



All Theses and Dissertations

---

2015-11-01

# Idioms as a Measure of Proficiency

Kyle Hugh Vanderniet  
*Brigham Young University - Provo*

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>

 Part of the [Linguistics Commons](#)

---

## BYU ScholarsArchive Citation

Vanderniet, Kyle Hugh, "Idioms as a Measure of Proficiency" (2015). *All Theses and Dissertations*. 5817.  
<https://scholarsarchive.byu.edu/etd/5817>

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact [scholarsarchive@byu.edu](mailto:scholarsarchive@byu.edu), [ellen\\_amatangelo@byu.edu](mailto:ellen_amatangelo@byu.edu).

Idioms as a Measure of Proficiency

Kyle Hugh Vanderniet

A thesis submitted to the faculty of  
Brigham Young University  
in partial fulfillment of the requirements for the degree of

Master of Arts

Cynthia L. Hallen, Chair  
Jesse Egbert  
Dan Dewey

Department of Linguistics and English Language

Brigham Young University

November 2015

Copyright © 2015 Kyle Hugh Vanderniet

All Rights Reserved

## ABSTRACT

### Idioms as a Measure of Proficiency

Kyle Hugh Vanderniet  
Department of Linguistics and English Language, BYU  
Master of Arts

This purpose of this thesis is to explore the relationship between idiomatic knowledge and second language proficiency. As the amount of research directly related to this topic is sparse, an in-depth discussion of relevant research and definitions comprises the first part of this paper. Two studies are then presented here that test the relationship between idiomatic knowledge and second language proficiency. A new definition for idioms proposes that all non-compositional phrases, popularized by usage, that is stored, retrieved, and employed as a single lexical unit. The results from more than 400 participants across two studies indicate that the two constructs are strongly correlated, but that the observable relationship between them is only modest. Additional results from the two studies also indicate that idiomatic knowledge is more strongly correlated with speaking skill than with writing or reading skills. The impacts of this study on existing research are discussed and directions for new research is suggested.

Keywords: Idioms, Idiomatic Knowledge, Decomposability, ESL, Proficiency, Fluency, TOFEL,

## AKNOWLEDGEMENTS

I would like to express thanks to my committee for their patience in working with me. I am grateful for all of the revisions, notes, and guidance from Dr. Hallen and especially for willingness to be my last minute chair. I would also like to thank Dr. Egbert who helped with my statistical analysis very late in my paper's development. I am also extremely grateful to Dr. Dewey who has agreed to continue on as my third reader.

I am also grateful to LoriAnne Spear who has helped me persevere through four chairs and four committee members. She has also extended far more time and energy than expected to make sure that I am compliant and ready for graduation. I must also acknowledge my wife, Tiffany, who has supported me through these four years of school and 18 months of thesis writing.

I am also grateful for the many staff and faculty that have lent assistance from the beginning of this project, from Dr. Anderson to Dr. Eddington. There have been a great many people who have helped me complete this project and I could not have done it without them. Finally, I am grateful for the many blessings from God, to continue through disaster and hardship.

## TABLE OF CONTENTS

TABLE OF CONTENTS.....	iv
LIST OF TABLES.....	v
LIST OF FIGURES.....	vi
<b>Chapter 1: Introduction .....</b>	<b>1</b>
<b>Chapter 2: Literature Review .....</b>	<b>5</b>
2.1    Definitions.....	5
2.2    Foundations.....	9
2.3    Decomposability .....	11
2.4    Idioms for ESL.....	16
2.5    Idioms and transfer .....	16
<b>Chapter 3: Experiment Methodology .....</b>	<b>21</b>
3.1    Study One – Intensive English Program.....	21
3.1.2    Participants.....	21
3.1.3    Measure.....	23
3.1.4    Procedures:.....	25
3.1.5    Data Gathering .....	25
3.2    Study 2 .....	27
3.2.2    Participants.....	27
3.2.3    Measure.....	28
<b>Chapter 4: Data Analysis .....</b>	<b>31</b>
4.1    Study One – Intensive English Program.....	31
4.2    Study Two.....	33
<b>Chapter 5: Discussion .....</b>	<b>35</b>
5.1    Limitations .....	37
5.2    Suggestions for Further Research .....	38
5.2    Conclusions.....	39
<b>APPENDIX A: Libben &amp; Titone 2008 Idiom Results .....</b>	<b>41</b>
<b>APPENDIX B: Survey for Studies One and Two .....</b>	<b>43</b>
<b>References.....</b>	<b>51</b>

## LIST OF TABLES

<b>Table 1:</b> Composition of Participants for Study One.....	22
<b>Table 2:</b> Results of Study One .....	26
<b>Table 3:</b> Number of results by Language for Study 2.....	29
<b>Table 4:</b> Correlation of Constructs from Study One.....	31
<b>Table 5:</b> Pearson Correlation from Study One.....	<b>Error! Bookmark not defined.</b>

## LIST OF FIGURES

<b>Figure 2-1</b> .....	6
<b>Figure 2-2</b> .....	6
<b>Figure 2-3</b> .....	7
<b>Figure 2-4</b> .....	11
<b>Figure 2-5</b> .....	12
<b>Figure 2-6</b> .....	13
<b>Figure 2-7</b> .....	14
<b>Figure 2-8</b> .....	14
<b>Figure 3-1</b> .....	23
<b>Figure 3-2</b> .....	25
<b>Figure 3: Loess Regression line</b> .....	<b>Error! Bookmark not defined.</b>

## **Chapter 1:**

### **Introduction**

A great deal of research has been done on idioms in a wide variety of fields. However, research into idioms as a measure of proficiency for second language learners is very sparse. Several researchers have tangentially approached the topic (Palmer 1925, Warren 2005, Gibbs et al 1989, Skehan 1998), and research on both idioms and second language proficiency abounds, but investigation into a possible relationship between the two concepts remains a field in need of exploring. In this paper I ask the question, what is the relationship between idiomatic knowledge and language proficiency?

In asking this question I am relying heavily on the work done by the researchers who have dealt directly with this question (Libben and Titone 2011; Irujo 1986). In order to establish support for my research on this topic, I branched out to other researchers who have produced work in either of the two fields that is relevant to my work here. I used research from the fields of idioms and second language proficiency to establish a framework to use in conducting my own study.

Idiom use by non-native speakers and idiom difficulty (Libben and Titone 2011, Gibbs & Nayak 1985) are two examples of relevant fields that I have explored in this pursuit. I also explored the various systems for categorizing idioms in order to better understand which idioms I should include in the studies. In the course of my research I reviewed many other topics for helpful studies and concluded that many of them were not pertinent to what I was trying to do here. Below I discuss the topics of idioms categorization and idiom flexibility because I have

found them helpful in establishing the premises for my and building a framework for my own experiment and research.

In an effort to support my research into a relatively uncharted field, I explored many topics on idiomatic knowledge to see which areas of research had consensus and supportive evidence. I found that many important concepts in many these fields had much disagreement. Researches had no agreed upon or commonly used definition for what an idiom is or how to categorize them. There are a few areas that researchers have established a consensus on, unfortunately many of those areas do not have any particular relevance to what I'm studying here.

In particular, all the research that I used recognized the value of idiomatic knowledge in second language learning generally and acknowledged that learning idioms would undeniably increase a learner's ability to sound more fluent. Jackendoff (1997) points out that idioms are pervasive in our communication and that we have nearly as many idioms in English as we do adjectives. It is also clear that idiom usage by non-native English speakers helps them sound more fluent, just as learning more vocabulary, improved grammatical accuracy, and better pronunciation (Palmer 1925, Irujo 1986). However, even these conclusions are only agreed generally that idioms are good for learning a second language, as none of these researchers set out with an agenda of exploring the relationship any further.

I also found that it is generally agreed that the nature of idioms makes them more difficult to learn and incorporate than other lexical items. Warren (2005) states that non-native speakers of English learn the meanings of idioms because they (the learners) "naturally memorize what is repeated" and because idioms are so pervasive in the discourse of native speakers that idioms

become memorized units, in essence a single lexeme for a complex idea. However, non-native speakers do not learn and incorporate these idioms in their normal dialog without assistance because idioms are typically only encountered in casual dialog between native speakers (Warren 2005). Additionally, since the meaning carried by the expression is “fixed” or “dead” (Warren 2005; Irujo, 1986) and not intrinsic, it makes idioms more difficult to learn than other lexical items. This unchanging and metaphorical nature of these idioms typically restricts their use to extremely narrow applications; which makes it even more difficult and unlikely that non-native speakers will learn idioms without assistance. (p. 288).

Inasmuch as my research bridges the gap between the idiomatic knowledge and language proficiency, I was faced with a staggering amount of information regarding each topic independently but very little directly related to my research question. Therefore, in regards to linguistic proficiency I decided that the best way to move forward was to simply use the standardized tests, such as the TOEFL and IELTS as measures for linguistic proficiency because they have had an enormous amount of work done to ensure they are an effective measure of language proficiency. These tests are generally agreed upon as effective in evaluating language proficiency and are rigorously maintained. Therefore, I did not deem it wise or prudent to sift through all the information on linguistic proficiency only to contrive my own way of determining how to measure the language proficiency. Unfortunately, the situation for idioms is not similarly simple.

Many important points regarding idioms are far from a consensus, and therefore I had to expand my area of focus in order me to establish a basis for my research question. Research in the area of idioms does have some agreement in areas such as retrieval speed area of storage, but

they are mostly irrelevant to my research. However, agreement concerning some of the most important areas related to my research like, "what is an idiom" are far from decided.

In the next chapter I will discuss the research that builds the foundation of my argument and lay out the definitions for the parameters I will be using. I will evaluate current definitions of idioms and evaluate what research has been done on the idiomatic knowledge as it applies to proficiency. In chapter 3 I will explain the layout of my experiment and explain the results. Chapter 4 contains my evaluations of the outcomes and I will discuss what could be learned from my study. Finally in Chapter 5 I will have my conclusions on how the data impacts what we know about the relationship between the two constructs as well as suggestions for further research.

## Chapter 2:

### Literature Review

#### 2.1 Definitions

Since the use of idioms is central to this thesis, exactly what an idiom is needs to be well defined. The existing definitions vary widely, and so, I will first examine several of these approaches to idioms and discuss their implications. One challenge with defining idioms is that even most laypeople have a working understanding of what an idiom is, but when someone tries to describe them based on lexical or syntactic characteristics, they encounter difficulty with being precise as well as accurate. Often the result is that each definition excludes groups of idioms and is objected to by fellow researchers based on these counter cases.

The *Oxford English Dictionary* (OED) states that an idiom is: “a group of words established by usage as having a meaning not deducible from the meanings of the individual words.” This definition gives finite criteria for deciding what is and isn’t an idiom, but it is an imperfect definition because the definition does not allow for idioms that do not fit the definition, but are, nevertheless idioms. I will be referring to idioms that are difficult to fit into the parameters established by any definition as ‘edge cases’. As I will explain with each definition, edge cases are problematic for each approach in their own way. Some examples of edge cases for the OED definition are shown in Figure 2-1. In each of these cases the meaning that is established by usage could probably be deduced by a native English speaker, even if they had never heard it before.

- a. Resist temptation*
- b. Second thoughts*
- c. Back to the drawing board*

**Figure 2-1**

Nunberg et al. (1994 p.492-493) developed a different definition for idioms that is more inclusive and flexible than the OED definition because it illustrates what a prototypical idiom is based on conventionality, inflexibility, figuration, proverbially, informality, and affect instead of establishing finite, and somewhat arbitrary, boundaries. This alternative definition is more precise than the OED approach because it does not establish any fixed criteria that establish a forced, dichotomous decision of either a phrase being or not being an idiom.

He explained that if a given phrase is a fixed expression that is used popularly in the same fashion as idioms, but does not share one or more of the characteristics that his definition propose, then usage itself would still provide some basis for labeling the phrase as idiomatic (p 493-494), like those seen in Figure 2-2.

- a. jealous much?*
- b. Yada yada yada.*
- c. QQ more.*

**Figure 2-2**

As our language is ever evolving and with idioms becoming antiquated and retired while others are regularly invented and institutionalized, we still strive for a way of incorporating these kinds of changes that do not rely on the antiquated or imprecise definitions.

This definition's strength is also its own weakness in that it has no method for deciding if something qualifies as being an idiom. Using this approach, idioms can be evaluated on several characteristics but there are no thresholds or standards for inclusion or exclusion into idiomhood. Nunberg acknowledge this shortcoming and continued that the main difficulty with devising a definition for idioms is that researchers are working with an ever changing subject matter. Old idioms become antiquated and fade from usage while new ones are constantly being invented; some of which catch on and are canonized.

A great deal of research has been done to define and categorizes idioms, but due to ongoing disagreement on what constitutes an idiom and how to categories them, it is difficult to find idioms that are universally agreed upon. Nevertheless, there are still some phrases that are regularly used by researchers that are undisputed, such as those found in Figure 2-3, that stand as unambiguous starting points.

- a. Kick the bucket.*
- b. Shoot the breeze.*
- c. Meet your maker.*

**Figure 2-3**

These phrases are among the phrases that the vast majority of people know wholesale as well as having been thoroughly used in research across the fields of idioms research, including Gibbs & Nayak (1989) to Libben and Titone (2011) Maher (2013). Another example of established idioms is found in Libben and Titone (2008), where they produced a list of 212 common VP idioms that all fall clearly within the boundaries of any accepted definition of idiom. While the

boundaries of what an idiom is still remain clouded, there are certainly idioms that can be employed as a focal point from which to study all the ill-defined edges.

These and more definitions are available for use, but each one has its own flaws. The focus of this experiment is not to get bogged down in defining what constitutes an idiom. Therefore I have elected to not use any of the existing definitions of idioms, but rather, I will posit a working definition that will allow the reader to understand what I am including when I refer to idioms.

For the purpose of this paper, I am using the term *idiom* more broadly to refer to any non-compositional phrase that is popularized by usage and that is stored, retrieved, and used as a single lexical unit. This would include any phrase that has a meaning beyond its constituent words and is used regularly and widely enough to be general knowledge to a given community. This remains as imprecise a measure as any, but it includes the above mentioned universally-acknowledged idioms I have used in my research without excluding large groups of other idioms. I will defer to other researchers to continue the work of defining idioms.

Because I use several technical terms in my thesis I will define them here. By the terms “linguistic proficiency” I mean the measurable skill that a speaker possesses with their second language, such as that measured by the TOEFL test. I am referring to proficiency on a scale rather than as an absolute. I chose this usage because my research focus is on ESL (English as

a Second Language) students and their language skill, or proficiency, as it compares against their idiomatic knowledge. Idiomatic knowledge is, for the purposes of this paper, the skill which one possesses in knowing, recognizing, and being able to produce or utilize idioms. I will be comparing idiomatic knowledge and language skill, or linguistic proficiency, throughout the paper.

Although I use a wide definition for idioms, I further narrow down the idioms used in my studies presented here to just verb phrase idioms. This is done primarily for simplification so that I only had one type of idiom, rather than having to account for different idioms types influencing the number of correct answers. A second reason for using only verb phrase idioms is that the researchers Libben & Titone (2008), upon whose research my idioms selection is based, compiled only verb phrase idioms as part of their research.

Additionally, I frequently use the term “transparent” or “transparency” to indicate whether or not an idiom is understandable from its constituent parts. Many idioms can be heard for the first time by a native speaker and understood perfectly while other idioms require an explanation, or at least context clues. As I will explain in a later chapter, this transparency is not necessarily dichotomous, but runs on a continuum. I use the term “transparent” to mean the opposite of opaque on the continuum. Opaque idioms are those that are not able to be understood from just the constituent words or without explanation.

## **2.2 Foundations**

A great deal of research has been done on idioms from many different fields including retrieval speeds, categorization, flexibility, literality judgments, and transfer in the second language. However, since research on idioms as they relate to a measure of linguistic aptitude is

lacking, it is necessary for me to first establish what is known about idioms and language proficiency before moving forward. I will discuss some of the research about idiom storage, categorization, idiom retrieval, idioms flexibility, and interference with idioms between the L1 and L2. The areas are closely linked with idiom knowledge and second language acquisition and are paramount to my selection of idioms and design of the experiment

The effort to define what qualifies as an idiom is closely tied to research on how idioms are stored, processed, and categorized in our brains. What is clear from the research is that idioms are all prefabricated phrases, meaning that they are memorized as a unit, or lexical item, and not created on the spot (Boers et al 2006). Additionally, in order to be useful for processing and facilitating communication idioms must be generally known and recognizable. These widely known idioms are often employed in speaking as a way to “facilitate fluent language production under real-time conditions” (Boers et al 2006) and gain processing time for constructing our next “chunk of content” (Skehan, 1998 p.40).

Researchers have also shown that idioms are stored as a single, special lexical item that has to be stored separately and processed differently from regular lexical items (Bobrow & Bell, 1973; Gasser & Dyer, 1986; Warren, 2005). These researchers assert, based on their results, that idioms are “opaque invariant word combinations”. Warren offers two definitions that illustrate that idioms are “native like selection[s] of expression” and “that which one has to know over and above rules and words” (p. 35). Thus, all idioms differ from normal lexical items in how they are stored, how they are comprehended, and the rules necessary for producing them correctly. For instance, a native speaker can take the sentence “a” in Figure 2-4 and may alter it using their specialized idiomatic knowledge to make sentence “a” in to sentence “b”, “c”, or “d”.

- a. He laid down the law.*
- b. He was laying down the law*
- c. What he laid down was the law.*
- d. The law was laid down by him.*

**Figure 2-4**

These qualities are what led me to believe that idioms make an excellent candidate for measuring a person's proficiency. The speaker must usually have equivocal knowledge of lexemes, syntax, and morphology to properly employ an idiom. Additionally, because idioms are graded from opaque to transparent, the level of idiom that is able to be processed and produced by a second language learner might be able to indicate what their proficiency is.

### **2.3 Decomposability**

The classification of idioms has been the subject of much research and speculation. The prevailing theory for more than 30 years has been the Idiom Decomposability Hypothesis (IDH) proposed by Gibbs et al (1980) and Nunburg (1978). This theory proposes that all idioms can be sorted into categories based on how transparent they are; or rather if their meaning can be obtained by analyzing the constituent words within the phrase. The topic of idiom categorization is vital to my research question because I need to establish how idioms are categorized in order to be able to effectively select a spread of idioms on the transparency continuum for my research tool. I need this spread in order to get an accurate gauge to use in judging the relationship between knowledge of these idioms and language proficiency.

- a. *Pop the questions.*
- b. *Kick the bucket.*
- c. *John laid down the law.*
- d. *The law was laid down by John.*
- e. *The old man kick the bucket.*
- f. *\*The bucket was kicked by the old man.*

**Figure 2-5**

The classification system put forth by Gibbs (1980) and Nunberg (1978) examines whether or not the expression is “normally decomposable” (p.114), meaning that the constituent parts of the expression give an idea to the overall meaning, or not. They proposed that there were normally decomposable idioms such as Figure 2-5a, where each part of the saying gave an indication of the meaning of the whole expression, and Figure 2-5b where the constituent parts give no idea to the expression as a whole (p. 115). Through their research, Gibbs & Nayak (1985) determined that another feature of normally decomposable idioms is that they are syntactically productive, meaning that Figure 2-5c can be changed to Figure 2-5d without any loss in meaning, whereas Figure 2-5e is not so flexible and cannot be put into the same configuration as seen in Figure 2-5f without losing meaning and is therefore non-decomposable. While these researchers found that some idioms are flexible, diverse, and have varying levels of decomposability, they also see that they behave very much like other lexical items in how they are learned, stored and employed in language. Gibbs et al. proposed that because of this semantic flexibility in decomposable idioms (Gibbs, Nayak, Cooper, 1989) they were more productive and are easier to learn for ESL students (p. 577-578).

*a. Push the panic button*

*b. Carry the torch*

*c. Pop the question*

**Figure 2-6**

As Gibbs et al. further investigated this line of study, another category emerged for idioms that are not syntactically productive and whose constituent parts do not give an idea to the meaning of the whole, but when the constituent words are taken together, they do give an idea to the meaning of the idiom (Gibbs, Nayak, Cooper, 1989). These "abnormally decomposable" idioms include idioms, like Figure 2-6a, but also include more figurative idioms such as Figure 2-6b. In these idioms the individual constituent words are not decomposable like Figure 2-6c, but when taken together, give an idea of the overall meaning.

The research done regarding this theory is plentiful and provides abundant information about retrieval speed and establishes that all idioms are not equally flexible (Callies & Butcher, 2007; Gibbs, 1992; Gibbs & Nayak, 1989; Gibbs, Nayak, Bolton, & Keppel, 1989; Gibbs, Nayak, & Cutting, 1989). However, even though the use of this theory has been pervasive for categorizing idioms, the Idiom Decomposability Hypothesis is not without its critics (Titone & Connine, 2011; Maher 2013). This approach to classifying idioms is somewhat rigid and encounters the same problem with edge cases as trying to define idioms. Gibbs (1985) also points out that there are many researchers who have attempted to explain the difference in idiomatic flexibility being due to syntactic deficiencies (Dong, 1971; Fraser, 1970; Katz, 1973; Newmeyer, 1974; Weinrich, 1969). However, others believe that relationships between syntax, semantic, and pragmatics are the more likely culprit for these differences (Nunberg, 1978;

Strassler, 1982). In any case, the answer is not entirely clear; however, it is clear that the IDH does not have the complete answer for these variations.

- a. Kick the bucket.*
- b. Over the moon.*
- c. Hit the nail on the head.*

**Figure 2-7**

- a. Give up the ghost.*
- b. Lay down the law.*
- c. Draw a line in the sand.*

**Figure 2-8**

In the last decade a few researchers have begun to challenge the decomposable theory in favor of an explanation that is more dynamic and all encompassing (Titone & Connine 2011; Maher, Z. 2013). One of the main concerns with the IDH is its dependence on the intuitions of people sorting idioms into these three categories. While the decisions based on intuition may be remarkably congruent, much of the time the agreement is not universal and intuition is not, itself, infallible (Gibbs et al 1989). An alternative is to consider idioms as being on a continuum rather than in rigid categories. On the one end of the continuum are inflexible idioms like those in Figure 7 while on the other end of the continuum are idioms like those shown in Figure 8.

The researcher Titone has frequently used the IDH as a basis for his own research (Titone & Connine 1994a; Titone & Connine 1994b; Titone & Connine 1999). In Libben & Titone (2008) they ran three experiments, and sighted others, (Abel 2003; Tabossi et al. 2008) that studied several aspects of idioms relating to their categorization according to the IDH. Libben &

Titone found that the results of their own studies had trouble supporting the IDH because the idiom retrieval times in their responses did not correspond to the crisp categories established by the IDH.

Libben & Titone continue to conjecture that another reason the results of their study did not support the IDH may be due in part to a myriad of uncontrolled and unmeasured factors like familiarity or plausibility. Libben also admits that using the IDH proposed by Nunberg (1978) and Gibbs (1980) as a meter against which to measure their experiment may have been a mistake considering the inconsistency in the IDH results. Especially since Libben & Titone had other theories to choose from when conceptualizing idioms, such as that proposed by Geeraerts (1995). Maher (2013) reached the same conclusion regarding the IDH citing the work of Tabossi et al. (2008) which showed a systematic failures of the IDH.

Tabossi conducted a serous of experiments that demonstrated that the use of intuition for the basis of categorizing idioms is unreliable and was consistent on in a small number of idealized cases. Tabossi was also unable to find any evidence of syntactic flexibility being affected by decomposability when recreating the (1994) Gibbs and Nayak (1989) (Experiment 2). As intuition judgments was the basis for the IDH these results throw substantial doubt on the IDH and convince me to find another measure.

The IDH was an admirable start at quantifying idioms and providing a conceptual system by which to categorize them. However, I am convinced that the opposition raised by Tabossi et al (2008), Titone & Connine (1994), and Maher (2013) that this system is out dated and the data does not support its validity. Therefore I will be using the gradation system put forth in Libben & Titone (2008) (see appendix A). This system doesn't seek to put idioms into a few sharply

defined categories. Instead, they simply assign a number 1-5 on a variety of characteristics based on the results from their study. Although I am rejecting the longstanding benchmark of the Idiom Decomposability Hypothesis, I feel that accepting the Libben & Titone system will provide a much better basis for categorizing idioms for the purposes of my study as it gives more reliability against which to measure my results.

#### **2.4 Idioms for ESL**

While research on defining, classifying, and processing idioms is plentiful, there is comparatively very little research on idioms as they relate to second language learning. Some of the earliest research on idioms and second language acquisition simply concluded that idioms were important for non-native speakers of English in order for them to sound fluent (Palmer, 1925; Clarke & Nation, 1980). However, in recent years, more research has been done on their importance to L2 proficiency. Hinkel (2013 p.16) suggests that to “increase learners linguistic repertoire they should devote a great deal of attention to...idioms...and even their appropriate pronunciation and intonation”. Many others have also shown the importance of using Idioms in regular education for ESL learners (Foster, 2001; Howarth, 1998; Nattinger and DeCarrico, 1992; Wray, 2002).

#### **2.5 Idioms and transfer**

Irujo is one of the few researchers who directly addressed idioms as they pertain to second language acquisition found that interference between the L1 and L2 was prominent (Irujo, S. 1986). She explained that if idioms are more identical in form and meaning in both the first and second languages they become easier for the ESL learners to comprehend and produce. As they begin to differentiate, it becomes more difficult for speakers to understand and produce

them correctly. Usage of idioms that have no resemblance between the L1 and L2 is extremely difficult, and most speakers judge these kinds of idioms as ungrammatical in the second language and simply avoid using them. It is equally difficult for the speakers to comprehend these idioms as the speaker usually tries to impose the grammar and syntax of the L1 on the target language (Gass, 1979).

Starting in the 50's researchers assumed that interlingual transfer was the most important factor in second language (Poltzer, 1965). They also believed that negative transfer could be overcome by simple repetition and reinforcement; which included idiom understanding and usage. However, by the mid 60's the emergence of generative grammar shifted the paradigm and established a new field of psycholinguistics (Irujo, S., 1986). This change brought with it a reduced focus on repetition and positive/negative transfer in language acquisition; instead, there was a focus on language as a creative process that needed learning strategies and an understanding of the target language structure for reducing errors in ESL students (Richards, 1974; Dulay & Burt, 1975).

Since then there has been some back and forth on the importance of interference being an important factor in second language acquisition, but researchers have found that interference is, in fact, a substantial influence on the language learning process (Irujo, S., 1986). While contrastive analysis, reinforcement, and habit formation have been disproven as a "magic bullet" to overcoming the obstacles when learning a second language, subsequent research does support a moderate version of this theory. Oller and Ziahosseiny (1970) propose that there is more difficulty in language a learning when the difference between the L1 and the L2 is only modest, as opposed to when the languages are very dissimilar due to interference.

Substantial research has been done that supports the idea that interference and positive and negative transfer between the L1 and L2 greatly affect the difficulty in language learning, (Sajavaara, 1976; Whitman 1970). In an experiment by Gass (1979) participants were able to use their L2 to correctly produce idioms that had the same form and meaning as the target L1. However, when the idioms began to become more and more different the participants were unable to produce a correct idiom in the second language. This demonstrates how much of an affect interference from the L1 has on the L2 when producing idioms.

Jordens (1977) and Kellerman (1977) conducted two studies in the Netherlands on idioms as they relate to second language acquisition. They asked second language learners to decide if correct and incorrect sentences that contained idioms were grammatical. In both of the experiments the participants judged the sentences to be ungrammatical even when the idiom in the sentence had an equivalent in their first language. The results of these studies showed that these participants were unwilling to use idioms as the language specific items they are. This showed how difficult idioms are to use and how resistant speakers can be to employ them due to this difficulty. However, these studies did not deal with production or comprehension of idioms, only with recognition. Additionally, they only used idioms as they relate to grammaticality judgments.

As I begin discussing the next paper, I must first define a few more items that are relevant here. This paper will discuss positive transfer which is when a speaker is able to use experience and information from their first language in helping them process information in the second language. Negative transfer is the opposite effect when the first language interferes with processing information in the second language.

In her paper, Irujo S. (1986) conducted an experiment where participants were tested on recognition, comprehension, recall, and production of idioms. She gave a four part test to twelve Spanish L1 participants that were advanced English speakers. These participants were tested on completing an idioms with a missing word, translate an idioms from Spanish to English, defining the idiom, and lastly, to select the correct meaning of the idiom in a multiple choice test.

The participants were able to produce identical idioms better than similar or different idioms in the target language. The participants were also able to comprehend identical and similar idioms equally well and both better than different idioms. In comprehension of similar and identical idioms they demonstrated positive transfer from the L1 while also showing negative transfer from the L1 on production of similar and different idioms. Positive transfer is using information from the L1 to help correctly process information in the L2, while negative transfer is using information from the L1 incorrectly process information in the L2.

Irujo's study demonstrated that second language learners use their L1 to process idioms in the L2 which let them correctly identify, comprehend, and produce idioms that are the same in the L1 and L2. Idioms that are similar between the L1 and L2, but not identical, are just as easy to comprehend as the identical idioms, but are more difficult for language learners to produce. This shows that the interference of the L1 help the learner process these idioms in the second language, but also produces negative transfer when trying to produce the same idioms. Finally, the different idioms that have no related counterpart in the L1 are both difficult to comprehend and produce.

In Irujo's study, the subjects had measured the student's TOEFL scores before testing to ensure that all the participants met a minimum level of English proficiency, but not for the

purposes of comparing the standardized test score to the results of the experiment. Therefore, while the study was very enlightening regarding learning strategies and L1 – L2 for interference for idioms, it was not able to examine language proficiency by comparing TOEFL scores and idiomatic knowledge results. The author acknowledges that a study of level of proficiency compared to idiomatic knowledge should be conducted with more diverse levels of English learners than were used.

Irujo (1986) surmises that these results can be taken to mean that overtly teaching idioms to ESL students that have a similar meaning in their L1 would be possible and beneficial. The students can also be instructed to use their knowledge of idioms in their L1 to comprehend and produce idioms in the target language. However, Irujo points out that the effect is diminished as the cultures of the two languages diverge, as the common idioms are likely to be proportionally dissimilar. This effect is compounded in ESL settings as the student body is likely to be comprised of several different L1s with varying areas of overlap with English. However, the teacher can still instruct the students on the how to recognize and process idioms in the L2 to obtain their non-literal meanings in order to increase the students' proficiency level and to sound more fluent.

## **Chapter 3:**

### **Experiment Methodology**

This chapter is about the experiments that I conducted to determine if linguistic proficiency has a relationship with idiomatic knowledge. In this chapter I will discuss the methods for two experiments. I begin by explaining about the participants and design used to conduct each experiment. Finally I will discuss the methods for data analysis and any shortcomings of the design.

In order to generalize my results to ESL learners as a whole, my two studies use two target populations of general ESL learners and those enrolled in intensive English programs. The first study sampled intensive ESL learners and the second study sampled a larger body of ESL learners. I also believe that this will help substantiate any statistical data by having two different pools of candidates from which to draw data and compare results/

#### **3.1 Study One – Intensive English Program**

In this experiment I test a group of participants on their idiomatic knowledge by administering a survey to them. These participants provided their demographic information for analysis against their performance. My goal is to see if their performance on the idiom survey is similar to the results on their standardized tests, thus showing that idiomatic knowledge is representative of language proficiency.

##### **3.1.2 Participants**

For this study I intended to draw as large of a participant group as possible in order to reach a minimum of one hundred results. However the participants all needed to have taken a

standardized test to determine their language proficiency so that I was not comparing results across disparate testing methods. Therefore, I used all 141 students at the Brigham Young University English Learning Center (ELC). The participant's ages ranged from 18-45 with a median age of 24. Participants were approximately equal numbers of male and female and represented a variety of countries and a total of 10 different native languages shown in Table-1 below.

**Table 1: Composition of Participants for Study One**

Country	Chinese	French	Japanese	Korean	Mongolian	Portuguese	Russian	Spanish	Thai	Ukrainian
<b>Male</b>	3	0	2	9	0	3	0	17	0	0
<b>Female</b>	6	1	1	2	1	4	2	16	2	1
<b>Total</b>	9	1	3	11	1	7	2	33	2	1

The participants were sorted into class levels 1-5 depending on the results of their LAT scores. The LAT was designed to so that the scores would correspond to the various ACTFL levels, allowing the students to be sorted into the different class levels. The LAT is a test that has been carefully constructed by the administration at the ELC at BYU. The test itself is a reliable testing instrument.

This method of gathering participants didn't allow for random sampling, but as there were a large number of participants with a wide variety of languages, countries and ages I believe the pool of participants is a good sample of my target population of ESL learners enrolled in intensive English programs. This was the most practical way of getting a large sample of intensive ESL students with the same standardized testing scores. Although the participants were not randomly selected, the ESL program participants varied widely in their demographics and

offer a fairly diverse population for the sample size. As the target population is intensive ESL learners, I feel that the diver the cooperation of the ELC in granting me total access to the students and having them all in the same building were two very important considerations.

An important note is that the students involved in the test have all been admitted to the ESL program based on the standardized test used by the ELC known as the Language Aptitude Test (LAT) test. This is integral to my process of computing my results because it will be the standard against which I will measure their performance on my survey. Using this measure, along with data from their demographics, I will see if my test can be as accurate at predicting, or at least evaluating their language proficiency as the standardized test.

### **3.1.3 Measure**

In order to construct a suitable multiple choice test for gauging the idiomatic knowledge of the participants I needed to first obtain realistic, natural and plausible distractors for the content. Therefore, I did a preliminary experiment where I selected 46 verb phrase idioms and presented them to three classes of beginning ESL learners at Utah Valley University comprising 60 respondents. The respondents were randomly given a list of 22 of the idioms and were asked to simply read down the list and provide a definition for each idiom like the one seen in Figure 3-1.

What does, “He lost his nerve” mean?

Answer: \_\_\_\_\_

**Figure 3-1**

I read through all the responses and compiled the answers for each idiom so that all similar answers were tallied. Some idioms had only one answer (allowing for variation) while others had a six or more and still other answers were unintelligible. I threw out idioms that had fewer than three different, intelligible, incorrect responses. Finally, I was left with 24 idioms that had enough plausible distractor answers.

From this list of idioms I constructed a multiple choice questionnaire on Qualtrics (appendix B). The survey asked the participants to provide their student number so that they would remain anonymous to me, but the school staff was able to provide the testing scores and demographic information; age, gender, nationality, native language, and entry test score. The survey also recorded the amount of time each participant spent on each page.

The survey was constructed with six blocks of questions, each block being comprised of four questions, totaling 24 questions to each participant. Each participant was only presented with three of the six blocks, randomly assigned by the Qualtrics programming, so that each block was seen about the same number of times. Within each block, the four questions was presented in random order so that even if two students had the same block the questions within that block would not be in the same order. Each question was presented in an identical format; asking the same question with only the idiom being different from questions to question. A sample question is shown below in Figure 3-2.

What does, "Lay down the law" mean?

- Put down something heavy
- Obey a police officer
- To give commands
- Commit a small crime

**Figure 3-2**

#### **3.1.4 Procedures:**

When I took this newly constructed Qualtrics survey to the intensive English students at BYU I wanted the entire student body to participate. So, the administration sent two emails to the teaching staff instructing them to email the survey link to the students and ask them to complete the survey. The teachers did not offer extra credit and the participants were not compensated in any fashion by me or the institution. The participants were able to take the survey at their leisure with-in a two day time period with no time limit. The participants were instructed by their teachers in class to take the survey by themselves without the help of friends, roommates, or the internet, though there was no overt monitoring to ensure that they followed those instructions.

#### **3.1.5 Data Gathering**

The survey was administered to 141 participants, but I had some participants that either did not answer all the questions that were shown to them so I discarded these results as invalid. Then I removed any respondents who provided an incomplete student number because their demographic information was unavailable without the complete number; further dropping the number of results to 93. Lastly I removed any participants who were not tested on the standardized test during the most recent semester because their results would have been more

than 16 weeks old compared to the 8-10 weeks of all the other students. This brought my final total of usable results to 72.

I used the demographic data, the standardized test scores, and survey results in my calculations to answer the research questions to this study. I first organized the data from the responses by ordering the respondents by their randomly generated ID and totaling up the number of correct answers they had on the survey (Table 1). I then converted all the constructs to SPSS, including dummy coding the respondents' Native Languages in order to run bivariate correlations. Finally, after checking the data for a normal distribution, I ran a Pearson test to evaluate the strength of the correlation.

The number of participants whose results I was able to use were overwhelmingly in the level 3 classes; which corresponds to the ACTFL Intermediate-low, Intermediate-middle, and Intermediate-high levels. The results to the survey varied with all levels, but as you in see in Table-2, the level 1 class varied the most widely.

**Table 2: Results of Study One**

# correct	Lvl 1	Lvl 2	Lvl 3	Lvl 4	Lvl 5	Lvl 6
1	1	0	1	0	0	0
2	0	0	3	1	0	0
3	0	0	2	1	0	0
4	2	0	3	3	0	0
5	0	0	3	2	0	0
6	0	0	5	9	1	1
7	0	0	5	2	0	0
8	0	0	5	6	2	0

9	0	0	5	1	1	1
10	0	0	1	0	0	0
11	1	0	1	1	0	2
# Subjects	4	0	34	26	4	4

## 3.2 Study 2

In order to see if the trend in the data could be substantiated on a larger scale and help substantiate the results of my first study I conducted this follow up study. In this case I targeted an even larger set of participants to help even out any data anomalies that might occur and confirm the general relationship between language proficiency and idiomatic knowledge.

### 3.2.2 Participants

I chose to use Amazon’s Mechanical Turk (MTurk) for quick access to large groups of participants at a low cost with a high diversity. MTurk workers have been shown to be as reliable as well paid participants with experience into the field (Mason & Suri, 2011). MTurk workers were also shown to be comparable to laboratory subjects in quality and reliability (Paolacci, Chandler, & Ipeirotis, 2010). The risks and downsides for using MTurk is the same for using any online survey, namely that the test population isn’t representative of any geographic population. While that is true, I am looking to generalize these results to ESL speakers as a whole and not to any geographic or even demographic population.

Low quality is also a concern, but I was able to set the parameters for who was able to qualify to take the survey so that only those people who had previously completed one hundred other tasks and had completed them with a 95% approval rating. This means that each participant could only have 5% of their tasks rejected by administrators for any reason. I choose these

cutoffs because they were recommended by MTurk for getting reliable participants, but not setting the standards so high as to have trouble attracting participants.

I received 640 results at first, but had to discard down to 361 to eliminate people who identified their native language as English, who did not finish the survey, or did not include their English proficiency level. I then had to double check all the remaining responses for duplicate MTurk IDs to make sure that each response was unique which brought me down to 340 actual results. The remaining respondents were composed of 35 languages (though most had less than 5), shown below in Table-3. The participants were compensated with one dollar for the completion of the study, but had only a maximum of 20 minutes to complete the survey.

### **3.2.3 Measure**

I used the same survey in this study as I did in the previous study with some changes to how it was delivered. Instead of giving each participant randomized blocks of questions to discourage sharing among participants, which is preferred, I simply had all of the idiom questions listed on one page for them. I made this change because MTurk suggested that single page surveys are received better by the workers when they can see the entirety of their task instead of only receiving portions of it. While I prefer to give randomly ordered surveys, I believed that with such a large group it would not impact the data.

I asked the participants to self-evaluate their English level on a scale of one to ten. Self-evaluation is problematic with accuracy because it has been shown that subjects who evaluate their own ability can over or under estimate their ability. However, some research on the subject indicates that subjects can be fairly accurate on tasks like writing and speaking. Having specific tasks related to their evaluation is helpful and the task asked of them, evaluating an idiom, is

most closely related to speaking (Strong-Krause 2006). Additionally, this is reflective of their feelings of how confident they are in their proficiency, which does impact ability as well as how likely they are to employ their language skills regardless of what their actual, tested proficiency level is. Additionally, this kind of self-evaluation is very practical for the purposes of this study as I needed a quick and efficient way to gauge ability. I have shown the breakdown of the demographics below in Table-3.

**Table 3: Number of results by Language for Study 2**

Telugu	Tamil	Russian	Spanish	Malayalam	Hindi	Bengali	Others *
19	98	7	39	48	54	8	67
<p><u>*All of the 28 Languages below had less than 5 respondents and totaled 67:</u>            Yoruba, Vietnamese, Urdu, Thai, Tagalog, Swahili, Serbian, Slovenian, Sourashtra, Romanian, Polish, Punjabi, Odia, Marathi, Mandarin, Macedonian, Korean, Kannada, German, Gujarati, Filipino, Finish, French, Chinese, Creole, Bulgarian, Assamese, Arabic</p>							



## Chapter 4: Data Analysis

### 4.1 Study One – Intensive English Program

I first ran a regression analysis on all constructs that I collected; age, L1, gender, number correct on the survey (henceforth called number correct), the full LAT score, and the component testing parts: reading, writing, and speaking. In my first analysis of the constructs, I was able to conclude that language, gender, and age played no significant roll on number correct or their language proficiency score based on the non-significant correlations they produced. I considered all results where  $p < .05$  to be significant. I removed all the insignificant constructs from further evaluations and comparisons and continued my analysis with the remaining constructs: LAT Score, The three component scores of the test, and the Number Correct on the idiom survey.

**Table 4: Correlation of Constructs from Study One**

		FullLATScore	# correct	ReadingFA	WritingFA	SpeakingFA
FullLATScore	Pearson Correlation	1	.287*	.740**	.890**	.840**
	Sig. (2-tailed)		.015	.000	.000	.000
	N	71	71	71	71	71
# correct	Pearson Correlation	.287*	1	.153	.176	.284*
	Sig. (2-tailed)	.015		.201	.141	.016
	N	71	71	71	71	71
ReadingFA	Pearson Correlation	.740**	.153	1	.562**	.551**
	Sig. (2-tailed)	.000	.201		.000	.000
	N	71	71	71	71	71
WritingFA	Pearson Correlation	.890**	.176	.562**	1	.632**
	Sig. (2-tailed)	.000	.141	.000		.000

	N	71	71	71	71	71
SpeakingFA	Pearson Correlation	.840**	.284*	.551**	.632**	1
	Sig. (2-tailed)	.000	.016	.000	.000	
	N	71	71	71	71	71

\*. Correlation is significant at the 0.05 level (2-tailed).

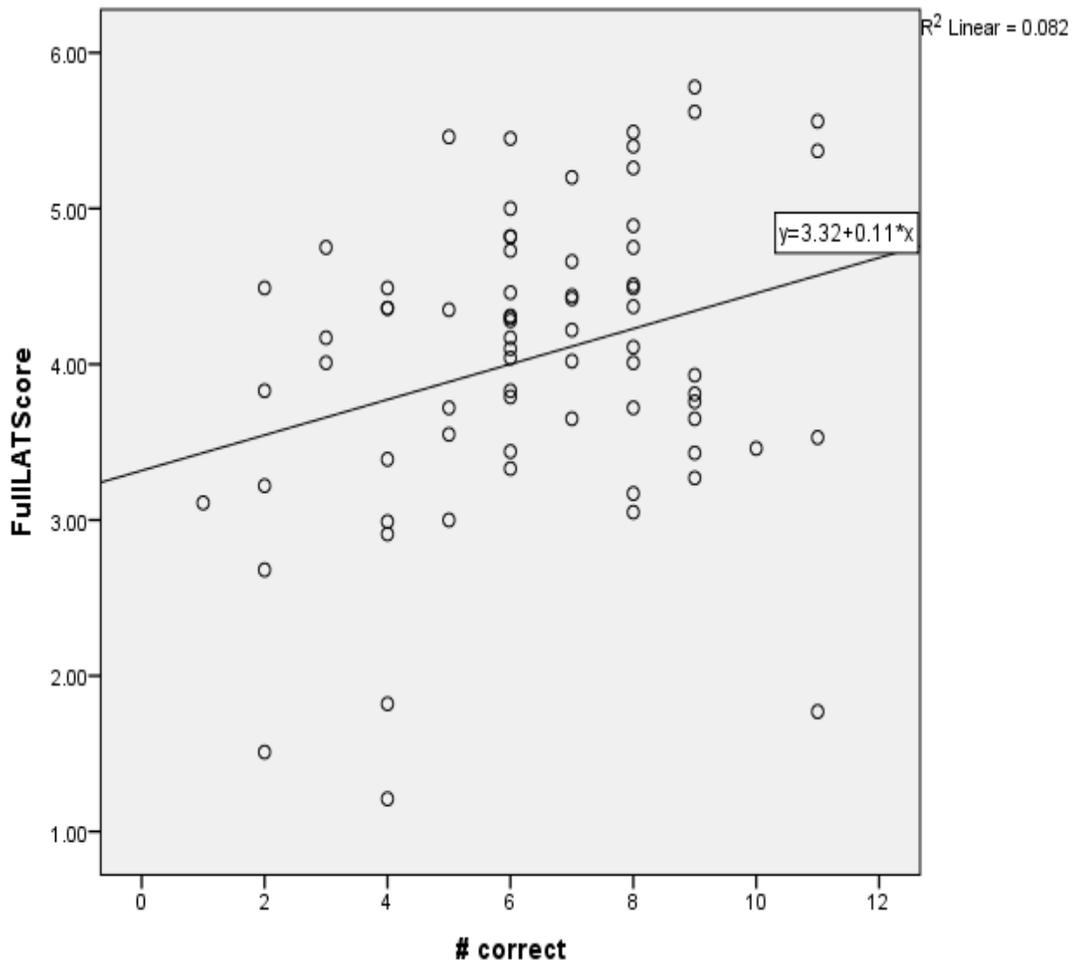
\*\*. Correlation is significant at the 0.01 level (2-tailed).

While I removed the three individual scores of the LAT from the relevant variables and proceeded with just the overall score, I feel it important to note the relationship between those three scores and the number correct to be interesting. The score for reading is least significant to the number correct, writing is more so, but still non-significant, and speaking is statistically significant to the number correct, almost the same score as the full LAT score.

I then ran a regression analysis on the constructs: full LAT score, number correct, reading score, writing score, and speaking scores as shown in Table-3. This table also shows that the two constructs that were most important to my thesis question, number correct and LAT Score, had a significant relationship. Once I determined that that data was distributed regularly, I decided to run a Pearson correlation on the LAT score and the number correct. The results, show in Table-4, show a significant relationship between the two constructs.

The wide spread of the data points produced a relatively low R squared which calculated a percent of variance of only 8.06%, as can be seen in Figure 4-1. The data produces a wide spread, though fairly regular, graph, but there are some more extreme outliers that may be affecting the results.

**Figure 3: Loess Regression line**



## 4.2 Study Two

In the follow up study I did with the MTurk participants, my data was not as complex as my first study. In this study I gathered a much larger sample of participants in order to be able to rely on the size of data and the correlations rather than the careful sampling that was done in the first study. As this was a follow up study, I have already established that the number correct and the subjects English Proficiency were the only two variables that were statistically significant. Therefore, I constructed the survey around only these two variables to find the strength of the correlation in this second demographic.

**Table 5: Bivariate Correlation from Study Two.**

		SelfEnglishScore	# correct
SelfEnglishScore	Pearson Correlation	1	.207*
	Sig. (2-tailed)		.000
	N	361	361
# correct	Pearson Correlation	.207*	1
	Sig. (2-tailed)	.000	
	N	361	361

\*. Correlation is significant at the 0.05 level (2-tailed).

As with the first set of data I first checked for a normal distribution and then ran a Pearson test to see the strength of the correlation, if any, between the self-ascribed English level and the number of correct answers. I found that the correlation was highly significant as shown in Table-5 with a p value far more significant than  $P > .05$ . However, as in the first study, the relationship between the constructs is modest. The results from the two studies were nearly identical and established a correlation between the two constructs for further investigation.

This data set is different in the make-up of the participants though. The people taking this survey were unlikely to be enrolled in an intensive English program and are very likely to be living in another country. Thus, this set of data provides a very different demographic than the first study. Having two dissimilar groups of ESL learners produce very similar outcomes on the same survey substantially helps validate the results shown above.

## **Chapter 5: Discussion**

This study set out to explore the relationship between language proficiency and idiomatic knowledge. With these aims in mind, I constructed a set of surveys to obtain authentic answers to act as distractor answers in the experiment survey. The resulting relationship was statistically significant and did establish a definite relationship between the two constructs in question. The subsequent data supported my thesis in establishing this relationship and also gave evidence of where to go in future studies.

While the strength of the correlations between number correct and the three sections scores of the LAT test is not very strong, as seen in Figure 2, the increasing correlation pattern between the constructs provides very interesting information for idioms. The data suggests that reading skill and being able to correctly identify idioms is very unrelated, but writing skill and being able to identify idioms is somewhat correlated, albeit very weakly. The speaking portion of the LAT test and idioms identification is very highly correlated, even though the data shows wide variability in the individual responses. What this means is that the relationship between idiomatic knowledge and standardized test scores definitely exists, but as this was only an observational study, it is still unknown how strongly each of the variables influences the other.

The escalating correlations between the three parts of the LAT test and idiomatic knowledge are somewhat to be expected as idiom usage, according to some, is primarily a speaking strategy to buy time for us to process more information, organize our thoughts, and convey specified meaning in short, universally understood chunks (Lakoff & Johnson 1980). While idioms can, and are, used in reading/writing, they are employed for different reasons in

writing than in speaking as there is no need to slow down for processing needs like there is when speaking.

While the correlations between idiomatic knowledge and language proficiency are promising, and the correlation is strong, the wide spread data points indicate a wide variability on an individual level. The positive correlation that was confirmed between speaking proficiency and idiomatic knowledge is another promising and unexpected, albeit unsurprising, result that helps show where to go next in this line of questioning. I believe that with modified test administration it would be possible to correct for several factors that may have led my initial data astray and produce stronger, more reliable results. Additionally, a larger group would also allow for a more complete sampling of the lower and higher echelons of the ELC program, as the bulk of students were in the mid-range level.

The disparity in the results from the survey, “24 Idioms as a Measure of Proficiency” indicates that some of the student’s exposure to idioms does not reflect their level of English. Thus, it is entirely possible that someone who has low level of English will know many idioms, and someone who has high level of English knows very few. Even more possibilities open up if we look at different types of learners. Some people are heritage learners, some are visual, tactile, passive or active learners. Some people have learned a language through immersion and some have never left their country of origin and still speak competently, at least in academic registers and professional registers. These factors may make a great deal of difference in how they process and deal with idioms they encounter.

This kind of variance comes from the methods that people use when learning English. Students who focus on an academic approach with classes, instructors, and approved materials

will likely have much less exposure to idioms as they are primarily a function of communication in the lower registers between friends, family, and other close relationships. If another student places a great importance on pop culture, movies, books, the internet, and other materials, they are likely to have a much greater exposure to idioms or at least the culture of the target language. Thus the difference in idiomatic knowledge can differ greatly. However, overwhelming anecdotal evidence suggest that the normal process of language learning does not contain an abundance of exposure to the target language's idioms or culture in general.

## **5.1 Limitations**

This study was limited by the number of usable results. This is the most important potential limitation to my study as I feel a large and diverse group of participants is required for the most reliable results. Also, the lack of supervision may have been problematic as there was no oversight to ensure that each participant did not seek help from others or the internet. The standardized testing scores used in evaluating each participant were between eight and 10 weeks old which may not accurately reflect the language growth that occurred in the intervening time, but I felt was an acceptable time lapse.

With more than 140 responses, I expected to have a larger data set, but the dramatic reduction of nearly 50% of usable responses resulted in a data set that was smaller and significantly less varied; especially in the lowest two levels of the intensive English program. While I expected that I was overestimating the number of respondents needed, I did not account for the students providing incomplete student numbers. While I could have used the respondents with older test scores, I felt that they no were no longer accurate enough and could invalidate any correlations drawn from the data.

As the entirety of the survey was less than ten minutes and the participants were not going to be graded on their results I thought it safe to trust them to not supervise them. I believe that the vast majority did not receive help, but doubt is cast on the level 1 student with eleven correct answers – as his time on each page was substantially higher than any other participant, but I cannot prove that he use outside help and that the extra time isn't attributable to something else – such as slow reading skills. Regarding the test scores, I feel that the students did not progress too far to make their tests outdated in just ten weeks. Additionally, even if the students did progress quickly, they would all have progressed at a similar level; meaning that the scores would still be proportionally relevant to the participants.

The survey presented in the second study was intended to be presented in a random fashion similar to the first, but with the complication of the administration of the survey that was no longer possible. I would like to have seen the subjects also asked a few short evaluation questions on their self-assessment related to a particular task instead of a single scale of ability. The research done by Strong-Kruase (2000) and Brown, Dewey, & Cox (29

In the future, all three issues would be easily dealt with by simply having the students complete the survey as part of their beginning of semester or end of semester testing; where the conditions are already controlled. This would ensure that the student numbers would be complete, there would be no outside help and the testing scores would be as recent and applicable as possible.

## **5.2 Suggestions for Further Research**

The next logical step would be to run a set of experiments to see what kind of predictive power these two variables have on each other. In exploring this area further it would behoove us

to know if additional instruction in strategies for recognizing and comprehending idioms increases a learner's language proficiency. I believe that this would be beneficial to many areas of learning, but primarily to speaking skills.

Leaving the narrow confines of this study, there are any number of interesting and promising research topics for this line of questioning. One area of particular interest is perceived proficiency. It is clear that idiom usage helps people communicate, but could additional instruction help the speaker sound more fluent than their counterparts of the same skill level? Could we compare students of similar levels, but teach one group a handful of idioms to incorporate into speech and see if native judges rate the recorded speakers as more fluent than their non-idiom using counterparts?

I also am interested in the correlation between reading, writing and speaking on idiomatic knowledge. From the data gathered in the first study that these three scores are related at varying levels of significance to idiomatic knowledge. I would like to see if learning idioms in the different environments would yield different results. For instance, does learning about idioms as a writing skill produce any positive affect on language proficiency as compared to learning idioms in speech? I believe that there is still much to do on this particular area of research.

## **5.2 Conclusions**

This study set out to show if there is a relationship between idiomatic knowledge and language proficiency. There is little direct research or evidence in this topic and none that deliberately state that there is a relationship or how it might work. However, building on the work that has been done and drawing from the areas of idiom research and language proficiency measures, I executed an experiment to explore what kind of relationship existed. The experiment

was a simple survey given to more than one hundred ESL students of varying proficiencies, according to standardized testing, and to see if the number of correct answers was commensurate with their estimated skill level.

It is clear that there is a relationship between these construct, but the strength of that relationship is modest at best. There is a relationship between language skill and idiom knowledge, but it does not seem to be strong enough to be predictive or indicative of language skill except that as one generally become better at language they are also better at idioms. The second set of data I collected corroborated this finding and showed that while the two constructs are strongly correlated the relationship between them is not very as strong.

My data also produced an interesting relationship between idiomatic knowledge and the different ways we process language. I showed that while speaking skill and idiom knowledge was strongly correlated, writing was less significant and reading seemed completely unrelated. Not only have I been successful in supporting my hypothesis, I have produced several interesting data points that need further investigation.

## APPENDIX A: Libben & Titone 2008 Idiom Results

Idiom	Familiarity	Familiarity SD	Meaningful	Meaningful	Literal Plausibility	Literal Plausibility SD	Verb Relatedness	Verb Relatedness SD
She changed her mind	4.97	1.24	4.97	1.25	2.03	1.20	4.70	0.25
It crossed his mind	4.83	0.46	4.97	0.18	1.40	0.89	2.13	1.07
He cracked a joke	4.81	0.85	4.69	0.26	0.81	1.27	1.01	1.25
They spread the word	4.80	0.41	4.97	0.18	2.07	1.31	4.00	1.07
She drove him nuts	4.77	0.77	4.97	0.18	1.40	0.97	2.07	0.94
He hit the books	4.73	1.22	5.00	1.19	3.83	1.12	1.63	1.38
She hit the sack	4.73	1.17	4.97	0.83	3.97	1.12	2.13	0.74
He threw a fit	4.73	1.21	4.97	0.75	1.20	1.14	1.97	1.24
He got the picture	4.67	1.59	4.93	1.79	4.33	1.13	3.23	1.11
He took a hint	4.67	1.38	5.00	1.32	1.43	1.25	2.47	1.18
She kicked some butt	4.63	0.84	4.97	0.18	3.70	1.24	2.40	0.48
She pushed his buttons	4.63	0.61	5.00	0.00	2.67	1.42	2.83	1.29
She gave her word	4.60	0.72	5.00	0.00	2.33	1.45	3.23	1.17
It slipped his mind	4.60	0.86	5.00	0.00	1.70	1.06	3.00	1.02
She spoke her mind	4.60	0.85	4.97	0.26	1.90	1.27	4.70	0.65
They tied the knot	4.60	0.81	4.97	0.18	4.17	1.23	2.60	1.25
She got the message	4.57	0.66	4.80	0.25	4.37	0.84	3.50	1.17
He used his head	4.57	1.01	4.90	0.68	3.50	1.21	3.13	1.37
She called the shots	4.50	0.63	4.90	0.31	2.77	1.36	3.30	1.24
It hit the spot	4.50	1.13	4.80	0.92	4.07	1.38	1.77	1.32
He popped the question	4.50	0.97	4.97	0.18	1.73	1.28	2.17	1.05
She used her head	4.47	0.86	4.77	0.40	3.90	1.43	3.11	1.11
He broke the ice	4.43	0.97	4.97	0.25	4.20	0.94	2.83	1.19
She had a ball	4.43	0.63	4.97	1.34	3.50	1.47	3.03	0.88
She packed her bags	4.43	0.77	4.93	0.25	4.80	0.66	3.63	1.25
He swallowed his pride	4.43	0.77	4.80	0.76	1.07	0.25	2.17	0.91
She broke her word	4.40	0.97	4.93	0.18	1.43	1.27	2.97	1.21
He lost his cool	4.40	0.92	4.93	1.40	1.53	0.94	2.60	0.90
She racked her brains	4.40	0.97	4.90	0.40	1.47	0.73	2.40	0.93
He lost his nerve	4.37	1.37	4.87	1.49	1.50	1.24	3.97	1.00
He kicked the habit	4.30	0.81	4.93	0.18	1.47	1.09	2.31	1.28
He cleared his name	4.17	1.32	4.87	1.24	2.30	1.08	4.03	0.92
He hit the roof	3.97	1.41	4.53	1.36	3.80	1.19	3.20	1.01
She changed her tune	3.80	0.18	4.37	0.18	4.00	1.35	4.93	0.92
He stole her thunder	3.77	1.45	4.67	0.88	1.33	0.92	4.27	0.83
He threw in the towel	3.67	0.64	4.70	0.18	4.13	0.76	2.10	0.93
She got the ax	3.37	0.86	4.07	0.61	3.87	0.96	2.73	1.11
She took the heat	3.33	0.61	4.40	0.00	2.07	0.86	4.10	0.90

She faced the music	3.23	1.28	4.27	1.08	2.80	1.42	4.13	0.68
They shot the breeze	3.17	1.56	3.80	1.71	1.50	0.68	1.40	0.86
He dropped the ball	3.03	1.25	3.60	1.59	4.80	0.61	3.04	1.00
They sang her praise	2.93	0.91	4.30	1.12	3.90	1.03	3.27	1.14
He learned the ropes	2.87	1.57	3.20	1.68	1.20	1.31	4.56	1.29
He showed his teeth	2.70	0.89	3.53	1.28	4.70	0.31	2.03	1.24
They mended their fences	2.24	1.46	3.48	1.43	4.20	1.10	4.23	0.86
He greased the wheels	2.10	0.94	2.08	1.41	4.70	1.06	2.88	1.27
He split a gut	2.03	0.89	2.90	1.56	2.37	1.00	1.57	0.77
He saved his bacon	1.97	1.22	2.93	1.72	3.23	1.28	3.90	1.32
He showed the flag	1.97	1.42	2.40	1.41	4.90	0.70	2.33	1.13
He sugared the pill	1.97	1.30	2.60	1.50	3.60	1.16	3.93	0.87
She dusted his pants	1.93	1.20	2.23	1.38	3.93	1.26	1.77	1.04
She raised the devil	1.90	1.09	3.27	1.46	1.43	0.77	2.43	1.04
He got a toehold	1.87	0.94	2.50	1.63	2.43	1.28	2.38	0.98
She skimmed the cream	1.86	0.83	2.31	1.20	4.13	1.17	2.20	1.45
He got his goat	1.77	0.94	2.40	1.43	3.60	1.45	2.00	1.12
He curled the lip	1.73	1.01	2.13	1.11	3.83	1.21	2.13	1.33
They sat on thorns	1.73	0.83	2.57	1.33	4.10	0.96	2.57	1.48
He rushed his fences	1.70	0.79	1.70	0.95	1.83	1.23	4.03	1.15
He changed his neck	1.67	0.84	2.77	1.48	1.33	0.55	4.00	1.14
He lost the thread	1.67	1.00	2.10	0.37	4.47	1.11	3.77	1.33
She took the veil	1.67	1.54	2.23	1.55	4.13	0.84	2.27	1.06
He fanned the breeze	1.60	0.97	2.03	1.19	2.73	1.39	1.57	0.77
She had a lark	1.57	0.94	2.17	0.18	2.67	1.66	1.72	1.25

## **APPENDIX B: Survey for Studies One and Two**

My name is Kyle Vanderniet, I am a graduate student at Brigham Young University and I am conducting this research under the supervision of Professor David Eddington, from the Department of linguistics. You are being invited to participate in this research study of idioms as a measure of ESL proficiency. I am interested in finding out about whether idiomatic knowledge is as good of a measure of English proficiency as standardized testing.

Your participation in this study will require the completion of the following survey. This should take approximately 15 minutes of your time. Your participation will be anonymous and you will not be contacted again in the future. You will not be paid for being in this study. This survey involves minimal risk to you. The benefits, however, may impact society by helping increase knowledge about current and future testing parameters that may change how we measure English proficiency.

You do not have to be in this study if you do not want to be. You do not have to answer any question that you do not want to answer for any reason. We will be happy to answer any questions you have about this study. If you have further questions about this project or if you have a research-related problem you may contact me, Kyle Vanderniet at [clyguy@gmail.com](mailto:clyguy@gmail.com) or my adviser, David Eddington at [eddingon@byu.edu](mailto:eddingon@byu.edu).

If you have any questions about your rights as a research participant you may contact the IRB Administrator at A-285 ASB, Brigham Young University, Provo, UT 84602; [irb@byu.edu](mailto:irb@byu.edu); (801) 422-1461. The IRB is a group of people who review research studies to protect the rights and welfare of research participants.

The completion of this survey implies your consent to participate. If you choose to participate, please complete the following survey and submit it when you are finished. Thank you!

- I Agree
- I do NOT Agree

What is your 9 digit school number?

What does, "Under the weather" mean?

- To be cold outside
- To be sick
- To be angry
- To be worthless

What does, "Worth its weight in gold" mean?

- To be very valuable
- To be very heavy
- To be very rare
- To be very misleading

What does, "Lay down the law" mean?

- Put down something heavy
- Obey a police officer
- To give commands
- Commit a small crime

What does, "Two left feet" mean?

- Good at soccer
- Bad at dancing
- Good at karate

- Bad at hiking

What does, "Hit the hay" mean?

- Start a fight
- Gain weight
- Go to sleep
- Be afraid of water

What does, "Kick the bucket" mean?

- To make a mess
- To be angry
- To make a mistake
- To die

What does, "In the nick of time" mean?

- In the middle of something
- At the exact right time
- Very poor timing
- At the very last possible moment

What does, "Changed your mind" mean?

- Get some rest

- Switch your cloths
- Make a different decision
- To be emotional

What does, "Get off my back" mean?

- Stop criticizing me
- You are lazy
- I am more important
- I am very tired

What does, "Sleep on it" mean?

- I don't care
- Make a dream come true
- To think about it
- To be very tired

What does, "Freak out" mean?

- Be my friend
- Be too hopeful
- Let me down
- Overreact

What does, "He drove her nuts" mean?

- He took her to the countryside
- He was irritating her a lot
- She liked him very much
- She wanted to be together

What does, "Drop a line" mean?

- Send a letter
- Go fishing
- Go shopping
- Give a complement

What does, "She twiddled her thumbs" mean?

- She wanted to party
- She made a sweater
- The rest will be simple
- She was doing nothing

What does, "Crash course" mean?

- Very noisy place
- A bad classroom experience
- Rapid learning

- An argument

What does, "Tie the knot" mean?

- Tie your shoes
- Have stomach problems
- Get in trouble
- Get married

What does, "Change your mind" mean?

- Get a new brain
- Choose something else
- Have a new emotion
- Go to sleep

What does, "Big mouth" mean?

- Yawn often
- Eat a lot
- Very noisy
- Tell secrets

What does, "He has no spine" mean?

- He is flexible

- He is cowardly
- He is relaxed
- He is tired

What does, "He threw a fit" mean?

- She had a party
- She hit someone
- She was excited
- She complained

What does, "She stole the show" mean?

- She was the center of attention
- She ruined the performance
- She really like the movie
- She was the manager of them all

What does, "She made a killing" mean?

- She was allowed to enter
- She was very mean
- She forced them to do something
- She made a lot of money

What does, "He stretched his legs" mean?

- He is good at sports
- He walked around
- He is flexible
- He ran a long way

What does, "He held his tongue" mean?

- His mouth was full
- He didn't say anything
- He started feeling sick
- He became very angry

## References

- Able, B. "English idioms in the first language and second language lexicon: A dual representation approach." *Second Language Research*. 19 (2003): 329-358.
- Caillies, Stephanie, and Kirsten Butcher. "Processing of Idiomatic Expressions: Evidence for a New Hybrid View." *Metaphor and Symbol* 22 (2007): 79–108. URL: [http://www.leaonline.com/doi/abs/10.1207/s15327868ms2201\\_3](http://www.leaonline.com/doi/abs/10.1207/s15327868ms2201_3).
- Cutler, Anne., Swinney, David. "The access and processing of Idiomatic Expressions." *Journal of Verbal Learning and Verbal Behavior*. 18 (1979): 523-534.
- Dong, P. "The applicability of transformations to idioms." *Papers from the Seventh Regional Meeting of the Chicago Linguistic Society*. (1971): Chicago: Chicago Linguistic Society.
- Fraser, B. "Idioms within a transformational grammar." *Lang*. 6 (1970): 22-42.
- Gass, S. "Language transfer and universal grammatical relations." *Language Learning*. 29 (1979): 327-344.
- Gibbs, Raymond W. "Idiomaticity and Human Cognition." *Idioms: Structural and psychological perspectives*. (1995): 97–116.
- Gibbs, Raymond W. and Nadini P. Nayak., "Psycholinguistic Studies on the Syntactic Behavior of Idioms." *Cognitive Psychology*. 21 (1989):100–138.
- Gibbs, Raymond W; Nadini P. Nayak; J L Bolton, and M E Keppel. "Speakers' assumptions about the lexical flexibility of idioms." *Memory & cognition*. 17 (1989): 58–68.
- Gibbs. Nayak., Cutting. "How to kick the Bucket and Not Decompose: Analyzability and Idiom Processing." *Journal of Memory and Language*. 28.5 (1989): 576.
- Irujo, S. "Don't Put Your Leg in Your Mouth: Transfer in the Acquisition of Idioms in a Second Language." *TESOL Quarterly*. 20(2) (1986): 287-304.
- Jordens, P. "Rules, grammatical intuitions and strategies in foreign language learning." *Interlanguage Studies Bulletin*. 2(2) (1977): 5-77.
- Katz, J. "Compositionality, idiomaticity, and lexical substitution." In S. Anderson & P. Kiparsky. *A Festschrift for Morris Halle*. New York: Holt, Rinehart, & Winston (1973).
- Kellerman, E. "Towards a characterization of the strategy of transfer in second language learning." *Interlanguage Studies Bulletin*. 2(1) (1977): 58-145.
- Mason, W.; & Suri, S. "Conducting behavioral research on Amazon's Mechanical Turk." *Behavior research methods*, 44(1), (2012) 1-23.

- Nattinger, James R.. "Lexical Phrases and Language Teaching." *Oxford University Press Meta*. 38.3 (1992): XVI-218.
- Newmeyer, F. "The regularity of idiom behavior." *Lingua*. 34 (1974): 327-342.
- Nunberg, G. "The pragmatics of reference." *Indiana Linguistics Club* (1978).
- Oiler, J.W. Jr., & Ziahosseiny, S.M. "The contrastive analysis hypothesis and spelling errors." *Language Learning*, 20 (1970): 183-189.
- Paolacci, G.; Chandler, J.; & Ipeirotis, P. G. "Running experiments on Amazon Mechanical Turk." *Judgment and Decision Making*, 5 (2010) 411–419.
- R.L., Politzer. *Teaching French: An introduction to applied linguistics (2nd. ed.)*. Waltham, MA: Blaisdell Publishing, 1965.
- Sajavaara, K. "Contrastive linguistics past and present: A communicative approach." *Jyvaskyla Contrastive Studies*. 4 (1976): 9- 30. Jyvaoskyla, Finland: Jyvaiskylai University, Department of English.
- Strong-Krause, D., "Exploring the effectiveness of self-assessment strategies in ESL placement". In G. Ekbatani & H. Pierson (Eds.), *Learner-directed assessment in ESL* (2000) (pp. 49-73). Mahwah, NJ: Lawrence Erlbaum.
- Strassler, J., "Idioms in English". Tübingen: Gunter Narr Verlag (1982).
- Tabossi, Patrizia, Rachele Fanari, and Kinou Wolf., "Processing idiomatic expressions: effects of semantic compositionality." *Journal of experimental psychology. Learning, memory, and cognition*. 34 (2008): #313–27. URL <http://www.ncbi.nlm.nih.gov/pubmed/18315408>.
- Titone, D. A., & Connine, C. M., "Comprehension of idiomatic expressions: Effects of Predictability and literality." *Journal of Experimental Psychology: Learning, Memory, & Cognition*. 20 (1994a): 1126-1138.
- Titone, D. A., & Connine, C. M., "Descriptive norms for 171 idiomatic expressions: Familiarity, compositionality, predictability, and literality." *Metaphor & Symbolic Activity*. 9 (1994b): 247-70.
- Titone, D. A., & Connine, C. M., "On the compositional and noncompositional nature of idiomatic expressions." *Journal of Pragmatics*. 31 (1999): 1655-1674.
- Weinreich, U., "Problems in the analysis of idioms." In J. Puhvel (ed.), *Substance and structure of Language*. Los Angeles: University of California Press, (1969): 23-81.