical functions” (p. 71).

Clark and Clark (1977) describe content words as “the real meat of a sentence” (p. 79, as cited in Lam, 1984) and note that these words help signal what might be coming next. Harrison (1980) tells us that vocabulary is a key predictor of text difficulty, specifically word length and word frequency, and as such, is almost always included in readability formulas (as cited in Lam, 1984). Lam concludes that since function words are typically shorter and more frequent than content words, that “the density of new content words in a text is the best indicator of the difficulty of that text” (p.19). She also points out that in considering research from both first and second language learning, “there is evidence that content words and function words are processed differently and that content words are more important” (p. 28). This was the basis of my initial decision to begin my search of the corpus based on the most frequent nouns and verbs.

However, the most common word forms in the corpus were function words: prepositions, followed by forms of the verb ‘to be’ (content words). Other forms of content words were virtually absent from the top ten most frequent forms, with the exception of ‘program’ and ‘force,’ from the computers and mechanics texts respectively. Also prominent were letters used in mathematical equations (e.g. ‘x’) in the calculus texts

and ‘fig’ (short for figure) in the mechanics texts. After initial exploration into content words from the top 100 most frequent word types, the decision was made to search the top 100 collocates of the verb ‘is’ and the preposition ‘of.’ In the context of corpus analysis, Deignan (2005) defines collocations as the phenomenon when “two or more words regularly appear in each others’ environment” (p. 79).

From this survey, a list of 24 collocates of ‘is’ and ‘of,’ combined with 21 of the most frequent nouns and verbs were the focus of the final sampling of concordances that were analyzed, to represent the most frequent collocations that students would encounter in their reading material. Stratified random sampling (Dornyei, 2007) was used to ensure that each of the engineering sub-discipline subjects in the corpus were examined in proportion to their sizes. In essence, a proportionate sample was drawn at random intervals to ensure that all five sub-disciplines (or strata) were reached. Also considered was whether the collocations were typically formed on the left or right side. Sometimes this was clear cut; sometimes collocations occurred on both sides of the word being examined, and this was incorporated into the sampling strategy. Table 7 summarizes the final collocations that were sampled for metaphoric content within the corpus.

Table 7 *Final Collocates Sampled for Metaphoric Content*

Calculus	Chemistry	Computers	Mechanics	Physics
Is called	Reaction is	Number is	Is shown	Field is
Function is	Is called	Is executed	Force is	Is given
Value(s) of	Number of	Program is	Action of	Number of
Function of	Energy of	Value of	Of action	Of energy
	Of energy		Moment of	Energy of
				Value of

It is important to note that sometimes it was straightforward to judge whether or not metaphor was being used, and other times it was less clear due to considering different senses of words and the use of terminology in subject-specific contexts. One example of confusion caused by the subject-specific nature of the text was with the ‘function is’ collocation. There were many instances of ‘rational function’ in the calculus portion of the corpus, which at first glance seemed metaphoric. However, the concept ‘rational function’ has specialized meaning in this context as a mathematical concept. The naming of this concept as ‘rational’ may still constitute metaphoric use of the word; however, someone with both linguistic and subject knowledge could better judge.

In other instances, metaphoric content was found not in the target collocation, but in the surrounding text. An example from the chemistry portion of the text is: “...strength of an acid depends on a number of factors.” The use of ‘strength (of an acid) depends’ is metaphoric, but not the collocation being examined, which was ‘number of.’ Similarly, in another example from chemistry, “Intermolecular forces give rise to a number of structural features...” the metaphor lies not in the collocation ‘number of,’ but in the text “(intermolecular) forces give rise.”

Phase 2

There were a number of analyses that took place in Phase 2 of the study. The first step was to look all of the responses to see what students did and did not attempt. Eleven students provided no response to any part of the test, and eleven attempted an explanation of all ten metaphoric expressions; the rest of the students fell in between. Had every student attempted an explanation for all ten metaphoric words/phrases, this would have

amounted to 420 responses to analyze. However, as demonstrated in Figure 3 below, not all students provided responses to all of the questions. There were a total of 209 (N=209) response attempts to analyze.

The next step was to analyze the ratings of the responses to answer the second research question as to whether metaphor affects reading comprehension of engineering texts. Descriptive statistics provided a picture of how students performed on each question and on the metaphor comprehension task overall. Because the metaphor test was piloted in the SAFE assessment, item facility and item discrimination analyses were undertaken. Item facility (IF) looks at the percentage of students who are rated as giving a correct response, and Item Discrimination (ID) analysis, considers “the degree to which

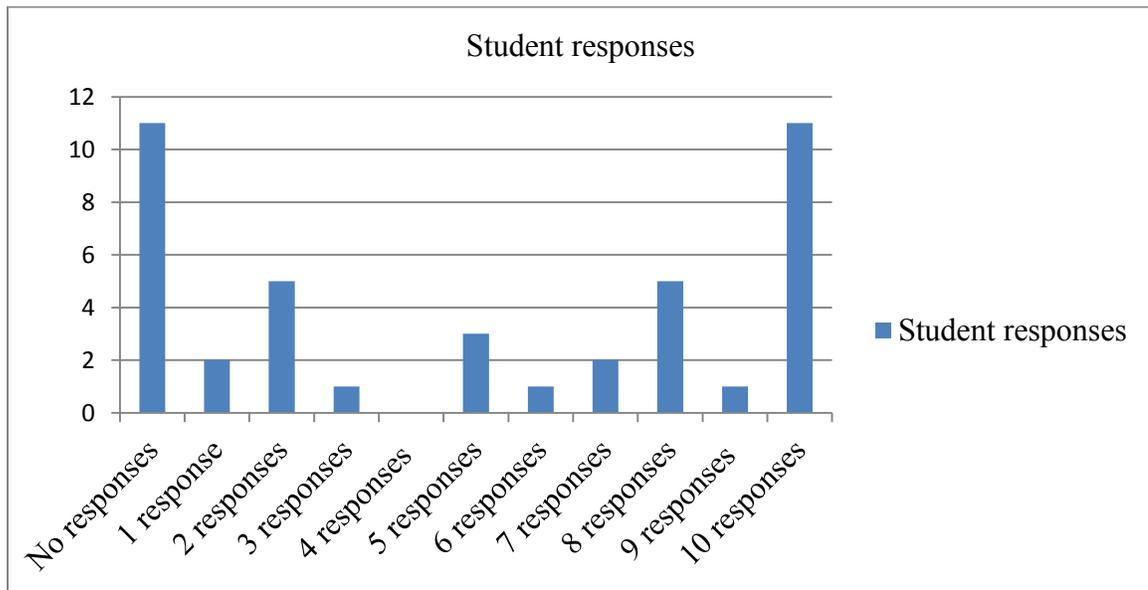


Figure 3 Student Responses by Test Item.

an item separates the students who performed well from those who performed poorly” (Brown, 1996, p. 66). As Brown explains, this is done by comparing the IF scores of the

top group of test performers with those of the lowest group. These groups are determined by taking the top and bottom third of students, although it does not always work out so neatly; the number can range as low as 25% (Brown, 1996). As previously mentioned, eleven of the 42 students who wrote the diagnostic assessment did not attempt any part of the metaphor comprehension task. Eight more students attempted between one and three questions only (no one responded to only four questions); their results were not included in the statistical analysis. This left the responses of 23 test takers who attempted at least half of the ten questions. As such, the top and bottom groups were comprised of approximately 30% of this group of 23 students (N=7 for each of the top and bottom groups).

The next stage of analysis involved comparing student responses and their first language, in order to investigate research question 3 regarding whether a student's first language is a factor in metaphor comprehension. A survey of individual written responses was undertaken to determine those that were examples of literal interpretations. This was followed by an examination of the results collectively to see the distribution of test scores by first language.

The final stage of analysis was to investigate the effectiveness of the metaphor task as a diagnostic tool. Statistically speaking, the five points of the scale used to rate student responses did not truly represent interval numbers. As such, non-parametric testing was undertaken (Dornyei, 2007). Spearman's rho was calculated to consider the relationship between the metaphor task and other tasks comprising the SAFE assessment. In addition, simple linear regression analysis was calculated to investigate the extent to which the results on the metaphor test could predict variance in the results from other,

related SAFE tasks (e.g. vocabulary, reading – cloze elide, reading – multiple choice).

The IF and ID analyses mentioned above also attested to the effectiveness of the metaphor task as a diagnostic tool.

Chapter Summary

This chapter provided details regarding the research design, participants, instruments, procedures and analysis that took place during both phases of the study. In Phase 1, a corpus of first-year engineering reading material, made up of texts from the sub-disciplines of calculus, chemistry, computers, mechanics and physics was qualitatively analyzed to get a sense for the prevalence and types of metaphoric content.

Using the Phase 1 findings and considering readability factors, a metaphor comprehension task was designed in Phase 2 and administered to 42 new incoming students in January, 2015. The majority of these test takers were international students, in Canada for one year or less, and almost exclusively multilingual. A 5-point rating scale was developed to rank the responses according to degree of correctness or incorrectness, and this scale was statistically analyzed, due to it being intuitively developed. A second rater also assessed a portion of sample responses and inter-rater reliability was determined to be satisfactory.

The students' responses were quantitatively analyzed using descriptive and inferential statistics. The ratings and individual test prompts (metaphoric words and phrases) were analyzed using item facility and item discrimination measures. Students' responses were also analyzed in comparison to their first language, both at the individual

and group levels. Finally, the metaphor task as a whole was analyzed using correlations and simple linear regression to determine its effectiveness as a diagnostic tool.

In the next chapter, the findings will be reviewed in sequence following the two-phase sequential exploratory design (Plano Clark & Creswell, 2008). These results will be discussed in view of the three research questions that guided the study, the theoretical framework, and previous research. Commentary will also be provided on the effectiveness and usefulness of the metaphor task as a potential tool to diagnose at-risk university students.

Chapter 5: Findings and Discussion

The thesis to this point has provided background and context to this investigation and detailed the heart of the study: the research design, participants, instruments, procedures and analysis. This chapter will synthesize the findings from both phases of the study. This includes the qualitative results of the corpus analysis, which informed Phase 2 of the study. The results of the analysis of the students' responses to the metaphor task are presented, as well as the statistical analysis of the metaphor task as a diagnostic tool.

The results will be discussed in light of the three research questions that guided this investigation and existing literature. In addition, the findings will be considered from the point of view of the study's theoretical framework. Finally, the effectiveness and usefulness of the metaphor task will be discussed according to the statistical analysis and Bachman and Palmer's (2010) test usefulness criteria.

In the final chapter that follows, implications will be discussed, including those for EAL and EAP pedagogy, post-entry diagnostic assessment in the post-secondary context, and supports for at-risk students to improve their chances for success and promote student retention. The limitations of the study will be outlined along with directions for future investigation.

Phase 1

Overall, there was metaphoric content found in every sub-discipline of the corpus, whether directly in the collocations examined, or in the surrounding text. This was to be expected and in line with the findings of Steen et al. (2010) and Dorst (2015). However,

the extent to which metaphor was present in each of the sub-disciplines was not consistent. The least amount of metaphoric content was found in the mechanics text, while the greatest concentration of metaphoric content was observed in the chemistry portion of the corpus. This was taken into account when designing the metaphor test task in Phase 2. The most common trends identified were personification (ontological metaphors), and metaphors relating to family and relationships. Table 8 outlines examples of metaphoric text from each of the five sub-disciplines of the corpus. The italicized text identifies collocations that were examined.

Table 8 *Examples of Metaphoric Content in the Engineering Corpus*

Calculus	Chemistry	Computers	Mechanics	Physics
...each <i>value of</i> the constant gives rise to...	One of the promising candidates <i>is called</i> ... ⁴	... <i>program is</i> translated by the computer...	...a <i>force is</i> completely characterized...	...a magnetic <i>field is</i> set up whose strength is proportional...
...the table <i>of values</i> suggests...	All forms <i>of energy</i> are capable of doing work...	Line 3 now calculates and stores the <i>value of</i> the variable...	The <i>moment of</i> a couple can also be expressed by the vector...	The conservation <i>of energy</i> is a powerful tool for solving problems.

There were two factors that seemed to affect the extent of metaphoric content, at least in the collocations examined. The first was the use of passive and active voice. In the mechanics text, one of the collocations examined was ‘is shown.’ No metaphoric instances were found because this constitutes a passive construction and the use of this

⁴ In the context of this statement ‘candidates’ referred to “CFC substitute” and not a person.

collocation was almost exclusively "...is shown in Figure...", such as in the example "...diagram of the pulley is shown in Figure..." Instances where metaphor was more prevalent were those where the verb was in active voice. For example, had the above statement been phrased as "Figure 2 shows that...", this use would be metaphoric (personification) since 'Figure 2' is not a physical being that can engage in the act of showing. There were many instances of this sort of metaphoric use in the corpus; however, specific collocations of 'figure' that would demonstrate this use were not formally examined.

Passive voice did not automatically result in literal versus figurative language, though. Examples in Table 7 from chemistry (one of the candidates is called) and both examples from mechanics (force is characterized; moment can be expressed) are all passive verb forms used metaphorically. More specifically, these constructions constitute active verbs with inanimate subjects. According to Master (1991) this construction is common in scientific text and causes great difficulty for some EAL learners. As such, there would be value in exploring it further.

The second factor affecting metaphoric use was context or subject-specific collocations. In the chemistry texts, the words 'reaction is' followed by a formula were common. In the mechanics texts, 'action of/of action' turned out to be part of a larger collocation, "line of action of" in almost every instance. Also from mechanics, 'moment of' collocated often as "the moment of a/the force," and "moment of inertia." None of these collocations were used metaphorically in the text sampled, but this would be another area for further research.

These results provide insight into the first research question: “To what extent are metaphors prevalent in textbooks used in first year engineering courses?” This question was answered, albeit weakly, since there was no attempt to quantify the metaphors that were found in the sampled parts of the corpus. The fact that metaphoric content was found across the corpus was to be expected in light of Steen et al.’s (2010) and Dorst’s (2015) findings. However, this may still be a surprising result since native speakers do not consciously think about producing and understanding the metaphors in their language. One might make assumptions about the formal nature of academic writing, especially in the STEM (science, technology, engineering and mathematics) subjects, underestimating the extent of metaphoric content in this genre.

Phase 2

The qualitative analysis results from Phase 1 facilitated the comprehension test designed in Phase 2 to be context-driven and authentically representative of metaphors that students may actually encounter in their first-year reading. The findings of the analysis of students’ interpretations of the metaphoric words and phrases are now presented.

Of the 31 students who attempted at least one response, Figure 4 (below) shows the overall distribution of students’ scores out of a possible 50, which was obtained from a rating between one (no understanding) and five (complete understanding) for each of the ten metaphoric words and phrases that were tested. As evidenced in the histogram, the scores were not distributed normally; a bi-modal distribution is revealed, with two groups at either end of the scale. There are distinct groups of high and low achievers on this task,

or those who did and did not have an understanding of the metaphoric words and phrases. This may be explained by the disproportionate number of EAL students in this cohort. In addition, the small sample size likely plays a role.

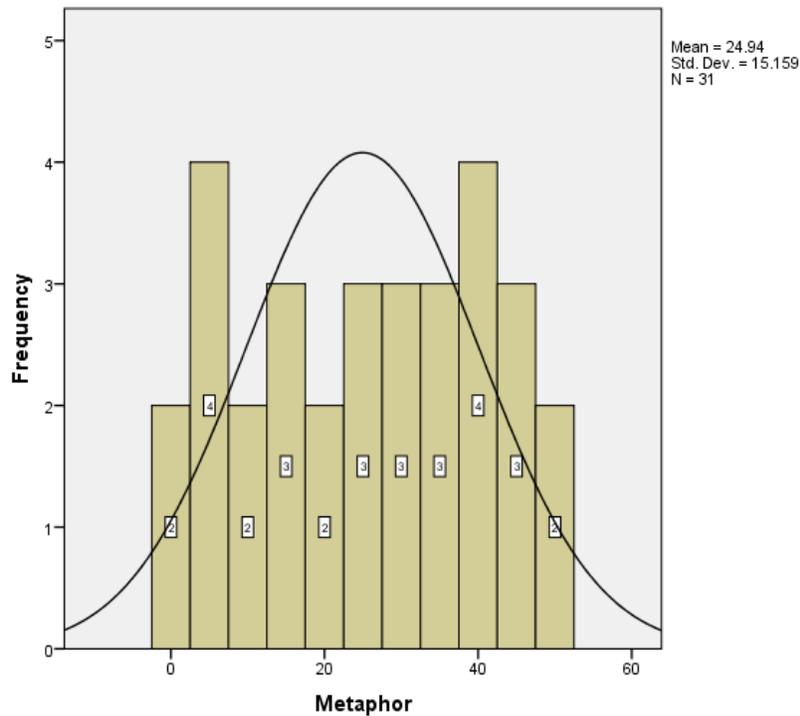


Figure 4 Distribution of Student Metaphor Test Scores.

Whereas Figure 4 demonstrates the distribution of students' scores on the metaphor task in its entirety, Table 9 (below) provides a breakdown of each of the ten test items individually, including how student responses were rated, and which questions were or were not answered. As should be expected, students had varying degrees of success in the responses they attempted. If we consider a rating of 4 or 5 as generally correct, while ratings of 1 – 3 as incorrect, it appears that students had the most success

with item 2 (chemical connection). This item also had the least number of skipped attempts. Items 1 (introductory level), 3 (experience gained) and 6 (led to a dead end) also seemed fairly straightforward for students. Item 9 (harnessing) had the least number of '5' ratings; however, the majority of those who attempted it were still rated as giving correct responses. Finally, although the last four questions had the most nil responses,

Table 9 *Student Responses and Ratings*

Question	Responses per Scale Descriptor					Total Responses	Nil Responses
	1	2	3	4	5		
1	1	3	5	1	17	27	15
2	0	3	1	7	18	29	13
3	3	2	3	6	10	24	18
4	9	2	1	1	9	22	20
5	3	2	4	4	8	21	21
6	2	1	1	15	4	23	19
7	2	3	2	4	5	16	26
8	1	0	3	7	6	17	25
9	1	1	0	11	1	14	28
10	1	0	2	4	9	16	26
Totals	23	17	22	60	87	209	211

they were at the end of the test. We can't automatically assume that these presented the greatest comprehension challenges; some students may not have been able to complete the task in the allotted time.

The IF and ID results presented in Table 10 below provide more insight into how the test takers fared on each one of the ten metaphoric words and phrases. The group of test takers who attempted at least half of the questions (N = 23) performed the best on items 2 and 9. The latter comes as a surprise since it initially appeared to have caused

some difficulty. However, the difficulty appears to have occurred only for the poor-scoring test takers. The IF scores also confirm that item 4 proved difficult, but again, only for the low achieving group. The latter group also had difficulty with items 5 and 7. Item 7 did the best job of discriminating between the strong and weak performers, with items 4, 5 and 9 also doing a good job of this. Item 3 did the worst job of differentiating between high and low achieving groups, and item 2 also did a poor job in this area.

Table 10 *Item Facility and Item Discrimination for the 10 Test Words and Phrases*

Question Number	Item Facility – All	Item Facility – High Group	Item Facility – Low Group	Item Discrimination
1	.67	.86	.29	.57
2	.86	1.0	.71	.29
3	.67	.71	.57	.14
4	.45	1.0	.14	.86
5	.57	1.0	.14	.86
6	.83	1.0	.57	.43
7	.56	1.0	0	1.0
8	.76	.86	.29	.57
9	.86	.86	0	.86
10	.81	.86	.29	.57

The next step of the analysis was to consider the test results in comparison with students' first language. Looking at the all of the individual responses, there were not a great number that were obvious examples of literal interpretations. However, item 4, 'employ,' evoked the most number of literal responses. Examples included "worker," "give work to," and "apply to" instead of the non-literal meaning of 'use,' which was considered correct. Other examples of literal responses included "taking out" for item 5 (carrying out), "use whole life in project" for item 6 (led to a dead end) and "learning

people” for item 7 (body of knowledge). Important to note is that the students who provided these literal interpretations were all non-native speakers of English with the exception of one student who identified English as the first language, but also identified a different language spoken at home. One other student who provided a literal response did not submit the accompanying questionnaire to the SAFE test; the first language is unknown. EAL students making literal interpretations of the metaphoric language is consistent with Danesi’s (1995, p. 4) “textbook literalness” classification and the results of Gunderson, Slade and Rosenke’s (1988) study.

Further investigation into the differences between the native and non-native English speakers was more revealing. Figures 5 and 6 below provide insight into how students fared on the metaphor comprehension task according to their first language. There were only 2 native French speakers, so these results are inconclusive. However, the divide between English first language speakers and those who learned it as an additional language is clear.

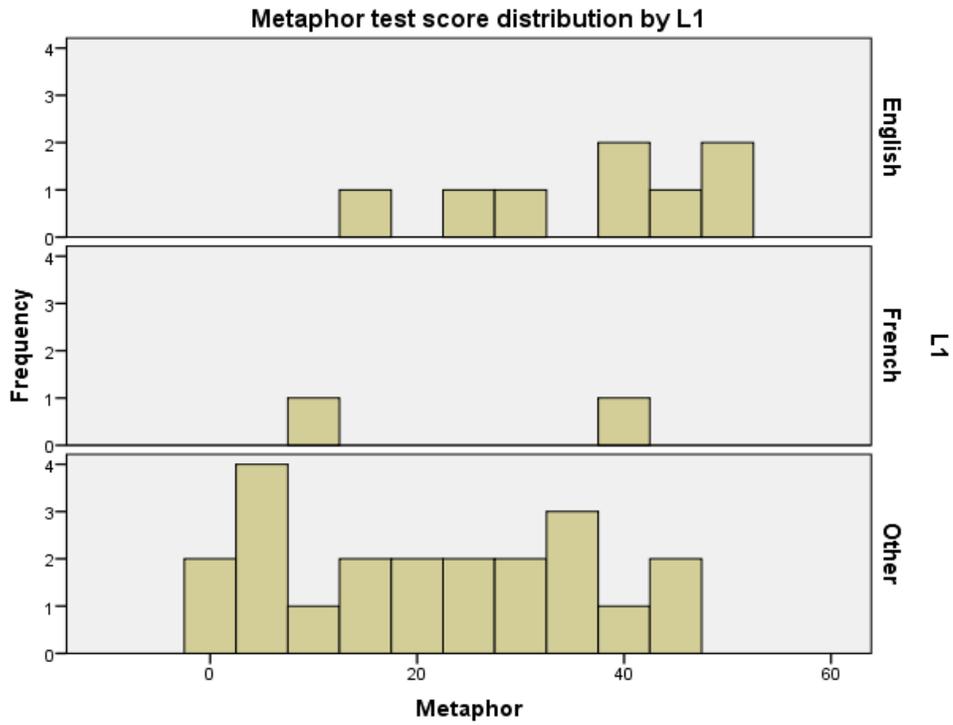


Figure 5 Metaphor Test Score Distribution by L1.

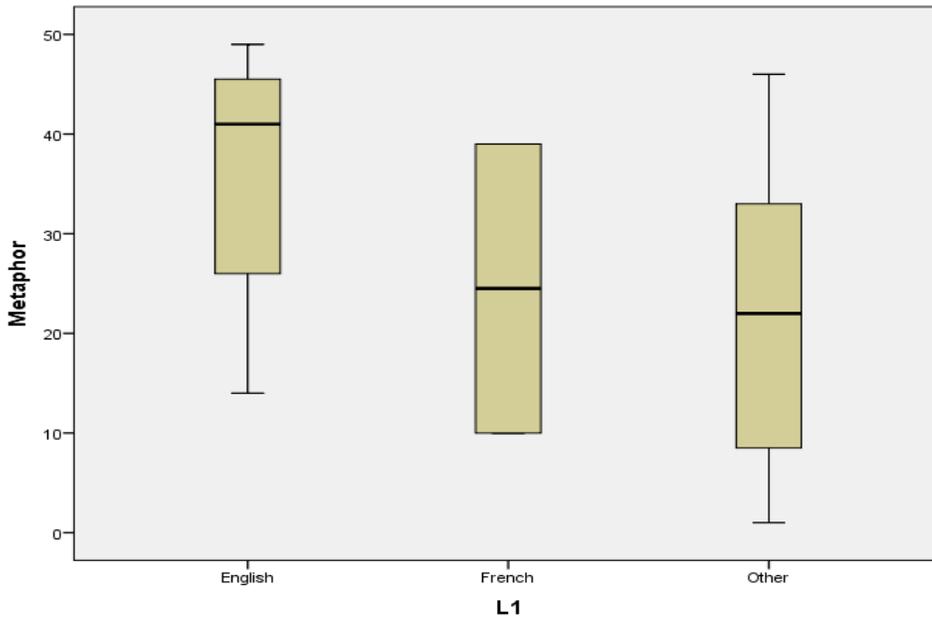


Figure 6 Box Plot of Metaphor Test Scores by L1.

In light of these results, the second and third research questions are better considered together. The second question “Is metaphor a factor in reading comprehension of engineering textbooks,” seems to depend on whether the test taker’s first language is English or another language. Some words/phrases were more easily defined than others and some students did better on the metaphor task than others, but ultimately, the students who had difficulties with certain metaphors and/or the task as a whole were primarily those whose first language was not English. What the results also showed is that the nature of this difficulty was based less on differentiating literal from figurative language than on misinterpreting the metaphoric words and phrases. This is evidenced by the small number of literal responses versus the overall scores. The item facility scores for the whole group of 23 students who attempted at least half of the test questions showed items that gave students the most difficulty: 4, 5 and 7. Even some of the high-achievers struggled with these words/phrases, although their overall results were not significantly impacted.

We can then conclude that the third research question, “Is a student’s first language a factor in comprehending English metaphors,” is the question most confidently answered by the results obtained in this study. Figures 5 and 6 showed a divide in how the English L1 students did as compared to students whose first language is a foreign language. In addition, all responses that were literal interpretations (and as such, incorrect) were provided by EAL students, which is consistent with existing research into metaphor comprehension. If we recall the results of Webster’s 2014 survey of faculty members’ perceptions of student strengths and weaknesses, they identified international student reading comprehension as a problematic area. The findings of this study suggest

that this may in fact be the case, particularly as relates more specifically to the comprehension of metaphoric words and phrases in reading materials.

The Results Considered Through the Lens of Conceptual Metaphor Theory

The two principles of CMT that shed the most light on this study are the systematicity and consistency in which conceptual metaphors are used, and the cultural-specificity of metaphors, or more specifically their referents. The systematicity and consistency mean that there are patterned ways in which we use language to talk or write about certain things. At first glance, this may seem helpful for EAL learners. However, these patterns may not be apparent or may be confusing if metaphoric referents are nonexistent or different in a learner's language and culture. The way we conceptualize things is very much dependent on how we are socialized within our culture. Considering the unconscious nature of how native speakers systematically produce and understand conventional metaphoric words and phrases, as we have grown to understand them through our cultural lens, it is not surprising that those learning English as an additional language would struggle to pick up these nuances. Even the novel metaphors invented by native speakers can be done so with automaticity, because of the cultural and background knowledge that can be summoned. Learning English as an additional language does not typically afford learners the same cultural exposure as native speakers receive when learning their mother tongue, although advances in technology and global exchange are attempting to lessen this gap.

The cultural specificity of metaphors has further implications for the background knowledge or schemata that students need to understand some metaphoric referents.

Although there are differences between cultural background and other forms of background knowledge that affects students' schemata that need to be teased apart, there is evidence in the literature relating to cultural background knowledge. One of the findings of Meissner's (2010) textbook study was that students had difficulty with metaphors that were not common in their language and culture. The results of Steffensen et al. (1979), and Floyd and Carrell (1987) also provided insight into the difficulties caused by a lack of shared cultural knowledge. We cannot take for granted the extent to which inadequate cultural and background knowledge may impede comprehension for international students, particularly those whose first language is not English.

The Metaphor Task as a Diagnostic Tool

Two last groups of statistics were calculated in order to compare the metaphor test results with the results of the other SAFE tasks. Spearman's rho confirmed significant, positive correlations with all parts of SAFE except the Math task: the metaphor test correlates with the writing task $r = .44$, $p < .014$; with the vocabulary task $r = .67$, $p < .001$; with the multiple choice (reading comprehension) task $r = .59$, $p < .001$; and with the cloze elide (reading comprehension) $r = .72$, $p < .001$. Although not the strongest correlation, the metaphor task results were specifically compared to the writing scores as these have proven to be the most dependable indicator in the SAFE diagnostic for predicting at-risk students (J. Fox, personal communication, March 23, 2015). As such, a significant correlation of the metaphor scores with the writing scores could speak to the potential of the metaphor task in diagnosing at-risk students. Figure 7 shows that the

students who did poorly on the writing task and were diagnosed as “at risk” also did not do well on the metaphor task.

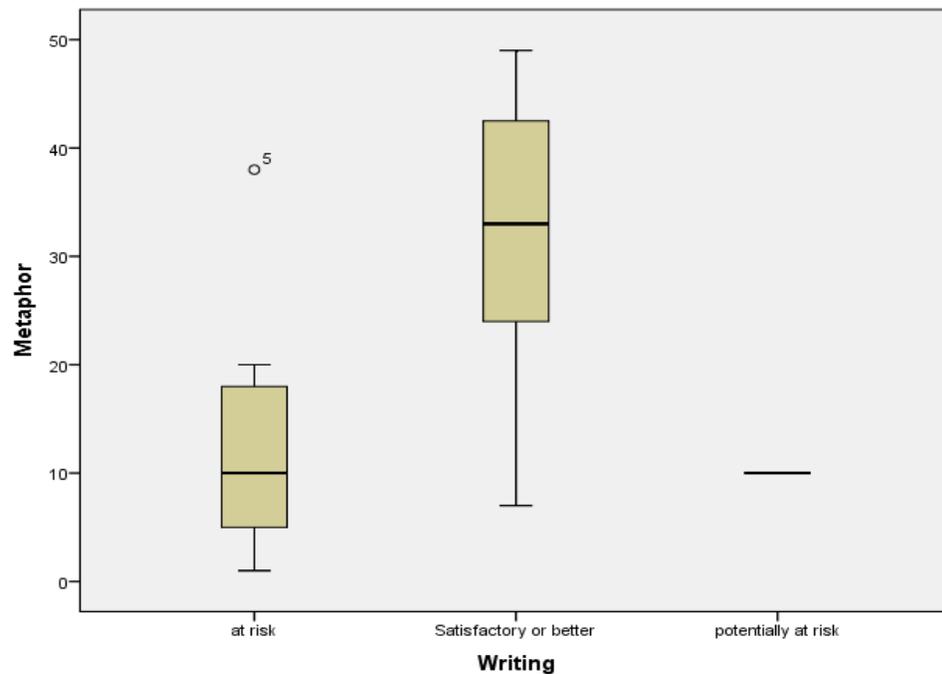


Figure 7 Comparison of Metaphor Test Results with SAFE Results.

However, as Figure 7 also shows, there is one outlier (Case 5), which is a puzzling case. Having just claimed that writing is a good predictor of academic success, this student appears to have done poorly on the writing task, but performed well on the metaphor task (38/50) – much better than her peers who did poorly on the writing portion. Upon closer inspection, this student did poorly on the writing task, vocabulary and multiple-choice reading, but performed well on the metaphor and cloze elide tasks, the latter also being traditionally difficult for EAL students (J. Fox, personal communication, March 23, 2015). The specifics of this case require further investigation;

however, one additional comment can be made. A look at this student's metaphor test responses revealed that although she performed well on this task overall, she did provide one literal response. Considering the low number of literal responses provided by all students, this is significant in demonstrating the persistent difficulties EAL learners have with English metaphor.

The final query was simple linear regression analyses to investigate the extent to which the variance in the overall scores of the vocabulary, reading - cloze elide, and reading - multiple choice tasks was accounted for by the metaphor task scores. As evidenced in Table 11, the metaphor test accounts for the most variance in the reading – cloze elide task at approximately 55% ($\beta = .74, p < .00$). Approximately 42% of the variance of the vocabulary task is accounted by the metaphor task ($\beta = .65, p < .00$) and the least amount of variance accounted for is in the reading – multiple choice task at approximately 20% ($\beta = .45, p < .03$). All of the beta results are significant, as are all of the model results with the exception of the cloze elide task, which may possibly be explained in part by the small number of students who completed the test. The regression findings suggest that the metaphor test results are good predictors of results on the vocabulary and reading tasks. In addition, the construct of metaphor comprehension as measured in this task shares similarities with the constructs of these three tasks as measured, although to a lesser extent with the cloze elide task. This may be due to similarities in the skills that these tasks measure, or an imprecise conception of the construct of metaphor comprehension; further empirical evidence would help to tease a part these skills and better define the constructs involved.

Table 11 *Simple Linear Regression: Metaphor Predicting Variance on Other SAFE Tasks*

Model	Unstandardized Coefficients		Standardized Coefficients	t	Significance
	B	Standard Error	Beta		
Dependent Variable: Vocabulary					
(Constant)	17.946	1.078		16.648	.000
Metaphor	.169	.037	.646	4.556	.000
Dependent Variable: Reading – Cloze Elide					
(Constant)	7.783	4.538		1716	.097
Metaphor	.930	.156	.742	5956	.000
Dependent Variable: Reading – Multiple Choice					
(Constant)	3.952	1.303		3033	.006
Metaphor	.092	.038	.452	2385	.027

Since the metaphor task was newly created, analysis was crucial in order to determine its effectiveness and usefulness as a diagnostic tool. Although testing conventions were followed where possible, this was a novel assessment that was constructed using as much intuition as empirical backing. Despite having to break new ground, the test performed reasonably well, and did what it was designed to do: differentiate between those who do and do not understand metaphoric words and phrases in reading material. There were a number of measures that confirmed this. The metaphor test positively and significantly correlated with all other tasks in SAFE with the exception of an unrelated mathematics task. The students who were diagnosed as at-risk according to (amongst other things) their scores on the writing task, also performed poorly on the metaphor test. And, simple linear regression showed that the metaphor test could predict

with significance anywhere from approximately 20% - 55% of the variance in how students would perform on the vocabulary and other reading tasks.

Item facility and item discrimination measures were helpful in determining the effectiveness of each of the ten metaphoric words and phrases in diagnosing students' comprehension of English metaphor. According to Ebel's (1979, p. 267, as cited in Brown, 1996) guidelines for ID scores, item 3 ('experience gained') should be eliminated or altered, item 2 ('chemical connection') is borderline acceptable and should be improved, and all other items are "very good" since their ID is above 4.0. These results are encouraging for the test pilot.

There were other, more practical issues with the test as a whole. Many students did not attempt some, or all, of the metaphor portion of the assessment. As Dornyei (2007) points out, we do not know if missing responses are meaningful, accidental or deliberate. The metaphor task was the final portion of a 3-part assessment, and Part 3 was only allocated 30 minutes to complete. Perhaps more time was required; however, SAFE is done during valuable class time. It may be that the time allotted for this assessment cannot be adjusted, in which case, the tasks must be adjusted to fit the timeframe. Further piloting also needs to take place to determine whether students understood the test instructions, or what was expected of them.

Bachman and Palmer's Usefulness Criteria

Bachman and Palmer (2010) insist that "The most important consideration in designing and developing a language test is the use for which it is intended, so that the most important quality of a test is its usefulness" (p. 17). The six criteria upon which they

gauge usefulness are briefly explained. Bachman and Palmer define ‘reliability’ as consistency in test results. ‘Construct validity’ is found when the scores are representative of the real-life task being measured. In other words, the test measures what it intends to measure and students’ results are representative of this. They explain ‘authenticity’ as the test tasks or what students are asked to do being characteristic of what they would do with the language being tested in real life. ‘Interactiveness’ is defined by Bachman and Palmer as the extent to which a test taker’s language skills, as opposed to other, unrelated skills, are used in completing the test task. ‘Impact’ is the influence of a test at the micro, personal level, and macro level within a society and its education system. And finally, ‘practicality’ represents the extent to which the test can be implemented with available resources (Bachman & Palmer, 2010).

There was some evidence for construct validity. The test was demonstrated to measure what it intended to measure, which is students’ ability to understand English metaphor. The bi-modal distribution of student scores delineated separate groups of low and high achievers, as did the item facility and item discrimination results with only two exceptions. However, questions remain regarding possible overlap in the constructs of metaphor comprehension, reading comprehension and vocabulary comprehension that should be further explored. Arguments can also be made for authenticity and interactiveness. Students read authentic text that came directly from first year engineering reading materials and the metaphors that were tested were naturally occurring and embedded in this text. However, it is questionable whether students would authentically write troublesome vocabulary in their own words as a strategy in reading comprehension. Similarly, in terms of interactiveness, students had to draw upon their own reading skills

to interpret the meaning of the metaphors in the given context in the same manner they would have to interpret course readings.

There was some evidence of reliability, although this is one area where further inquiry is necessary. There was evidence of strong links between the metaphor and other test tasks, including strong correlations with the vocabulary, reading – cloze elide, and reading – multiple choice tasks demonstrating overlap in these skills, and the regression analyses that showed that the metaphor test was a good predictor of how students would perform on the related tasks. However, a true test of reliability measures test tasks that are virtually identical (Bachman & Palmer, 2010), not just correlated or good predictors of variance. Stronger evidence would come, for example, from the same students rewriting the test at a later date, or them performing similarly on another version of the test that was determined equal to this version.

Similarly, impact and practicality are more difficult to determine at this juncture. Too little is known about this test and how it functions to determine its effect on students, their instructors, or other stakeholders such as university policy makers or EAL/EAP instructors. The hope, of course, is that this study has positive washback effects. Washback is defined by Alderson and Wall (1999) as the effects (positive or negative) of testing on the teaching and learning of language. We sometimes see this manifesting as “teaching to the test,” particularly when it comes to preparing students for English proficiency tests, university entrance/exit exams, and other high stakes tests (see for example Hamp-Lyons, 2007; Pan, & Newfields, 2011). Teaching to the test is often viewed negatively as it means teachers diverting class time to focus on skills needed to pass a test as the priority over building communicative competence. However, the effects

of washback are not always negative. Both Cheng (2008) and Green (2013) in reviews of testing literature relating to washback comment on studies that investigate washback as a change agent and suggest that more attention be paid to test design in order to elicit positive change or washback, while minimizing or avoiding the detrimental effects.

In terms of practicality, there was no cost involved other than the time to comb the corpus for readable passages to use. It was a relatively simple process to embed the metaphor test into the existing SAFE diagnostic; however, more research is required to determine why so many students did not attempt part, or all, of the metaphor comprehension task. Further piloting of this test in isolation would provide evidence on the amount of time it should reasonably take to complete the task – in addition to providing evidence for or against reliability.

Bachman and Palmer note that the usefulness of a test should not be judged by any one of the six criteria, but on the combination of the criteria – how they operate in unison, and in a given testing context. As with any test, particularly a new one, there were things that could be changed or improved in order to increase the effectiveness and usefulness of the metaphor comprehension task. In addition to the two test items revealed as weak in the ID results, the correlation and regression results indicated a certain amount of overlap in some of the tasks. As such, further investigation is needed to determine the best combination of tasks to most accurately diagnose students' academic skills. If the metaphor test is to be one of these tasks, some items would have to be changed or improved; namely items 2 and 3 ('chemical connection' and 'experience gained'). However, with adjustments and further testing, the metaphor task has potential as a diagnostic tool.

Chapter Summary

The findings presented here demonstrated that every sub-discipline contained metaphoric words and expressions, but some to a greater extent than others. This is not surprising considering the ubiquity of metaphor in English and prior corpus-based investigations; however, native speakers of a language do not consciously think about using and interpreting metaphor and may find surprising how common metaphor is in English academic writing.

Two principle findings were used in the building of the metaphor comprehension task administered in Phase 2: the Chemistry and Physics portions of the corpus contained the greatest concentration of metaphors, and personification and metaphors relating to family and relationships were the greatest patterns found. Not surprisingly, the students' whose first language was not English did not fare as well as their native English-speaking classmates, which is consistent with past research. All literal interpretations of the metaphoric test words and phrases stemmed from EAL students. In addition, those who did not do well on the metaphor comprehension task did poorly on the writing task (with one exception), which has proven a strong indicator of potential academic risk.

The metaphor test proved to be positively and significantly correlated to the other test tasks with the exception of a mathematics task. Regression analysis demonstrated that the metaphor test scores were significant and positive predictors of the vocabulary, reading – cloze elide, and reading – multiple choice task scores. Item facility and item discrimination analysis demonstrated that each of the ten test items, with the exception of 'experience gained' and possibly 'chemical connection' worked well to differentiate those who did poorly and those who did well on the metaphor test as a whole. These

statistical analyses and a review of the test in light of Bachman and Palmer's (2010) test usefulness criteria illustrate that the metaphor task has potential as a diagnostic tool.

The final chapter of this thesis will consider implications for EAL and EAP pedagogy and post-entry diagnostic assessment in the post-secondary context. The study's limitations will be summarized and directions for future investigation explored.

Chapter 6: Conclusion

Thesis Summary

The idea for this research grew from reading Lakoff and Johnson's *Metaphors We Live By* and investigating metaphors in TOEFL preparation reading materials in an undergraduate course. It evolved into a two-part study, bringing together inquiry on metaphor comprehension, reading comprehension and diagnostic assessment. In the first part of the study, a corpus of engineering reading materials was examined to get a broad sense of the quantity and types of metaphors present. The results from this phase of the study were used, along with readability considerations, to create a metaphor comprehension test that was embedded into an existing diagnostic assessment (Self-Assessment for Engineering or "SAFE") that all first year engineering students at this Canadian university write at the beginning of their program. The larger diagnostic examines subject knowledge to a small extent, and to a larger extent, language skills: integrated listening, writing and reading; vocabulary; and reading comprehension skills. The purpose of SAFE is to diagnose potentially at-risk students early in their program, and in a timely fashion so that intervention can facilitate student success and retention.

The results of the students' responses to the metaphor comprehension test, their own written interpretations of ten metaphoric words and expressions that were underlined in three paragraphs of text taken from the chemistry portion of the corpus, were statistically analyzed. The findings supported to a small extent previous research that demonstrated that EAL learners gravitate towards literal interpretations over figurative ones. To a stronger degree, further evidence was provided demonstrating the difficulties EAL students have with English metaphor comprehension. The native English speaking

students out-performed their EAL classmates, and the students who were diagnosed as at-risk according to their test results on the other SAFE tasks also performed poorly on the metaphor comprehension task. Evidence of the effectiveness and usefulness of the metaphor task as a diagnostic tool came in the form of item facility, item differential, correlation, and regression measures, as well as assessing the metaphor comprehension task according to Bachman and Palmer's (2010) usefulness criteria.

Implications of the Study

There are a number of implications that can be drawn from both phases of this study. The first relates to the metaphoric content found in the corpus. It is important for EAL, EAP and post-secondary instructors to be aware themselves, and raise students' awareness about metaphoric content they may encounter in what they read, along with the fact that this figurative language content may impede their reading comprehension. Drawing attention to figurative uses of words and expressions so that students are able to differentiate between literal and figurative language is a good first step, but it appears that focus on interpreting metaphoric terminology is also necessary. This may be in part, due to cultural differences in the use of metaphors and metaphoric referents.

Although a body of research exists on the importance of metaphoric competence and the difficulties EAL learners have mastering English metaphor, it still does not rank as a priority in most language classrooms. This study provides further evidence of how crucial it is for teachers preparing EAL students for post-secondary study in English-medium institutions to incorporate metaphor comprehension into their curriculum. Although it may be overly optimistic to hope for positive washback from this study since

this call is not a new one (see for example Roessingh, 2005), including metaphor comprehension in EAL curriculum would assist students to do better on assessments by virtue of improving their reading comprehension skills, and as argued in the literature (e.g. Littlemore & Low, 2006b), metaphor competence increases overall communicative competence.

EAL and EAP instructors can also use authentic reading materials that students may encounter and draw beyond the most common 2000 word families, from discipline-specific word lists to assist students in focusing their study time on the most relevant materials. I recently learned of the “Basic Engineering English Word List” that was developed to help students who may “fall short of the basic linguistic, and in particular lexical, knowledge necessary to read academic material in English” (Ward, 2009, p. 170). Ward describes the aims of this word list as summarizing terminology that is used across all engineering disciplines and being accessible both in terms of size and difficulty. Examining this targeted vocabulary in context and searching for literal and figurative uses of the language could go a long way in preparing students for their course reading materials.

Instructors also need to be aware of figurative language they may be using in their own written material, whether it be information or exercises distributed in class or test materials. I recently heard an anecdote related to a testing situation in which a student mistook the instruction “provide concrete examples” for being required to write something about physical concrete material in a construction context. This may have been a one-off, random example; however, it was in the context of a high-stakes English

proficiency test. As such, those creating and administering testing materials must take into account anything that could unfairly impede students' chances of success.

There are also implications in terms of diagnostic assessment, particularly in the context of university post-entry admission testing, which is becoming increasingly popular and important in improving student experience and retention (Fox, von Randow & Volkov, in press). Further exploration and tweaking of the metaphor test could result in it being a useful diagnostic tool. The study results also provide insight into the kinds of targeted reading assistance that could be provided to at-risk students whose first language is not English.

Universities have come a long way in providing student supports for both academic and non-academic issues; however, government and other funding is becoming increasingly scarce, and these programs are often scaled back or eliminated as a result (see for example Fox, 2009). In addition, or perhaps as a result of funding issues, often is the case that these supports are generic in nature, or "one size fits all" (Fox & Haggerty, 2014). Diagnostic assessment provides the opportunity to put in place subject-specific and student-specific supports. Fox and her team have been working with the Faculty of Engineering for a few years now to tweak and improve the SAFE diagnostic. A student support centre for engineering students was created and tutors use the diagnostic assessment as one of their tools (Fox & Haggerty, 2014). The more detailed information these tutors have regarding student strengths and weaknesses, the more efficiently they can target their advice and assistance. This subject and student-specific diagnostic information could also be used by instructors and teaching assistants in other contexts for efficient, targeted assistance.

Although students diagnosed as at-risk are not exclusively international students, nor do all EAL students struggle with their studies, this investigation suggests that metaphor comprehension may be a factor affecting EAL learner success, particularly with reading comprehension. These students already face a number of challenges, including traveling great distances away from their support systems to live and study within an academic culture that is often quite different than their own. Universities are increasingly competing to attract international students for purposes of “internationalization;” however, it is hard to ignore the fact that these students typically pay significantly higher tuition fees than their domestic classmates. In addition, English proficiency scores provided for applications to post-secondary institutions are not typically designed to predict academic preparedness, yet they are heavily depended upon as criteria for admission. This results in some students being admitted who are not sufficiently equipped with the academic tools they need to succeed in Western, English-medium institutions. This researcher feels strongly that it is the duty – in fact, moral responsibility – of institutions who recruit international students to do more than just pay lip-service when it comes to putting in place academic (and other) supports to facilitate student success.

Although the dollar value behind student retention should not be the driving force, evidence exists that this is a benefit of SAFE at this institution, despite the time, personnel, and financial costs associated with its development and maintenance. When the student support centre was first opened and two students were “talked down” from dropping out of the program by tutors, the Dean of Engineering was quoted as saying “Saving even two students pays for the entire cost of the diagnostic program” (Fox &

Haggerty, 2014, slide 15; Fox, Haggerty & Artemeva, forthcoming). Despite such strong evidence, the benefits of diagnostic assessment still seem to escape some departments and institutions.

Limitations of the Study

Although the findings of this study provide further recognition of the difficulties that some EAL learners have with English metaphor, there are a number of limitations that restrain the conclusions that can be drawn. There were issues faced during each phase of the project and decisions made that affected the outcomes that must be recognized. Broadly, it is a small-scale investigation that is not generalizable beyond the population considered in the study. The group of 42 students, over half of whom speak something other than English as their mother tongue, is not representative of the typical, larger intake of students that occurs in the fall term.

In the first phase of the study, the task of analyzing the corpus seemed overwhelming. The AntConc tools assisted in many ways, in both the exploratory and more fine-grained steps of the analysis. An entire study could have focused on a more in-depth examination of the corpus in order to quantify the metaphors found. However, the driving force behind this study was the metaphor comprehension task. Analyzing the corpus was important in order to find authentic occurrences of metaphor and to demonstrate that there is a sufficient amount of metaphoric content in the reading material to be concerned about it affecting students' reading comprehension. It was also important to use excerpts from the corpus to make the metaphor comprehension task more meaningful. However, since the amount of metaphoric content was not quantified,

the judgement of the prevalence of metaphors in the corpus overall, as well as statements regarding sub-disciplines having more or less metaphoric content, are subjective. In addition, the method by which the corpus was explored to arrive at the final collocations examined was awkward. A more systematic approach, such as random stratified sampling, may have been more efficient. This may have provided stronger and more precise evidence for the first research question.

There were also challenges in analyzing the text for metaphoric content. Sometimes this stemmed from separating different word senses from metaphoric uses of words or phrases in a given context. Other times the difficulty came from not understanding the subject content. As such, the instances of metaphoric content documented were conservative; only the instances when I felt certain the word/phrase in question was being used metaphorically did I make note. Although this would be more of a concern had the study design included quantification of metaphors, it provides ammunition for critics of Conceptual Metaphor Theory who protest the subjective nature by which they claim metaphors are determined.

The rating scale used to judge the students' interpretation of the metaphoric words/phrases was problematic in its intuitive development, determining the scale points, and implementing it to rate the responses. First, it was intuitively built because of the novel nature of this metaphor test. The data collected in Phase 2 of the study can now be incorporated to improve the scale and further testing can take place. One thing that could have been changed from the onset was the number of scale points. In hind sight, a 4-point scale would have been more useful, eliminating the "middle ground" point. Responses

would have been rated more clearly on degrees of correctness or incorrectness. This may have made rating simpler and easier.

The way in which the construct of metaphor comprehension was assumed to be demonstrated by students explaining metaphoric words and phrases in their own terminology is arguable. In addition, the rating of responses, which was the judgement of students' comprehension, was sometimes challenging due to unintelligible responses or those that were simply unclear in the context. Speaking to students directly would have allowed for follow-up questioning to clear up uncertainty and provide further evidence about their understanding of the metaphoric words and phrases. Given the evidence of overlap with reading and vocabulary tasks, further research is needed in order to better define the construct of metaphor comprehension. Other challenges with rating came from inconsistencies in rating similar responses. Reviewing the ratings more than once and documenting the rating process facilitated consistency, as did having a second rater assess some of the responses, but human rating always poses challenges in subjectivity and consistency (see for example, Lumley, 2002; Bejar, 2012).

The readability of the paragraphs containing the metaphoric words/phrases was also intuitively determined. The paragraphs were analyzed using Coh-Metrix in an attempt to strengthen the argument of readability; however, many of the indices provide no scale of reference in order to interpret the score. In some instances, one can determine that a low or high score increases or decreases readability; however, there is not typically any indication of what constitutes a low or high score. As such, the readability of the three paragraphs could only be determined in comparison with each other. A comparison

of text randomly drawn from other parts of the corpus may have given a basis of comparison to put these results into perspective.

Finally, missing data impeded the number of responses that could be analyzed, and in one case, the biographical information (e.g. L1) that could be linked to the responses. A number of test takers did not attempt some, or all, of the metaphor portion of the assessment. However, we cannot automatically assume that this was due to a lack of understanding of the metaphoric words/phrases. As previously discussed, the metaphor task was at the end of a 3-part assessment. Part 3, which contained two other tasks including a significant writing task, was limited to 30 minutes to complete. Some students may have run out of time or energy to respond. Some students may have opted not to respond to this or any other part of the assessment because there was no penalty. There were no grades assigned, or at least none that counted towards course/program credit, and as such, the stakes were low.

Directions for Future Research

Although there are a number of factors that limit or impede the conclusions that can be drawn in this study, it does contribute to the literature by providing further evidence of difficulties that EAL students have with English metaphor, specifically in the domain of reading and in the post-secondary context. In addition, it acts, in essence, as a pilot study of the potential of metaphor comprehension as a diagnostic tool in the domain of reading. As this was a small-scale study, there is potential for future research on a larger scale. This includes further corpus-driven studies. Corpora are increasingly being used in metaphor identification studies in part, due to technical innovations (see for

example Deignan, 2005; Berber Sardinha, 2007; Semino, Hardie, Koller, & Rayson, 2009; and Shutova, Teufel & Korhonen, 2013). Although still developing as a means of inquiry, this field holds promise for large-scale investigation into metaphor in use.

This study raises a number of questions that could fuel related inquiries. As previously mentioned, this entire study could have been based on a more detailed, quantitative analysis of this corpus. The corpus itself could also be enlarged to contain more texts within or beyond what is used this institution. This would provide a more accurate picture of the extent metaphor and the types of metaphors present in engineering texts more broadly. The sub-disciplines could also be more closely investigated, as students typically specialize after their first year of study. Similar studies in other disciplines, particularly business/commerce, would provide EAL and EAP instructors information on the types of metaphors to focus on that their students could encounter in their future studies. In addition, the domains of speaking and writing are grossly underrepresented in the literature; very little is known about how EAL students produce metaphor in their speech or academic writing.

At the beginning of this study I had hoped to interview some of the test takers to ask them about their experiences in learning English metaphors. However, it was beyond the scope of this thesis. This could be undertaken in an expanded version of this study, or a separate investigation taking a more qualitative (or mixed methods) approach. It would be valuable to learn how students do or do not identify and interpret metaphoric words and phrases in what they read and what they hear. It would also be of value to learn teachers' attitudes about metaphor in the curriculum and the ways in which it is introduced into the EAL and EAP classrooms since this is not done regularly or

systematically. Finally, I suspect from my own searches at conferences over the last two years that teaching metaphors happens with very little support materials. However, one example of EAP textbook that addresses metaphor competence is Caplan and Douglas' (2011) *Q: Skills for Success Reading and Writing 5*. Development of learning materials as well as glossaries in academic textbooks that pay attention to figurative uses of the terminology are crucial areas for further exploration.

There is a great deal of potential in the idea of pairing metaphor comprehension inquiry with research on other language forms. For example, data on the extent to which some metaphors are formulaic language or collocation that may be metaphoric in certain contexts may further inform EAL pedagogy. A natural link already discussed, but that could benefit from more exploration is with vocabulary comprehension, as well as academic and discipline-specific word lists. Finally, when looking at metaphor use in specific contexts, such as post-secondary study, and even more precisely in certain subject areas, a link with genre research could be explored. Deignan, Littlemore and Semino's 2013 book *Figurative Language, Genre and Register* is a positive step in this direction.

In the field of testing, it would be interesting to explore more about metaphor comprehension (or use) as a test construct. More evidence is needed to understand how to define metaphor comprehension in order to appropriately and accurately develop ways to test it. A related endeavour is to investigate the ways in which metaphor is used in testing materials and any effects this may have. As demonstrated in the "concrete example" anecdote, test writers must be conscious of factors that could alter the test construct or add unnecessary challenges for test takers. This neatly circles back to what first piqued

my interest in EAL student comprehension of English metaphors. When I looked at reading passages in TOEFL prep materials in my undergraduate studies, my hypothesis (or fear) was that they contained metaphoric content that would impede understanding of the text and increase the gate-keeping nature of the test. There are many tests and assessments created and used for many different purposes, some with higher stakes than others. Anyone creating and using a test has a responsibility to ensure its fairness. If inquiry into metaphoric terminology in task instructions or test content revealed comprehension issues unrelated to what was being tested, or other negative effects on student success, raising awareness of this issue is an important contribution of this study.

References

- Alderson, J. C., & Wall, D. (1999). Does washback exist? *Applied Linguistics*, *14*(2), 115-129. doi:10.1093/applin/14.2.115
- Altbach, P. G. (2004). Globalisation and the university: Myths and realities in an unequal world. *Tertiary Education & Management*, *10*(1), 3-25.
doi:10.1023/B:TEAM.0000012239.55136
- Anderson, N. J., Evans, N. W., & Hartshorn, K. J. (2014, March). U.S. national survey. In N. J. Anderson, N.W. Evans, K. J. Hartshorn, R. Keogh, & E. Webster, *Reading and writing expectations of matriculated university students*. Colloquium conducted at the TESOL International Convention & English Language Expo, Portland, OR.
- Anderson, A., & Lynch, T. (1988). *Listening*. Oxford: Oxford University Press.
- Anthony, L. (2014). *AntConc 3.4.3 - Readme*. Retrieved from <http://www.laurenceanthony.net/software.html>
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L., & Palmer, A. (2010). *Language assessment in practice*. Oxford: Oxford University Press.
- Bartlett, F. C. (1932). *Remembering*. Cambridge: Cambridge University Press.
- Bejar, I. I. (2012). Rater cognition: Implications for validity. *Educational Measurement: Issues and Practice*, *31*(3), 2-9. doi:10.1111/j.1745-3992.2012.00238.x

- Berber Sardinha, T. (2007). Metaphor in corpora: A corpus-driven analysis of applied linguistics dissertations. *Revista Brasileira de Lingüística Aplicada*, 7(1), 12-35.
- Boers, F. (2000). Enhancing metaphoric awareness in specialised reading. *English for Specific Purposes*, 19(2), 137-147. doi:10.1016/S0889-4906(98)00017-9
- Brown, J.D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall Regents.
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Cameron, L. (2003). *Metaphor in educational discourse*. London: Continuum.
- Cameron, L. (2010). The discourse dynamics framework for metaphor. In L. Cameron & R. Maslen (Eds.), *Metaphor analysis. Research practice in applied linguistics, social sciences and humanities* (pp. 77-94). Oakville, CT: Equinox Publishing Ltd.
- Canadian Federation of Students (2015). *International students*. Retrieved from <http://cfs-fcee.ca/the-issues/international-students/>
- Caplan, N. & Douglas, S. (2011). *Q: Skills for Success Reading and Writing 5*. New York, NY: Oxford University Press.
- Carter, R. (1997). *Investigating English discourse*. London: Routledge.
- Cheng, L. (2008). Washback, impact and consequences. In E. Shohamy, & N. H. Hornberger (Eds.), *The encyclopedia of language and education, volume 7, language testing and assessment* (pp. 349 – 364). NY: Springer Science+Business Media, LLC.

- Cheng, L., & Fox, J. (2008). Towards a better understanding of academic acculturation: Second language students in Canadian universities. *Canadian Modern Language Review*, 65(2), 307-333.
- Clark, E. V. (1981). Lexical innovations: How children learn to create new words. In W. Deutsch (Ed.) *The Child's Construction of Language*. London: Academic Press.
- Clark, E. V. (1982). The young word maker: A case study of innovation in the child's lexicon. In E. Wanner & L.R. Gleitman (Eds.), *Language Acquisition: The state of the art* (pp 390-426). Cambridge: Cambridge University Press.
- Clark, H. H., & Clark, E. V. (1977). *Psychology and language: An introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.
- Coxhead, A. (2012). Academic vocabulary, writing and English for academic purposes: Perspectives from second language learners. *RELC Journal*, 43(1), 137-145.
- Crossley, S. A., Allen, D. B., & McNamara, D. S. (2011). Text readability and intuitive simplification: A comparison of readability formulas. *Reading in a Foreign Language*, 23(1), 84-101.
- Creswell, J. W. (2015). *A concise introduction to mixed methods research*. Thousand Oaks, California: SAGE.
- Cummins, J. (1981). *Bilingualism and language minority children*. Toronto, ON: Ontario Institute for Studies in Education.
- Danesi, M. (1995). Learning and teaching languages: The role of "conceptual fluency". *International Journal of Applied Linguistics*, 5(1), 3-20. doi:10.1111/j.1473-4192.1995.tb00069.x

- Deignan, A. (2005). *Metaphor and corpus linguistics*. Philadelphia, PA: John Benjamins Publishing Co.
- Deignan, A., Littlemore, J., & Semino, E. (2013). *Figurative language, genre and register*. Cambridge: Cambridge University Press.
- Delpit, L. (1988). The silence in dialogue: Power and pedagogy in educating other people's children. *Harvard Educational Review*, 58(3), 280-298.
- Dörnyei, Z. (2007). *Research methods in applied linguistics. Quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.
- Dorst, A. G. (2015). More or different metaphors in fiction? A quantitative cross-register comparison. *Language and Literature*, 24(1), 3-22.
doi:10.1177/0963947014560486
- Douglas, S. R. (2013). After the first 2000: A response to Horst's "Mainstreaming second language vocabulary acquisition". *The Canadian Journal of Applied Linguistics*, 16(1), 189-199.
- Duffy, G. G. (2009). *Explaining reading: A resource for teaching concepts, skills, and strategies*. New York: The Guilford Press.
- Dunworth, K. (2010). Clothing the emperor. Addressing the issue of English language proficiency in Australian universities. *Australian Universities Review*, 52(2), 5-10.
- Ebel, R. L. (1979). *Essentials of educational measurement* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

- Floyd, P., & Carrell, P. L. (1987). Effects on ESL reading of teaching cultural content schemata. *Language Learning*, 37(1), 89-108. doi:10.1111/j.1467-1770.1968.tb01313.x
- Fox, J. (2009). Moderating top-down policy impact and supporting EAP curricular renewal: Exploring the potential of diagnostic assessment. *Journal of English for Academic Purposes*, 8, 26-42. doi:10.1016/j.jeap.2008.12.004
- Fox, J. (2015). Editorial, Trends and issues in language assessment in Canada: A considerations of context. Special Issue on Language Assessment in Canada. *Language Assessment Quarterly*, 12, 1-9. doi:10.1080/15434303.2014.999921
- Fox, J., & Haggerty, J. (2014, March). *Mitigating risk in first-year engineering: Post-admission diagnostic assessment in a Canadian university*. Paper presented at the Annual Conference of the American Association of Applied Linguistics, Portland, OR.
- Fox, J., Haggerty, J., & Artemeva, N. (forthcoming). Mitigating risk: The impact of a diagnostic assessment procedure on the first-year experience in engineering. In J. Springer volume: *Post-admission assessment of university students*. New York, NY: Springer.
- Fox, J., von Randow, J., & Volkov, A. (in press). Identifying students-at-risk through post-entry diagnostic assessment: An Australasian approach takes root in a Canadian university. In V. Aryadoust & J. Fox (Eds.), *Current trends in language testing in the Pacific Rim and the Middle East: Policies, analyses, and diagnoses*. Newcastle upon Tyne, UK: Cambridge Scholars Press.

- Fulcher, G. (1996). Does thick description lead to smart tests? A data-based approach to rating scale construction. *Language Testing*, 13, 208-238. doi: 10.1177/026553229601300205
- Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books Inc.
- Grabe, W. (2009a). Epilogue: Reflections on second language reading research and instruction. In Z. H. Han, & N. J. Anderson (Eds.), *Second language reading research and instruction: Crossing the boundaries* (pp. 192-205). Ann Arbor, MI: The University of Michigan Press.
- Grabe, W. (2009b). *Reading in a second language. Moving from theory to practice*. New York: Cambridge University Press.
- Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, 36(2), 193-202. doi:10.3758/BF03195564
- Green, A. (2013). Washback in language assessment. *International Journal of English Studies*, 13(2), 39-51.
- Green, A., & Weir, C. (2004). Can placement tests inform instructional design? *Language Testing*, 21(4), 467-494. doi:10.1191/0265532204lt293oa
- Gunderson, L., Slade, K., & Rosenke, D. (1988). The cloze response patterns of L1 and L2 students to literal and idiomatic text. *TESL Canada Journal*, 6(1), 60-66.
- Hamp-Lyons, L. (2007). The impact of testing practices on teaching. In J. Cummins, & C. Davison (Eds.), *International handbook of English language teaching* (pp. 487-504). Boston, MA: Springer US.

- Harrison, R. T. (1980). *Readability in the classroom*. Cambridge: University Press.
- Henning, G. (1991). *A study of the effects of contextualization and familiarization on responses to the TOEFL vocabulary items*. TOEFL Research Reports, 35. Princeton, NJ: Educational Testing Service.
- Horst, M. (2009). Developing definitional vocabulary knowledge and lexical access speed through extensive reading. In Z. H. Han, & N. J. Anderson (Eds.), *Second language reading research and instruction: Crossing the boundaries*. Ann Arbor, MI: The University of Michigan Press.
- Horst, M. (2013). Mainstreaming second language vocabulary acquisition. *Canadian Journal of Applied Linguistics*, 16(1), 171-188.
- Institute of International Education. (2013a). Fields of study of students from selected places of origin, 2012/13. *Open Doors Report on International Educational Exchange*. Retrieved from <http://www.iie.org/opendoors>
- Institute of International Education. (2013b). Top 25 places of origin of international students, 2011/12-2012/13. *Open Doors Report on International Educational Exchange*. Retrieved from <http://www.iie.org/opendoors>
- Jang, E. E. (2009). Demystifying a Q-Matrix for making diagnostic inferences about L2 reading skills. *Language Assessment Quarterly*, 6, 210-238.
doi:10.1080/15434300903071817
- Johnson, P. (1982). Effects on reading comprehension of building background knowledge. *TESOL Quarterly*, 16(4), 503-516.
- Kathpalia, S. S., & Carmel, H. L. H. (2011). Metaphorical competence in ESL student writing. *RELC Journal*, 42(3), 273-290. doi:10.1177/0033688211419379

- Keogh, R. (2014, March). Academic challenges: The international student point of view. In N. J. Anderson, N. Evans, J. Hartshorn, R. Keogh, & E. Webster, *Reading and writing expectations of matriculated university students*. Colloquium conducted at the TESOL International Convention & English Language Expo, Portland, OR.
- Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition*, 2(1), 15-47. doi:10.1016/0010-0277(72)90028-5
- Kövecses, Z. (2005). *Metaphor in culture: Universality and variation*. Cambridge: Cambridge University Press.
- Kövecses, Z. (2008). Conceptual metaphor theory. Some criticisms and alternative proposals. *Annual Review of Cognitive Linguistics*, 6(1), 168-184. doi:10.1075/arcl.6.08kov
- Kövecses, Z. (2010). *Metaphor: A practical introduction* (2nd ed.). New York: Oxford University Press.
- Lakoff, G. (1994). Conceptual metaphor home page. <http://www.lang.osaka-u.ac.jp/~sugimoto/MasterMetaphorList/MetaphorHome.html>
- Lakoff, G. (2008). The neural theory of metaphor. In R. Gibbs (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 17-38). Cambridge: Cambridge University Press.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

- Lam, A. S-L. (1984). *The relative importance of content words and function words as related to syntactic complexity, English proficiency and language transfer in the reading comprehension of English as a second language (ESL) learners.* (Unpublished doctoral dissertation). University of Pittsburgh, Pittsburg, PA.
- Learned, J. E., Stockdill, D., & Birr Moje, E. (2011). Integrating reading strategies and knowledge building in adolescent literacy instruction. In S. J. Samuels, & A. E. Farstrup (Eds.), *What research has to say about reading instruction* (4th ed.) (pp. 159-185). Newark, DE: International Reading Association.
- Littlemore, J. (2001a). Metaphoric competence: A language learning strength of students with a holistic cognitive style? *TESOL Quarterly*, 35(3), 459-491.
- Littlemore, J. (2001b). Metaphoric intelligence and foreign language learning. In *Humanising Language Teaching*, 3(2). Retrieved from <http://www.hltmag.co.uk/mar01/mart1.htm>
- Littlemore, J. (2001c). The use of metaphor in university lectures and the problems that it causes for overseas students. *Teaching in Higher Education*, 6(3), 333-349.
doi:10.1080/13562510120061205
- Littlemore, J., & Low, G. (2006a). *Figurative thinking and foreign language learning.* New York: Palgrave Macmillan.
- Littlemore, J., & Low, G. (2006b). Metaphoric competence, second language learning, and communicative language ability. *Applied Linguistics*, 27(2), 268-294.
doi:10.1093/applin/aml004

- Littlemore, J., Trautman Chen, P., Koester, A., & Barnden, J. (2011). Difficulties in metaphor comprehension faced by international students whose first language is not English. *Applied Linguistics*, 32(4), 408-429. doi:10.1093/applin/amr009
- Low, G. D. (1988). On teaching metaphor. *Applied Linguistics* 9(2), 125–47. doi:10.1093/applin/9.2.125
- Lumley, T. (2002). Assessment criteria in a large-scale writing test: What do they really mean to the raters? *Language Testing*, 19(3), 246-276. doi:10.1191/0265532202lt230oa
- MacArthur, F. (2010). Metaphorical competence in EFL: Where are we and where should we be going? A view from the language classroom. *AILA Review*, 23(1), 155-173. doi:10.1075/aila.23.09mac
- MacLennan, C. H. G. (1994). Metaphors and prototypes in the teaching and learning of grammar and vocabulary. *International Review of Applied Linguistics in Language Teaching*, 32(2), 97-110.
- Master, P. (1991). Active verbs with inanimate subjects in scientific prose. *English for Specific Purpose*, 10(1), 15-33. doi: 10.1016/0889-4906(91)90013-M
- McCarthy, P.M., Lewis, G.A., Dufty, D.F., & McNamara, D.S. (2006). Analyzing writing styles with Coh-Metrix. In G. Sutcliffe, & R. Goebel (Eds.), *Proceedings of the Nineteenth International intelligence Research Society Conference* (pp. 764–769). Menlo Park, CA: AAAI Press.
- McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge, M.A.: Cambridge University Press.

- McNamara, D.S., Louwarse, M.M., Cai, Z., & Graesser, A. (2005, January 1). Coh-Metrix version 1.4. Retrieved [2015, April 2], from <http://cohmetrix.memphis.edu>.
- Meissner, V. H. (2010). *Metaphorical comprehension in grade 8 English language learners from one town in Norway* (Unpublished master's dissertation). Hedmark University College, Elverum, Norway.
- Moon, R. (1998). *Fixed expressions and idioms in English: A corpus based approach*. Oxford: Clarendon Press.
- Mukundan, J., & Roslim, N. (2009). Textbook representation of prepositions. *English Language Teaching*, 2(4), 13-24. doi:10.5539/elt.v2n4p13
- Nunberg, G. (1987). Poetic and prosaic metaphors. In *Proceedings of the Sixteenth National Conference on Artificial Intelligence, AAAI 99* (pp. 121-128). Menlo Park, CA: AAAI Press.
- Pan, Y., & Newfields, T. (2011). Teacher and student washback on test preparation evidenced from Taiwan's English certification exit requirements. *International Journal of Pedagogies and Learning*, 6(3), 260-272. doi:10.5172/ijpl.2011.6.3.260
- Piekarz, J. (1956). Getting meaning from reading. *Elementary School Journal*, 56, 303-309.
- Pollio, H., Barlow, J., Fine, H., & Pollio, M. (1977). *The poetics of growth: Figurative language in psychology, psychotherapy, and education*. Hillsdale, N.J: Lawrence Erlbaum Associates.

Plano Clark, V. L., & Creswell, J. W. (2008). *The mixed methods reader*. Los Angeles: Sage Publications.

Pragglejaz Group. (2007). MIP: A method for identifying metaphorically used words in discourse. *Metaphor and Symbol*, 22(1), 1-39. doi:10.1080/10926480709336752

Rasinski, T. V., & Samuels, S. J. (2011). Reading fluency: What it is and what it is not. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about reading instruction* (4th ed.) (pp. 94-114). Newark, DE: International Reading Association.

Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language Testing*, 18(4), 429-462.
doi:10.1177/026553220101800407

Read, J. (2008). Identifying academic needs through diagnostic assessment. *Journal of English for Academic Purposes*, 7(3), 180-190. doi:10.1016/j.jeap.2008.02.001

Read, J. (2013). Issues in post-entry language assessment in English-medium universities. *Language Teaching*, 48(2), 217-234. doi:10.1017/S0261444813000190

Renton, N. E. (1990). *Metaphorically speaking. A dictionary of 3,800 picturesque idiomatic expressions*. New York: Warner Books Inc.

Roessingh, H. (2005). The intentional teacher: The mediator of meaning making. *The Journal of Educational Thought*, 39(2), 111-134.

Roessingh, H., & Douglas, S. R. (2012). Educational outcomes of English language learners at university. *Canadian Journal of Higher Education*, 42(1), 80-97.

- Roessingh, H., & Kover, P. (2003). Variability of EAL learners' acquisition of cognitive academic language proficiency: What can we learn from achievement measures? *TESL Canada Journal*, 21(1), 1-21.
- Semino, E., Hardie, A., Koller, V., & Rayson, P. (2009). A computer-assisted approach to the analysis of metaphor variation across genres. In *Proceedings of Corpus-Based Approaches to Figurative Language: a Corpus Linguistics 2009 Colloquium*. University of Birmingham.
- Shutova, E., Teufel, S. & Korhonen, A. (2013). Statistical metaphor processing. *Computational Linguistics*, 39(2), 301-353. doi:10.1162/COLI_a_00124
- Skorczyńska Sznajder, H. (2010). A corpus-based evaluation of metaphors in a business English textbook. *English for Specific Purposes*, 29(1), 30-42.
doi:10.1016/j.esp.2009.05.003
- Steen, G. J., Dorst, A. G., Berenike Herrmann, J., Kaal, A. A., & Krennmayr, T. (2010). Metaphor in usage. *Cognitive Linguistics*, 21(4), 765-796.
- Steffensen, M. S., Joag-Dev, C., & Anderson, R. C. (1979). A cross-cultural perspective on reading comprehension. *Reading Research Quarterly*, 15(1), 10-29.
- Swan, M. (1988). *Practical English Usage*. Oxford: Oxford University Press.
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam: John Benjamins.
- Turner, C. E., & Fox, J. (2015, March). *Situating mixed methods research in language assessment: Evolution and practice*. Pre-Conference Workshop given at the Language Testing Research Colloquium, Toronto, ON.
- Upsher, J. A., & Turner, C. E. (1995). Constructing rating scales for second language tests. *ELT Journal*, 49(1), 3-12. doi:10.1093/elt/49.1.3

Ward, J. (2009). A basic engineering English word list for less proficient foundation engineering undergraduates. *English for Specific Purpose*, 28, 170-182.

doi:10.1016/j.esp.2009.04.001

Webster, E. (2014, March). What do students need? MSU faculty perspectives. In N. J. Anderson, N. Evans, J. Hartshorn, R. Keogh, & E. Webster, *Reading and writing expectations of matriculated university students*. Colloquium conducted at the TESOL International Convention & English Language Expo, Portland, OR.

Willis, A. I. (2008). *Reading comprehension research and testing in the U. S.*

Undercurrents of race, class and power in the struggle for meaning. New York: Taylor & Francis Group, LLC.

Wood, D. C., & Appel, R., (2012). [*First Year Engineering Texts*]. Unpublished corpus.

Wood, D. C., & Appel, R., (2014). Multiword constructions in first year business and engineering university textbooks and EAP textbooks. *Journal of English for Academic Purposes*, 15, 1-13. doi:10.1016/j.jeap.2014.03.002

Appendix A

Metaphor Comprehension Task

Part 3.3. Reading for detail in Engineering

DIRECTIONS: The paragraphs below (A., B., and C.), are taken from first-year chemistry for engineering textbooks. In the space provided on the right-hand side of the page (see Comment boxes), please describe in your own words the meaning of each of the underlined words/phrases in the paragraphs.

A. Compared with other subjects, chemistry is commonly believed to be more difficult, at least at the introductory level. There is some justification for this perception; for one thing, chemistry has a very specialized vocabulary. However, even if this is your first course in chemistry, you already have more familiarity with the subject than you may realize. In everyday conversations we hear words that have a chemical connection, although they may not be used in the scientifically correct sense. Examples are "electronic," "quantum leap," "equilibrium," "catalyst," "chain reaction," and "critical mass." Moreover, if you cook, then you are a practicing chemist! From experience gained in the kitchen, you know that oil and water do not mix and that boiling water left on the stove will evaporate.

B. All sciences, including the social sciences, employ variations of what is called the scientific method, a systematic approach to research. For example, a psychologist who wants to know how noise affects people's ability to learn chemistry and a chemist interested in measuring the heat given off when hydrogen gas burns in air would follow roughly the same procedure in carrying out their investigations. The development of science has been irregular and sometimes even illogical. Great discoveries are usually the result of the cumulative contributions and experience of many workers, even though the credit for formulating a theory or a law is usually given to only one individual. For every success story, however, there are hundreds of cases in which scientists have spent years working on projects that ultimately led to a dead end, and in which positive achievements came only after many wrong turns and at such a slow pace that they went unheralded. Yet even the dead ends contribute something to the continually growing body of knowledge about the physical universe.

C. Energy production and energy utilization are closely tied to the quality of our environment. A major disadvantage of burning fossil fuels is that they give off carbon dioxide, which is a greenhouse gas. That is, it promotes the heating of Earth's atmosphere. Carbon dioxide, sulfur dioxide and nitrogen oxides result in acid rain and smog. Harnessing solar energy has no such detrimental effects on the environment. The energy contained in tidal waves can also be harnessed to perform useful work, but the relationship between tidal waves and chemistry is minimal.

 tinabeynen
Introductory level

 tinabeynen
Chemical connection

 tinabeynen
Experience gained

 tinabeynen
(All sciences) employ

 tinabeynen
Carrying out (investigations)

 tinabeynen
(Projects) led to a dead end

 tinabeynen
Body of knowledge

 tinabeynen
(Energy production and utilization) are closely tied to

 tinabeynen
Harnessing (solar energy)

 tinabeynen
Relationship (between tidal waves and chemistry)

Appendix B

Stop Word List

a	ours	if
an	they	or
the	them	neither
I	their	nor
me	theirs	
my	it	
mine	its	
you	this	
your	that	
yours	these	
he	those	
him	who	
his	whom	
she	whose	
her	and	
hers	or	
we	but	
our	because	