

St. Cloud State University

theRepository at St. Cloud State

---

Culminating Projects in TESL

Department of English

---

5-2020

## A Study of ESL Learners' Depth of Knowledge of Figurative MWUs

Eleanor Jefferson

Follow this and additional works at: [https://repository.stcloudstate.edu/tesl\\_etds](https://repository.stcloudstate.edu/tesl_etds)

---

### Recommended Citation

Jefferson, Eleanor, "A Study of ESL Learners' Depth of Knowledge of Figurative MWUs" (2020).  
*Culminating Projects in TESL*. 31.  
[https://repository.stcloudstate.edu/tesl\\_etds/31](https://repository.stcloudstate.edu/tesl_etds/31)

This Thesis is brought to you for free and open access by the Department of English at theRepository at St. Cloud State. It has been accepted for inclusion in Culminating Projects in TESL by an authorized administrator of theRepository at St. Cloud State. For more information, please contact [tdsteman@stcloudstate.edu](mailto:tdsteman@stcloudstate.edu).

**A Study of ESL Learners' Depth of Knowledge of Figurative MWUs**

by

Eleanor Jefferson

A Thesis

Submitted to the Graduate Faculty of

St. Cloud State University

in Partial Fulfilment of the Requirements

for the Degree of

Master of Arts

in English: Teaching English as a Second Language

May, 2020

Thesis Committee:

Choonkyong Kim, Chairperson

Edward Sadrai

Maria Mikolchak

### **Abstract**

Multiword units (MWUs) make up a large portion of the language used by native and highly proficient speakers of English in both spoken and written discourse (Erman & Warren, 2000; Pawley & Snyder, 1983), yet second language vocabulary instruction often focuses on the meaning of single words in isolation. This is an insufficient way to teach vocabulary because many MWUs are figurative in nature and therefore “deceptively transparent” (Laufer, 1989) – the meaning of the words as a unit is different from the meaning of the individual words. Because of this, second language learners experience difficulties in learning and even recognizing MWUs when they encounter them in a text (Martinez & Murphy, 2011; Kim, 2016). The purpose of this study was to investigate the nature of ESL learners’ understanding of figurative MWUs and what semantic features of these MWUs might contribute to their comprehensibility – MWUs that described an action were compared with those that describe a situation. Using a test method designed by Laufer and Goldstein (2004), this study took the form of two tests – a Passive Recall test followed by a Passive Recognition test. The goal was to see at what depth of knowledge Arabic L1 ESL learners knew the correct meaning of the given MWU. It was found that in the Passive Recall test the participants scored lower, on average, when presented with MWUs that describe a situation than on those that describe an action. However, in the Passive Recognition test, there was no statistically significant difference between participants scores when the MWU was an action or a situation. From these findings, it is clear that ESL learners cannot simply be said to not know MWUs. Instead, these findings show that ESL learners have a complex understanding of figurative MWUs. Additionally, these findings confirm the need for ESL learners to receive instruction in MWUs.

### **Acknowledgements**

My sincerest gratitude to my participants who generously gave their time and energy voluntarily to be a part of my research. I would also like to thank those who assisted me in my pilot testing and recruitment processes. Additionally, this work was made possible with the wisdom, guidance, and time of Dr. Kim and the other members on my committee. I could not have done this without the help of many. It truly takes a village to raise a thesis.

## Table of Contents

	Page
List of Tables .....	5
List of Charts .....	6
Chapter 1: Introduction .....	7
Chapter 2: Literature Review .....	10
a. Defining & Classifying MWUs .....	10
b. Deceptive Transparency .....	14
c. Semantic & Syntax of Figurative MWUs .....	16
d. Importance of MWUs .....	17
e. Empirical Studies on MWUs in L2 .....	20
f. Research Questions .....	25
Chapter 3: Method .....	26
a. Participants .....	26
b. Materials .....	26
c. Procedure .....	30
Chapter 4: Results & Analysis .....	32
Chapter 5: Discussion .....	38
Chapter 6: Conclusion.....	42
a. Limitations .....	43
b. Pedagogical Implications .....	43
References .....	45
Appendices .....	48
a. Appendix A: IRB Approval .....	48
b. Appendix B: Pilot Test .....	49
c. Appendix C: Informed Consent .....	51
d. Appendix D: Background Survey .....	52
e. Appendix E: Test Instrument .....	53

**List of Tables**

	Page
Table 1: Four degrees of strength of knowledge .....	23
Table 2: MWUs Used in this Study .....	27
Table 3: Accuracy in Supplying the Meaning of MWUs - Passive Recall .....	32
Table 4: Accuracy in Recognizing the Meaning of MWUs - Passive Recognition .....	32

**List of Charts**

	Page
Chart 1: Situation MWUs - Distribution of Choices Among Options .....	33
Chart 2: Action MWUs - Distribution of Choices Among Options .....	34
Chart 3: Situ- and Act- Scores Compared Between Recall & Recognition .....	35

## Chapter 1: Introduction

Vocabulary acquisition is universally acknowledged as a vital part of language learning. It may even be the most important part (Folse, 2011). Students and teachers of language alike know that the more words a language learner knows, the more they can accomplish with a language and participate in that language community. Traditionally, vocabulary acquisition has been thought of in terms of connecting the individual words in a language to their meanings. The knowledge necessary to use vocabulary appropriately is much more complex than this form-meaning connection of single words, though. What is missing in these considerations of what constitutes “vocabulary” is the relationships between these individual words and their neighboring words in sentences and phrases. If English language learners are only taught the literal meanings of the words “hold,” “your,” and “horses,” how are they to know that the phrase “hold your horses” is frequently used, as a single unit, to request for the someone to be patient? Or, in a subtler example, if a language learner is only taught the literal meanings of the words “on,” “other,” and “hand,” how can they be expected to understand, without explanation, that “on the other hand” can be an expression of contrast?

“Hold your horses” and “on the other hand” are just two examples of multiword units (MWUs) that exist in the English language. MWUs are groupings of words, common phrases, in which the meaning of the whole unit can be separate from the sum of the meaning of the individual words involved. Such phrases are often described as being “non-compositional” MWUs (Martinez & Murphy, 2011). MWUs are important for language learners to know because both native and highly proficient speakers of English communicate most often through such idiomatic phrases and habitual expressions. Very few of the sentences that proficient



language users say or write are unique or new. Because of this, comprehending and appropriately using MWUs is vital to reaching a high level of proficiency as an English language speaker.

However, learning and teaching MWUs is not as straightforward as learning individual words. These units of words present specific difficulties to learners and even to the researchers who study them. First and foremost, they present difficulties because MWUs do not exist in clear cut, easily defined categories (Liu, 2008). Many different pieces of language can be included under this umbrella term, varying in phrase structure, phrase length, and linguistic or semantic function.

Some MWUs are also easier to understand from their component words than others. Several studies, including work by Martinez and Murphy (2011) and Kim (2016) have established how the “deceptive transparency,” a term introduced by Laufer (1989), of non-compositional MWUs can negatively impact reading comprehension for ESL learners. MWUs can give learners a false sense of how well they have understood a text. If language learners recognize the individual words in an unfamiliar MWU, they can incorrectly believe that they understand that phrase, not realizing that the meaning of these words as a unit is different than the meaning of the individual words (such as in the MWU “on the fence” – an expression of indecision and not a description of a physical location). In their studies, Martinez and Murphy (2011) and Kim (2016) were both measuring ESL learner’s passive knowledge of MWUs – this terminology, borrowed here from Laufer and Goldstein (2004), refers to the learner’s ability to either supply or recognize the meaning (as opposed to the form) of a word when provided with the L2 form of that word in a prompt.

This present study aimed to build on the works of Martinez and Murphy (2011) and Kim (2016), using a method modeled after Laufer and Goldstein’s (2004) vocabulary size and depth

test. While Laufer & Goldstein's (2004) full test also included four degrees of strength of knowledge – Active Recall, Active Recognition, Passive Recall, and Passive Recognition – those two levels of active knowledge (which require supplying or recognizing the form of the L2 word as opposed to the meaning) are not included here. This is because the goal of this study is to add to the existing body of knowledge on ESL learner's passive knowledge of MWUs – how do ESL learners interpret MWUs when they encounter them in written texts? Unlike Martinez and Murphy (2011) and Kim's (2016) studies, Laufer & Goldstein's (2004) test method allows for a more detailed picture of the subtleties of the learner's passive knowledge of MWUs. In addition, while Laufer & Goldstein (2004) looked at single words, this study applies their test method to MWUs.

The MWUs investigated in this study are those that have a possible literal interpretation as well as a possible figurative interpretation. The goal was to further investigate the difficulties that this aspect of language presents to language learners and what factors might contribute to the comprehension and learning of MWUs for ESL learners. An important question for this study is whether there are factors – specific features of certain figurative MWUs, for example – that might make them easier or more difficult for language learners to understand. Through the data collected for this study, it is clear that ESL learners cannot simply be said to not know the MWUs in this study, though the average score for these participants was quite low on the Passive Recall part of the test, when they were asked to supply the meaning of the MWU presented. Instead, these findings show that ESL learners have a complex understanding of figurative MWUs, because they scored much higher on the Passive Recognition part of the test, when asked to choose the meaning of the MWU from a provided list.

## Chapter 2: Literature Review

### Defining & Classifying MWUs

Defining and describing the precise characteristics of MWUs is the greatest difficulty that these groups of words pose for both linguists and language teachers. How MWUs are defined influences both how they are studied in research as well as whether, and how, teachers incorporate them into their curriculum. Within the literature, conversations about MWUs inevitably start with the author choosing a label for these word groups and explaining the characteristics of these groups that are relevant to that specific discussion or study. MWUs – a term adopted from the writings of Nation (2013) – have also been called “lexical phrases” (Nattinger & Decarrico, 1992), “fixed expressions” (Moon, 1998), “formulaic language” (Wray 2002), “multiword expressions” (Martinez & Murphy, 2011), and “formulaic sequences” (Alali & Schmitt, 2012), among other names in various studies. Regardless of the different labels given to MWUs, these researchers are all talking about a similar idea: a piece of language made up of multiple words – such as an expression or idiomatic phrase – that is commonly used and easily understood by proficient speakers of that language. The term “idiom” has also been used throughout studies on this topic, but most often it is used to describe a particular type of MWU, as in the work of Fernando (1996), and Moon (1998). “The straw that breaks the camel’s back” would be an example of such an idiom. It is usually obvious that the meaning of this type of MWU is not to be understood literally from its component parts. However, even in his discussion of idioms, Liu (2008) argues that there is no definition or label that can possibly encapsulate all of the variant forms of the linguistic phenomenon of MWUs, particularly because researchers are asking so many different questions about MWUs regarding their behavior and importance in language (p. 3).

Despite variations in the specific details of how researchers classify MWUs, most seem to follow a version of either a “form-based approach,” “storage-based approach,” or “meaning-based approach” as described by Nation (2013). The basis for labeling a group of words as an MWU in a form-based approach is how frequently these words occur together. This can be investigated through searching various corpora (p. 485). One danger, here, however, is that even some things that would not be considered MWUs, i.e. groups of words that simply are frequently used together but otherwise have no unique linguistic function, would be called MWUs as well. This means that even incomplete phrases such as “is one of the” would have to be considered an MWU simply because these words occur together frequently in spoken and written English (Nation, 2013, p. 486). Fernando (1996) provides some clarification to this point in the terminology he uses to characterize idioms. One of his criteria is “institutionalization” (p. 3). Fernando says that idioms (MWUs) must be institutionalized phrases in order to be considered true idioms; they cannot merely be any common grouping of words. For example, “to fly off the handle” is an MWU, while “to get angry” is not, even though both of these are groups of words that commonly occur together. “To get angry” does not require any institutionalization of its meaning because its meaning is transparent from its parts; “to fly off the handle” does, though, because it is not transparent.

Storage-based approaches, on the other hand, seek to define MWUs on the basis of how they are stored in the speaker’s mental lexicon, whether as a whole unit or in their individual parts (Nation, 2013). This issue decides whether the MWU is treated as a single lexical item or as a collection of individual words by the language user and in research or language instruction. Ellis (1996) describes MWUs as behaving simply like “big words,” saying that working memory performs the same when learning MWUs as it does for single words (p. 111). The *lexical*

*representation model* further expands on this idea of MWU storage. This model, put forward by Swinney and Cutler (1979), suggests that when MWUs are encountered, they are processed both as single lexical items and as a string of individual words by language users simultaneously. This is until, using context clues, the language user can choose which meaning is most appropriate (p. 525-526). Some MWUs, those with a more obvious figurative meaning than others (such as “beat around the bush”) can more readily be seen and understood to be simply large words (Wray, 2000, p. 466). However, that is not the case for all MWUs.

A study conducted by Elmquist (2014), showed proof of MWUs being stored as whole chunks even by language learners. In her study, Elmquist had both native and nonnative speakers of English judge the acceptability of variations of MWUs. In the test instrument, the MWUs were modified in three ways: by adding or changing a modifier, quantifier, and key word. For example, “costs an arm and a leg” was modified into “costs a *huge* arm and a leg,” “costs an arm and *two* legs,” and “costs an arm and a *foot*” (p. 22). Participants then had to rate the acceptability of such modifications on a scale from clearly unacceptable to clearly acceptable. In her results, Elmquist found that non-native English speakers are able to make judgments about what type of variance is acceptable in MWUs that are similar to the judgements of native speakers. This was true for all three types of modification – changing or adding a modifier, a quantifier, or a key word. As stated in her conclusions, she believes that if idioms were not stored as whole chunks in the language learners’ mental lexicon the acceptability judgments of the language learners would be significantly different from the native speakers (p. 30).

The third approach that Nation says some linguists adhere to when defining MWUs is the meaning-based approach, which relies on the compositionality of the word group as the main deciding factor in whether it is an MWU or not – in other words, whether the meaning of the

word group can be understood from the meaning of the individual words or not. Word groupings with less decipherable meanings can be considered to be MWUs. Fernando (1996) also uses this criterion in his discussion on this topic. He describes some MWUs as “semantically anomalous” (p. 22) and, for this reason, “semantically opaque” (p. 3). They are semantically anomalous and opaque because they communicate meaning that is in opposition to the traditionally understood senses of the individual words within them. An example of such a semantically opaque MWU would be “on thin ice,” which, when taken literally, would mean that something is situated on a piece of thin ice and in danger of falling into cold water; however, in its institutionalized, idiomatic sense, this expression means to be “about to make a mistake or offend someone.”

Such semantically opaque MWUs are often figurative in nature and, for native and highly proficient speakers of the language, conjure up a mental picture that contributes to the understanding of the idiomatic meaning of the MWU. In the case of the MWU above, “on thin ice,” the mental image of someone or something being on a patch of thin ice over freezing water does carry a sense of risk or danger. The communicative power of such MWUs for native and highly proficient speakers of a language lies in the conflict between their “essential untruthfulness” and how, in a way, they are “representative of truth” (Moon, 1998, p. 193-194). Lakoff and Johnson (1980b) explain this in their work on conceptual metaphors. They claim that the human “conceptual system is metaphorically structured,” that many of the concepts people understand are at least “partially understood in terms of other concepts” (p. 475-476). Humans have the ability to conceptualize an object or experience as a different kind of object or experience because, as Lakoff and Johnson (1980a) explain, there is an entailment relationship between the two objects or experiences (p. 197). They offer the example of time and money in the idiom *time is money*. The concepts of both time and money, in western cultures, entail the

ideas of being valuable yet limited (p. 197-198). This is the reason that this idiom has communicative power, because of the entailment relationship between the two words. This is the reason that figurative or metaphorical MWUs have such communicative power. Being “on thin ice” and being on the brink of offending someone both entail the idea of being in danger or in a risky situation. Lakoff and Johnson (1980a) go on to say that the basis for understanding metaphors is experience. Metaphors cannot be understood or explained without some connection to the concrete and tangible experience that is being used as the vehicle to communicate the abstract idea (p. 205). It is the concrete experience that makes the metaphor powerful.

The creativity of metaphorical language, the “untruthfulness” of using words in ways that contradict their primary meaning, is precisely what is most difficult about figurative MWUs for language learners. When unaware of the presence of a figurative MWU, learners can only try to understand them as the combination of their individual parts (Martinez & Murphy, 2011). They do not know that they should do anything else. This is known as “deceptive transparency” (Laufer, 1989).

### **Deceptive Transparency**

Laufer (1989) describes deceptive transparency as a difficulty for language learners regarding single words as well as MWUs. She defines a deceptively transparent word or MWU as one that looks like it should mean one thing but in fact means something completely different (p.11). In the case of an MWU, this means that the individual words in this phrase are known to the learner, but the new meaning they take on as a unit is unfamiliar. Though not all MWUs are deceptively transparent and some can be understood from their parts – Nation (2013) refers to these as “literals,” such as “at the moment,” “you know,” and “I think (that)” (p. 490) – a large

number of MWUs are deceptive – “on the fence” or “miss the boat” (Laufer, 1989, p. 12). These are called figurative MWUs.

When learners come across something that is unfamiliar, they have three options, according to Laufer (1989): “ignore it,” “ask someone” or “try to guess it from context” (p. 16). But in order for the learner to use any of these strategies, they must first recognize the phrase or word as unfamiliar. With deceptively transparent words, however, learners are often unaware that these are unfamiliar. Because of the issue of deceptive transparency, teachers cannot rely on incidental learning as the primary way that their students will learn MWUs.

The main difficulties that learners experience with MWUs as a result of deceptive transparency can be divided into two areas: productive and receptive. Because language learners lack ‘native intuition’ and because of insufficient instruction in MWUs in the classroom, learners can have trouble discerning when a phrase they are using is, in fact, a holistic, commonly used chunk of words or whether it is simply a sentence created based on the grammatical rules of the language. In addition, Liu (2008) states that the acquisition of MWUs follows the traditional order of “comprehension before production” (p. 101). However, the issue of deceptive transparency would seem to prohibit acquisition from happening for language learners – because they do not know that they are not actually comprehending the MWUS – in cases where incidental learning is relied upon. Therefore, it can be assumed that production would be prohibited as well.

Additionally, there are often no clues that the input the learner is receiving contains MWUs. This is particularly obvious in written texts, where strings of words are merely separated and connected by white space and punctuation, regardless of whether it contains an MWU or not. MWUs can also take a variety of syntactic forms – verb phrase, prepositional phrase, noun



phrase, etc. – so there are no syntactic clues, either, that something is an MWU. This is also true for spoken English. Highly proficient or native speakers of English use prosody – “a coherent intonation contour” – in their speech to mark an MWU as a single unit within a discourse (Lin, 2012, p. 344). However, the factor of intonation is often overlooked in classroom instruction as well. This lack of clues contributes to the issue of deceptive transparency (Laufer, 1989).

### **Semantics & Syntax of Figurative MWUs**

The various forms that figurative MWUs can take – varying parts of speech, varying phrase lengths, varying semantic functions – while being a source of difficulty for language learners and researchers, is also what makes figurative MWUs so useful for communication. They can accomplish a wide variety of linguistic purposes within discourse.

In order to focus discussions and studies around MWUs, researchers have suggested categorizing them around specific features, whether semantic or syntactic. Liu (2008) uses Makkai’s (1972) “six subcategories” of what he calls “lexemic idioms” – those idioms which are longer than one word and whose meaning is separate from the meaning of the individual words – in his work on MWUs (p. 16). These categories are:

- 1) phrasal verbs – such as “*come across*”
- 2) tournures – such as “*take the bull by the horns,*”
- 3) irreversible binomials – such as “*friend or foe,*”
- 4) phrasal compounds – such as “*high-handed,*”
- 5) incorporating verbs – such as “*eavesdrop,*”
- 6) pseudo-idioms” – such as “*chit-chat*” (p. 17).

If syntax is emphasized in the research questions, MWUs may also be separated based on whether they are merely collocations, typically two or three words in length, or whole phrases, perhaps three words or longer, that can be called MWUs – such as verb phrases (“break the bank), noun phrases (“a fish out of water”), prepositional phrases (“in a tight spot), etc.

Though all of the above examples of MWUs are non-compositional, each of the six category varies greatly in the type of message they are communicating and whether they are indeed figurative and bring an image to mind or not. If investigating the figurative nature of non-compositional MWUs, it might be more useful to categorize them based on their semantic function. Liu (2008) also suggests several “message content types” that MWUs can fall into:

1. actions: *twist somebody's arm...*
2. events: *the bottom drops out...*
3. situations: *in hot water...*
4. people and things: *a back-seat driver...*
5. attributes and evaluations: *cut and dried...*
6. emotions: *green with envy...* (pg. 20).

Such divisions would be useful if semantics is of interest in the research questions being investigated. When MWUs are sorted in this way, the figurative nature of the MWUs is more obvious than if they were to be sorted based on syntax. Each example for the above categories brings to mind a clear mental image of the message it is communicating.

However, syntax and semantics are not mutually exclusive within language, creating further complexities in studying these pieces of language. Even in the examples of the different message content types, the syntax of each MWU is different. Action MWUs, commonly describing something the speaker is doing or having done to them, take the form of verb phrases; MWUs that describes events or people and things are noun phrases; situation MWUs, primarily describing social situations, take the form of prepositional phrases; and the attributes and emotions MWUs take the form of adjective phrases.

### **Importance of MWUs**

The issue of deceptive transparency is just one reason why MWUs are important for language learners. The second reason that MWUs must not be overlooked in language instruction is because of the frequency with which MWUs are used both in written and spoken English

(Erman & Warren, 2000; Pawley & Snyder, 1983). As Nation (2013), asserts, MWUs are important for ESL learners to be made aware of because most words in English have expected and habitual patterns in which they typically occur and that knowing this pattern can help learners use the words correctly (p. 481). This is evidenced in the frequency with which MWUs are used in English by both highly proficient and native speakers in spoken and written discourse. From the results of an analysis of several different written texts conducted by Erman and Warren (2000), they estimate that MWUs make up over half (about 55%) of any written text in English (p. 50). Pawley and Snyder (1983) also contend that a majority of an English speakers' lexicon is made up of "complex lexical items," and this includes thousands of MWUs (p. 215).

One reason why MWUs are so frequently used, according to Pawley and Snyder (1983) is that MWUs are "*social institutions*," meaning that MWUs are "conventional label[s] for conventional conceptual categor[ies]" (p. 209). MWUs are powerful tools for communication because they carry the "authority" of common, frequent and conventionalized use by the speakers of that language (p. 209). They are familiar and easily understood. MWUs can have a special ability to communicate things more vividly and powerfully than unique sentences, too (Liu, 2008, p. 101), as in the case of figurative MWUs.

Another reason that native and highly proficient speakers of a language know so many of these MWUs, the reason they are used so often, is that humans are "creatures of habit." Nattinger and DeCarrico (1992), make this statement, claiming that people like to use "routinized formulas" and "ritualized language" more than they like to create original sentences (p. 1). Pawley and Snyder (1983) say that in conversation "listeners are tolerant, up to a point," of unexpected or novel language use, but ultimately their patience for it is limited (p. 208). They

also argue that this “reliance on ready-made expressions” does not prohibit creativity in spoken language; in fact, having thousands of phrases or sentences memorized and ready to use allows the person speaking to “channel energies to other activities” in the discourse, such as affecting their tone, rhythm, or working creatively within the “ready-made” language chunks (p. 208).

Wray (2002) contributes to this discussion by bringing in Chomsky’s assertion of human linguistic creativity from his book *Aspects of the Theory of Syntax*, published in 1965 (p. 11). Wray believes Chomsky’s theory – that humans are capable of creating and understanding an infinite number of original sentences – is true, but that this ability has been overemphasized as it relates to real language use (p. 11). Though humans are capable of creating unique, never before used sentences every time they talk or write, they actually prefer to use phrases that have been used before, phrases that are used regularly by other speakers of the same language. If native speakers of a language find MWUs to be so communicatively convenient, the same would be true for language learners (Jones & Haywood, 2004, p. 269).

Pawley and Snyder (1983) take this even further in their claim that if native speakers of a language were to exercise their linguistic creativity to its fullest extent every time they spoke they would be seen as non-nativelike – “not...exhibiting nativelike control of the language” (p. 193). Even though the grammatically correct possibilities are endless, only some of those possibilities are considered “nativelike” by speakers of that language. Therefore, using completely novel sentences every time a person speaks or writes would appear “odd” or “foreign” (Pawley & Snyder 1983, p. 193). The decision that is made between all possible grammatically correct choices requires highly proficient pragmatic competence (Nattinger & DaCarrico, 1992, p. 11). Nativelike competence, then, is choosing the option that is grammatically correct as well as pragmatically appropriate in a given situation. This is why

learners need to be made aware of MWUs, because nativelike selection goes beyond the constraints of English grammar to add other restrictions to what is more commonly used and what is not. Nativelike selection adds the expectation that native speakers of English have of what is being said or what should be said together, including what expressions should be used for certain sentiments. This is not to say that every English language learner must reach nativelike proficiency, but MWUs must be learned if that level of proficiency is the goal.

### **Empirical Studies on MWUs in L2**

Two studies, one by Martinez and Murphy (2011) and the other by Kim (2016), were conducted to test learner awareness and understanding of MWUs in written texts, and both showed evidence of deceptive transparency significantly affecting reading comprehension. These studies had participants engage in a reading task and then a task designed to gauge reading comprehension, as well as a self-report of how the participants felt they did in the task. Martinez and Murphy's (2011) study tested 101 adult English language learners from Brazil with similar levels of proficiency in English (p. 275). The test took place in two parts. In each part, learners were presented with a text to read. Both texts contained identical lexical components, and all of the words that made up the texts came from the first 2,000 most commonly used words in the British National Corpus (p. 276). One of the two texts, however, did not contain any MWUs, while the other was made up primarily of MWUs – the same words from the first text but recombined into MWUs. The participants' comprehension was gauged through a true or false style test and the participants' own rating of how much they believed they understood was measured through a personal rating on a scale from 5-100% (p. 276). In their results, Martinez and Murphy found that participants' comprehension of the texts containing MWUs was significantly lower than their comprehension of the texts which did not contain MWUs (p. 281).

They also found in the comprehension self-report completed by the participants that the participants overestimated of their own comprehension of the texts containing MWUs (p. 281).

Kim's (2016) study was similar in some ways to Martinez and Murphy's (2011) test in that participants engaged in a reading test and self-reported on their own understanding. However, for her study, Kim (2016) made the MWUs in the text explicit to the participants. The participants of this study were international students, 21 female and 31 male, studying at a university in the US (p. 96). Participants were presented with both unfamiliar MWUs and unfamiliar words in short, paragraph-long stories (p. 97). The unknown items were underlined, therefore making them explicit for the participants. The first task participants were asked to do was to "Read & Identify" if the underlined items in texts 1) were unfamiliar, 2) sounded strange, or 3) were familiar; the second task was, for each underlined item in the texts, to "1) write what it meant in the story; 2) write a synonym that made sense in the story; or 3) leave it blank if they did not know what it meant" (p. 99).

The results of this study showed that even for the MWUs that learners marked as familiar in the first task, phrases the participants claimed to know, the definitions they wrote in the second task were low in acceptability (p. 104). Here, Kim (2016) found proof for Laufer's (1989) idea of deceptive transparency in that, despite the fact that these MWUs were explicitly pointed out to the participants, these falsely familiar phrases prevented learners from even noticing them as unfamiliar (p. 104). Kim also notes that for her study and for Martinez and Murphy's (2011) study, learners do not appear to perceive MWUs as whole units of language, but instead try to understand the meaning from the parts (p. 104).

In the self-reports from both studies, learners believed that they had understood the MWUs that were present, whether the MWU was explicit or not. However, the participants in

both studies did not truly understand as much as they thought they did. Both studies provide evidence for the deceptive transparency of MWUs and support the argument that MWUs need to be included in language instruction – because learners are unaware of this gap in their knowledge and the way it affects their comprehension of the language used around them.

Do these results mean that language learners have no understanding of MWUs at all? Or is their knowledge of the meaning of MWUs more complex than that? Laufer and Goldstein (2004) designed a Vocabulary Knowledge Size and Strength Test so that they could test the subtleties of vocabulary knowledge in language learners. Their participants were high school students studying English as their second language. Laufer and Goldstein believed that knowledge of vocabulary was more than just a simplistic form-meaning connection. One of the main assumptions that guided their study was that “the form-meaning link in the mental lexicon can have four degrees of strength” (p. 409). These four degrees of strength, in order from strongest to weakest, are: active recall, passive recall, active recognition, and passive recognition (p. 406-7). In their terminology, having “active” knowledge of the meaning of a word involves being able to supply or recognize the *form* of an L2 word when the meaning was provided in some kind of prompt. “Passive” knowledge involves supplying or recognizing the *meaning* of an L2 word, given the form in some kind of prompt. “Recall” refers to being able to supply or provide the L2 word or meaning, given a prompt of some kind, whereas “recognition” means merely recognizing and picking out an L2 word or meaning from a given list. In Laufer and Goldstein’s (2004) test, they first presented participants with test items at the active recall level, asking them to supply the L2 form of a word when giving the meaning. Knowledge of the same word was then tested at the lower levels if the participant did not answer correctly. This test was computerized, allowing participants to finish early if they proved they knew all the words at the

active recall level. However, for any words that were answered incorrectly at the active recall level, more questions were provided at the other levels – first passive recall, then active recognition, then passive recognition – to see at what level the words were known. They found that their participants scored lowest at the active recall level and highest at the passive recognition level (p. 421), supporting their idea that the four degrees of strength of knowledge are implicationally scaled (p. 417). The implicational scale of the four degrees of strength of knowledge can be seen in Table 1:

**Table 1**

*Four degrees of strength of knowledge*

	<b>Recall</b>	<b>Recognition</b>
<b>Active</b> (retrieval of form)	(1) strongest	(2) or (3)
<b>Passive</b> (retrieval of meaning)	(2) or (3)	(4) weakest

(Laufer & Goldstein, 2004, p. 408)

The implicational scale presented here shows that if the word is known at the active recall level, it can be safely assumed that it is also known at the active recognition, passive recall, and passive recognition levels. In contrast, if a word not known at the passive recall level, it will be known at the passive recognition level, but it will not be known at the active recall level.

What Laufer and Goldstein’s (2004) study shows is that “knowledge of form-meaning link is not an all or nothing phenomenon but depends on what the learner is required to do with the knowledge” (p. 426). They said that previous vocabulary tests had not be able to “capture this distinction” (p. 426). In the same way, previous tests have not captured this distinction specifically for MWUs. While Laufer and Goldstein’s (2004) study was based on single words, it is hypothesized here that their four degrees of strength of knowledge can also be applied to MWUs, especially if the storage-based approach to classifying and understanding MWUs, as



discussed above, is taken. Therefore, Laufer and Goldstein's (2004) test method is the basis for this present study, though only the passive recall and passive recognition levels of their test. This is because this study is primarily interested in ESL learners' passive knowledge of MWUs – when provided with the L2 form of the MWU, how do ESL learners interpret these MWUs?

### **Summary**

The evidence would suggest that instruction in deceptively transparent MWUs and their figurative meanings is as vital to vocabulary acquisition as instruction in single words and their meanings. Because MWUs are used so frequently in everyday language for highly proficient speakers, are so efficient for communication, and yet they provide no clues to language learners that they are more than their component parts, researchers and teachers must seek to more fully understand how language learners handle these linguistic phenomena when presented with them.

In the same way that, for this study, the term MWU has been chosen, out of the multitude of possible names, to describe this linguistic phenomenon, certain MWUs have also been chosen as the focus. Here, because the non-compositionality and semantic opacity of MWUs is of primary interest, the role of the semantic function of MWUs in learner comprehension is the focus. Therefore, MWUs have been chosen on the basis of their semantic function, and, using Liu's (2008) list of "message content types," MWUs describing actions (and taking the form of verb phrases) and MWUs describing situations (taking the form of prepositional phrases) have been chosen to compare and contrast in this study.

In an effort to build on the findings of Martinez and Murphy (2011) and Kim (2016), and add to research on ESL learners passive knowledge of MWUs, Laufer and Goldstein's (2004) test method was borrowed in order to study how language learners handle deceptively transparent MWUs when presented with them in a limited context. It was expected that the

findings of this research will agree with previous findings that non-compositional MWUs present specific difficulties to learners and that they need to be pointed out and explained in order for students to understand and really learn them.

The research questions for this study are as follows:

1. When presented with an unfamiliar, non-compositional MWU in the context of a sentence, how effective are second language learners at guessing the meaning of these MWUs?
2. Is the semantic nature of the MWU (whether it describes an action vs. a situation) a factor in the accuracy of their guesses?

## **Chapter 3: Method**

### **Participants**

For this study, twenty-three native Arabic speakers were recruited from the student body of a four-year public university in central Minnesota. These students were all from the same language background so that no additional variables related to their L1 were introduced into the study. These students were also, at minimum, in their second year of study at the university – past the level of any ESL specific classes – in order that their level of proficiency met certain minimum standards. There were twenty male participants and three female. Participants were recruited through contacts within the Arab community at the university.

### **Materials**

#### **Pilot Test**

In order to choose the MWUs for this study – those that were opaque and unfamiliar to the participants – a pilot test was given to students of the same demographic as the target participants. This pilot test can be found in Appendix A. The pilot test contained a list of 39 MWUs, seventeen of which describe an action, such as “drop the ball” or “beat around the bush”; another seventeen describe a situation, such as “on the fence” or “in hot water”; and 6 were intended to be easy filler items – MWUs that were expected to be known to the students already. The purpose of having filler items in the finalized test was to help lessen any test anxiety or fatigue the participants may experience by giving them some easy questions to answer.

Of the MWUS included in the pilot test only those which were found to be unknown to the participants were included as test items in the final test instrument. Those items which were found to be easy for the participants in the pilot test were included as the filler items. The finalized test instrument included 30 test items on each of the two tasks – 30 different MWUs within simple sentence frames. Ten of these items included target situation MWUs (Situ-

MWUs), another ten were target items including action MWUs (Act-MWUs), and the final ten items were filler items including MWUs.

Before administration of the finalized test instrument, these test items were also pilot tested with two native speakers of English to verify that the sentence frames were understood in the same way by multiple native English speakers and not just the researcher.

### **MWUs & Other Vocabulary Chosen for this Study**

The MWUs chosen for this study, as listed in the pilot test, were short – between three and five words in length – in order that the length of the expression might not be a salient variable in the results. The words that make up each of these MWUs are from the first 2000 most frequently used words in the English language, as analyzed using the Web VocabProfile Classic (n.d.) on *Compleat Lexical Tutor*. It was important that the individual words that make up the MWU be familiar to the students, but that the figurative meaning of the words together, as one unit, be unfamiliar. The final MWUs chosen for the study can be found in Table 2:

**Table 2***MWUs Used in This Study*

<b>Situation MWUs</b>	<b>Action MWUs</b>	<b>Filler MWUs</b>
On the spot	Brush up on	Show your true colors
On the fence	Hit the books	A piece of cake
In hot water	Kick a habit	Follow your heart
Out on a limb	See the light	Kill two birds with one stone
On thin ice	Test the waters	On the same page
Out of the woods	Pull some strings	Spill the beans
On shaky ground	Beat around the bush	Out of my hands
On the wrong track	Break the bank	Pull some strings
In over your head	Ask for the moon	Tell it like it is
On pins and needles	Drop the ball	Have the floor

In addition, all of the vocabulary used throughout the test instrument, both in the sentence frames for the test items and in the options for the passive recognition multiple choice questions was chosen based on the first 2,000 most commonly used words in the English language. This was also determined using the VocabProfile Classic on *Compleat Lexical Tutor*. The purpose of this was to mitigate the effect of unfamiliar vocabulary on the participants' performance. The sentence frames for the test items were also designed to contain no unnecessary context clues to aid the participants in making their decisions. This was because truly testing whether participants know the figurative and idiomatic meaning of the MWUs can only be done when the participants are not offered any assistance. If the participants knew the MWU, it was assumed that they would treat the MWU as a unit and write or select the figurative meaning during the test and not the literal interpretation of these words.

### **Preliminary Materials**

The informed consent form for this study, explaining the purpose of the study and how the data collected would be used, can be found in Appendix B. A Background Survey, found in Appendix C, was also created to collect some demographic information from participants. The background survey simply asked the participants to provide some information about their language learning background, such as length of stay in the US, and number of years studying English. In the discussion section below, the information collected in the background survey is compared to student results to see if these factors could have influenced the accuracy of answers given.

### **The Test**

Modeled after a portion of the vocabulary size and depth test designed by Laufer and Goldstein (2004), the test for this study, found in Appendix D, had two parts – a short answer style task followed by a multiple-choice style task. Each part of the test had 30 test items. The first part of the test, the Passive Recall test, was simply a list of sentences containing MWUs that have been marked, such as:

1. Sarah said she would need to **brush up on** her Spanish.

Participants were instructed to write the meaning of *idiom* that was underlined and bold in the space below each sentence. Participants could write their answer either in English or in Arabic, whichever was more comfortable for them. This first part of the test was meant to gauge how well the participants know these MWUs. If the participant can recall or supply the meaning of the MWU, they could be said to have a true understanding and deep knowledge of this MWU.

The second part of the test, the Passive Recognition test, used the exact same sentence with bolded expressions, arranged in a different order, but this time the participants were given a

list of three possible meanings for the MWU to choose from, as well as the option of “I don’t know.” For example, the following are the possible choices for the meaning of the underlined expression in sentence 1, as seen above:

- a. Forget about
- b. Sweep dirt off of
- c. Review
- d. I don’t know

Each of these options can naturally fit the syntactic frame of the sentence given. Of the three options preceding “I don’t know,” one is the literal interpretation of the MWU – in the example above, this would be “b. Sweep dirt off of.” Another of the options is the figurative or idiomatic interpretation – “c. Review.” The third is a distractor, something that can fit in the sentence frame but is neither the correct literal nor the figurative interpretation of the MWU. In the example above, this would be option “a. Forget about.” This second part of the test was meant to show which MWUs the participants knew at a lower level on the hierarchy of knowledge than the first test. If a participant had answered incorrectly for a certain MWU on the Passive Recall test, but could identify the correct, figurative meaning of the MWU from the list provided in the Passive Recognition test, it would mean that they had some level of knowledge of this MWU, but not very deep knowledge.

### **Procedure**

Before beginning the test, participants were given the informed consent form and then the background survey. After filling in these preliminary materials, the participants were given the first test, the Passive Recall test, to complete. Only after they had finished this part, and the Passive Recall test was collected, were they given the second test, the Passive Recognition test. There was no time limit for participants to complete the test; they were free to work at their own pace, but they had to complete both parts in one sitting. Two versions of the test, with test items

in different orders, were distributed during data collection so that the order effect could be analyzed at this part of the research process.



## Chapter 4: Results & Analysis

Once the data was collected, any answers provided in Arabic were translated with the help of a native speaker of Arabic. Rater reliability for answers provided on the Passive Recall test was verified by having another native English speaker judge a set of participant answers. This was compared to the researcher's rating of the same answers; they were found to be above 80% similarity. The remainder of the data set was then rated by the researcher alone.

To determine if there was any order effect within the test items, a paired t-test was conducted. Based on the subjects' answers for Situ-MWUs on the Passive Recall test across both versions of the test, a paired t-test showed no statistically significant difference ( $t [22]=0.751; p < .461$ ) in their accuracy between the first half of the test items ( $M=17.39; SD=18.64$ ) and second half of the test items ( $M = 21.30; SD = 20.74$ ). Based on the subjects' answers for Act-MWUs in the Passive Recall test across both versions of the test, a paired t-test showed no statistically significant difference ( $t [22]=0.462; p < .648$ ) in their accuracy between the first half of the test items ( $M =36.52; SD =29.33$ ) and second half of the test items ( $M =38.69; SD =25.64$ ).

The same paired t-test was run with the subjects' answers for the Passive Recognition test for Situ-MWUs across both versions of the test, and also showed no statistically significant difference ( $t [22]=0.364; p < .720$ ) in their accuracy between the first half of the test items ( $M = 65.00; SD = 31.44$ ) and second half of the test items ( $M =68.00; SD =27.59$ ). This was also true of the paired t-test run on subjects' answers in the Passive Recognition test for Act-MWUs across both versions of the test – no statistically significant difference ( $t [22]=0.816; p < .423$ ) in their accuracy between the first half of the test items ( $M =70.43; SD =24.77$ ) and the second half of the test items ( $M =74.78; SD =18.31$ ). These order effect analyses indicate that the

participants' performance, from start to finish, was consistent throughout both sets of tests.

Therefore, all subsequent analysis of the data was based on the entire set of data.

### Passive Recall

Table 3 displays the average score of the 23 participants ( $N$ ) for the Passive Recall test.

**Table 3**

*Accuracy in Supplying the Meaning of MWUs - Passive Recall*

	Average Score (Max = 100%)	$N$	Std. Deviation
Situ-MWU	19%	23	15.2
Act-MWU	38%	23	25.1

Participants provided acceptable definitions of the Situ-MWUs 19% of the time on average and 38% of the time on average for Act-MWUs. Based on the subjects' definitions in the recall test, a paired t-test showed a statistically significant difference ( $t [22]=5.474$ ;  $p<.000$ ) in the accuracy of their answers between the Situ-MWUs ( $M=19.35$ ;  $SD=15.25$ ) and the Act-MWUs ( $M=37.61$ ;  $SD=25.13$ ).

### Passive Recognition

Table 4 displays the average score of the 23 participants ( $N$ ) for the Passive Recognition test.

**Table 4**

*Accuracy in Recognizing the Meaning of MWUs - Passive Recognition*

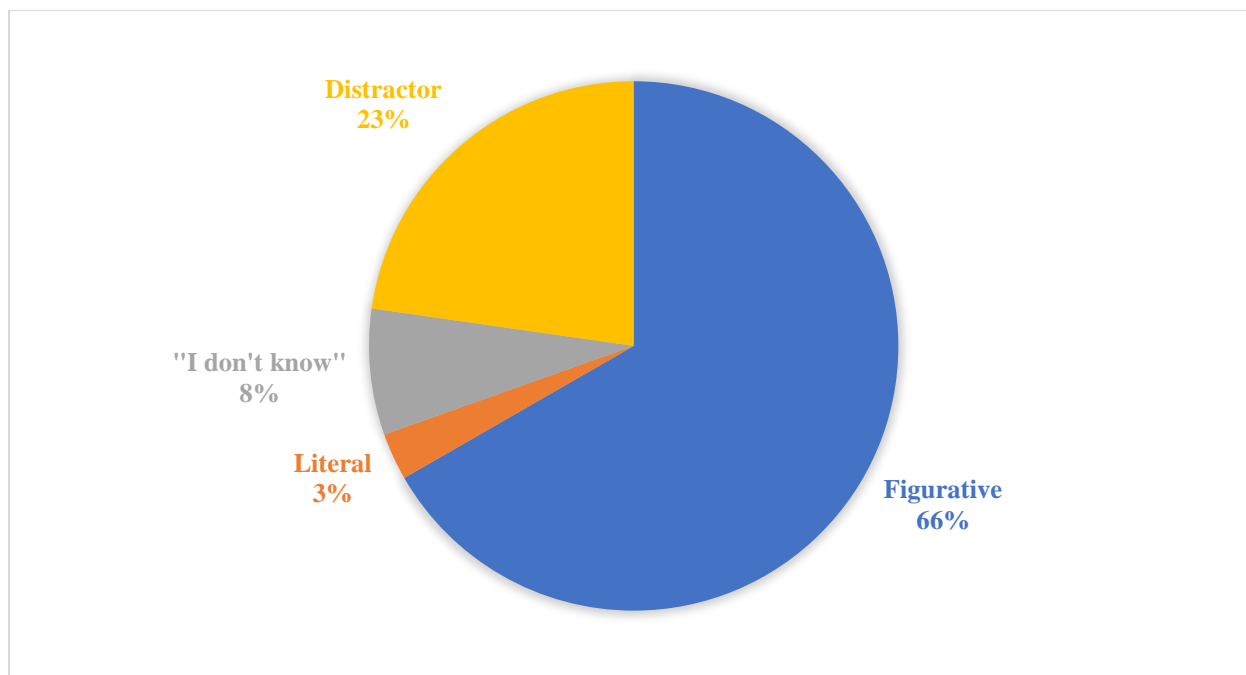
	Average Score (Max = 100%)	$N$	Std. Deviation
Situ-MWU	67%	23	21.9
Act-MWU	73%	23	17.6

Participants chose the correct definition for the Situ-MWUs 67% of the time on average, and 73% of the time when the MWU was an Act-MWU, on average. A paired t-test ( $t [22]=1.660$ ;  $p <.111$ ) showed that there was no statistically significant difference in the accuracy subjects' choices in the recognition test between test items with Situ-MWUs ( $M =66.5$ ;  $SD =21.98$ ) and those with Act-MWUs ( $M =72.61$ ;  $SD =17.64$ ).

Chart 1 shows the distribution of participant choices among the four choices (1 - the figurative meaning, 2 – the literal meaning, 3 – the distractor, 4 – “I don’t know”) for the situation items.

### Chart 1

*Situation MWUs - Distribution of Choices Among Options*



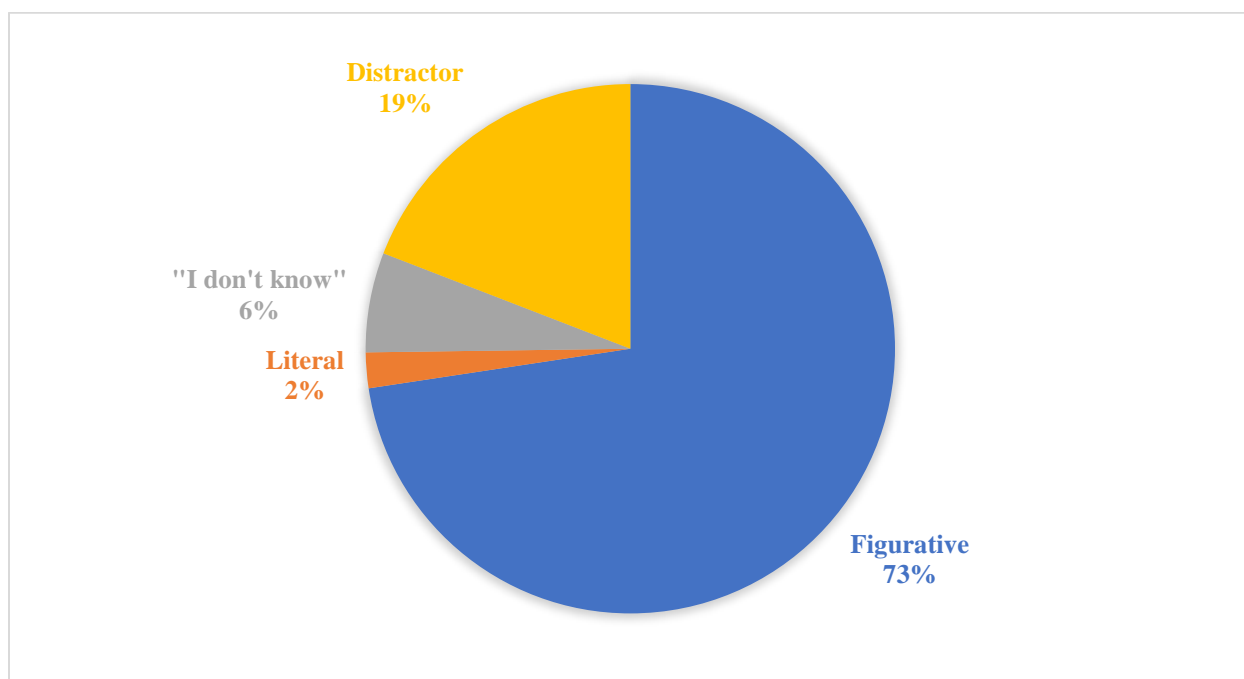
When presented with a test item in which the MWU was a Situ-MWU, participants chose the figurative interpretation (the correct answer) 66% of the time on average. They chose the distractor option 23% of the time on average; they chose “I don’t know” 8% of the time on average; and they chose the literal interpretation only 3% of the time on average. For example,

this means that in the case of a test item such as “Ben has been **on the fence** all week,” participants would have chosen the figurative meaning, “Undecided between two things” more frequently than the distractor, “Feeling sick,” or the literal “Sitting on the fence.”

Chart 2 shows the distribution of participant choices among the four choices (1 - the figurative meaning, 2 – the literal meaning, 3 – the distractor, 4 – “I don’t know”) for test action items.

### Chart 2

*Action MWUs - Distribution of Choices Among Options*



When presented with a test item in which the MWU was an Act-MWU, participants chose the figurative interpretation (the correct answer) 73% of the time on average. They chose the distractor option 19% of the time on average; they chose “I don’t know” 6% of the time on average; and they chose the literal interpretation only 2% of the time on average. For example, this means that for a test item such as “There are a lot of good ways to **kick a habit**,”

participants would have chosen the figurative meaning, “Stop a bad habit” more often than the distractor, “Try something new,” or the literal “Strike a habit with your foot.”

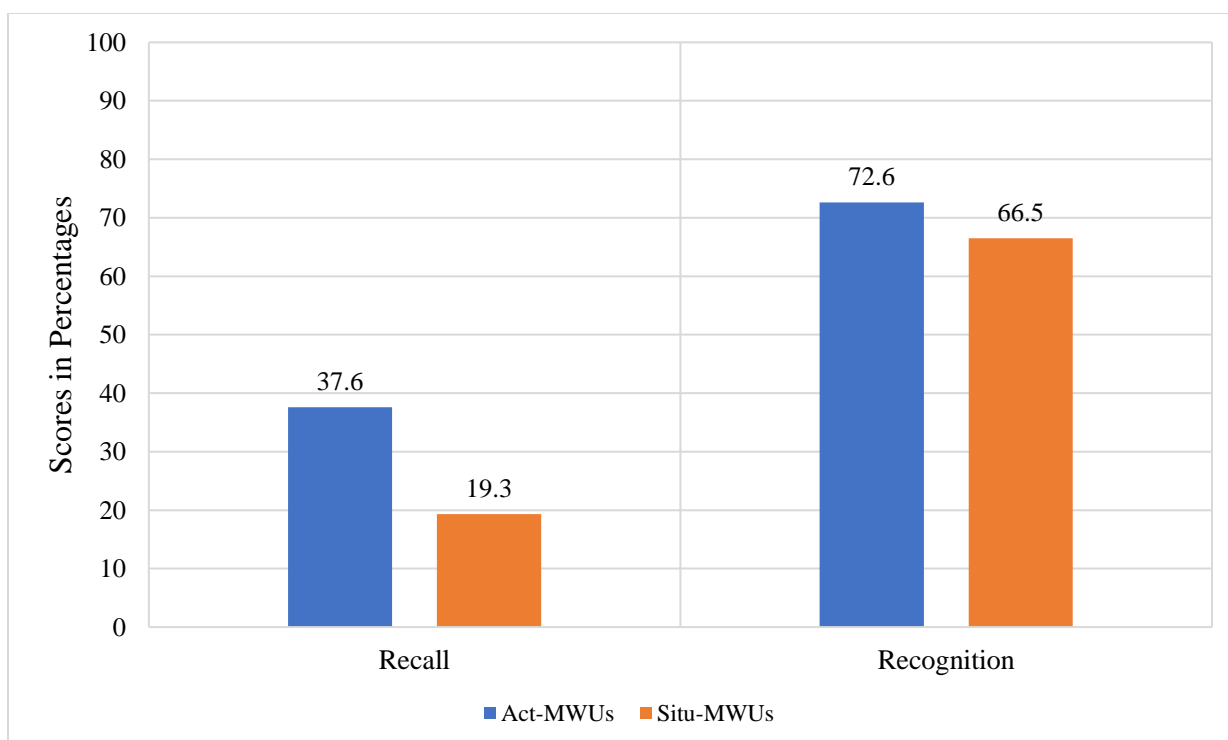
For both Situ-MWUs and Act-MWUs in the Passive Recognition test, participants most often chose the correct answer, the figurative interpretation of the MWUs. Additionally, for both Situ-MWUs and Act-MWUs, when participants chose the wrong answer, they chose the distractor options and “I don’t know” more often than they chose the literal interpretation.

### Passive Recall & Passive Recognition Compared

Chart 3 visualizes the results between both parts of the test, Passive Recall and Passive Recognition, compared.

#### Chart 3

*Average scores for Act-MWUs vs. Situ-MWUs compared for the whole test*



The average score of participant answers in the Passive Recall test were far lower than the average of participant scores in the Passive Recognition test. For the Passive Recall test, the average score for test items involving Act-MWUs was nearly 40%, but for the Passive Recognition test the average score was above 70%. The same trend is visible for Situ-MWUs. In the Passive Recall test, the average score for test items involving Situ-MWUs was around 20%, while the average score in the Passive Recognition test was nearly 70%. A set of paired t-tests were run to analyze the difference in scores between Situ-MWUs and Act-MWUs across both parts of the test – Passive Recall and Passive Recognition. The first t-test ( $t [22]=11.774; p <.000$ ) showed that there was a statistically significant difference in the accuracy subjects' answers for Situ-MWUs in the Passive Recall test ( $M =19.35; SD =15.25$ ) and the Passive Recognition test ( $M =66.50; SD =21.98$ ). The second t-test ( $t [22]=10.294; p <.000$ ) also showed that there was a statistically significant difference in the accuracy of subjects' answers for Act-MWUs in the Passive Recall test ( $M =37.60; SD =25.13$ ) and the Passive Recognition test ( $M =72.61; SD =17.64$ ).

## Chapter 5: Discussion

Prior to data collection, based on the findings of Martinez and Murphy (2011) and Kim (2016), it was hypothesized that the participants in this study would score higher on the second part of the test, the Passive Recognition section, than on the first, the Passive Recall section. This would also be in accordance with Laufer and Goldstein's (2004) implicational scale of the four degrees of strength of knowledge. The data shown above in the Results section clearly confirms this hypothesis. Just like the participants in the studies conducted by Martinez and Murphy (2011) and Kim (2016), when asked to provide the meaning of an MWU – as in the Passive Recall test – the answers given by the participants in this study were quite low in acceptability. Participant answers were, however, higher in accuracy on the second part of the test – Passive Recognition – when they only had to select the meaning from a provided list.

These findings confirm what Laufer and Goldstein (2004) suggested, that not only does Passive Recall – the ability to recall and supply the meaning of an L2 lexical item – require that the language learner have a deeper understanding of the lexical item than Passive Recognition – the ability to choose the meaning of an L2 lexical item from a list of options, but it also confirms that knowledge of lexical items does not exist in binary categories of complete understanding or no understanding at all. Though these participants could not supply the meaning of many of the MWUs in the Passive Recall test, they knew them enough to recognize the meaning and match it with the MWU in the Passive Recognition test. The findings of this study, however, add a new dimension to this body of knowledge in that where Laufer and Goldstein (2004) investigated single words, this study investigated MWUs. Yet the same principles were found to be true for MWUs as well. Therefore, Laufer and Goldstein's framework can be applied to ESL learners' knowledge of MWUs as well as to single words.

These results answer the first research question: When presented with an unfamiliar, non-compositional MWU in the context of a sentence, how effective are second language learners at guessing the meaning of these MWUs? The participants in this study, on average, were better able to give the correct meaning of the MWU when they had to choose the answer from a list provided for them than they were when asked to write in the meaning. The implications of that for language learners is that when faced with such MWUs in written or spoken English, they would have a hard time recognizing them as unique chunks of language and interpreting them by the figurative meaning, because rarely in life does vocabulary come with a list of possible meanings to choose from. This means that more often than not language learners would be interpreting what they are hearing or reading incorrectly, if it contains MWUs – which it most likely does, because MWUs are a way that native and highly proficient speakers of a language prefer to communicate (Martinex & Murphy 2011; Nattinger & DeCarrico, 1992; Pawley & Snyder, 1983; Wray, 2002).

What is of greater interest, however, because this was not a hypothesized result, is that there is a large difference in participant scores between Act-MWUs and Situ-MWUs for Passive Recall. On average, the participants scored twice as high on the definitions they provided for the MWUs that described actions than on those which described a situation. These results seem to provide the answer to the second research question: Is the semantic nature of the MWU (whether it describes an action vs. a situation) a factor in the accuracy of participant guesses? It is unclear, however, what role the semantic nature of these MWUs plays in the accuracy of their answers, because no such variance exists in the accuracy of their answers between action and situation MWUs in the Passive Recognition test. In the Passive Recognition test, there was no significant difference in the average score between the two types of MWUs. It seems that these participants



understand action MWUs at a higher level on Laufer and Goldstein's (2004) four degrees of strength of knowledge than situation MWUs. But why?

It is possible that MWUs that describe an action are more commonly used by English language speakers and writers than those that describe a situation. It could also be that situation MWUs are more misleading than action MWUs. For example, the phrases "on thin ice," "in hot water," and "out on a limb" (all included as test items in the test instrument), all communicate a level of danger, so that even if someone were to be unfamiliar with these MWUs but was trying to interpret them figuratively they would know there is some risk involved in what is being talked about. However, the risk that each MWU describes is a bit different, though they are all social risks of a kind. "On thin ice" describes the danger of being close to offending someone; "in hot water" describes being in trouble or being outside of someone's good graces already; and "out on a limb" describes knowingly taking a risk in hopes that it will have positive results. These subtleties of meaning could make such MWUs more difficult for language learners than those like "break the bank" or "hit the books" which, in having less transparent meanings as idioms, seem to point more towards a figurative interpretation than a literal interpretation.

There is some debate within discussions of conceptual metaphor as to whether, and to what degree, people use their knowledge of an MWU's "underlying conceptual metaphor" to interpret MWUs when they encounter them (Gibbs, Bogdanovich, Sikes & Barr, 1997, p. 150). Gibbs, Bogdanovich, Sikes & Barr (1997) suggest that familiarity with an idiom may render this connection unnecessary (p. 150). This processing may be more automatic. It could be that more of the Act-MWUs were familiar and processed more automatically than the Situ-MWUs. The Act-MWUs could have been processed as whole chunks, with no reference to the underlying metaphor. In contrast, if the Situ-MWUs were less familiar, processed more as individual words

in a string, it is still possible that the underlying metaphor was understood through a mental picture, but the subtle differences in the meaning of each of these MWUs was not readily apparent even from their metaphors. Therefore, if these MWUs were never explained to these language learners, it makes sense that they would provide incorrect answers when asked for their meaning.

Another point of interest in the data collected from this study is that in Passive Recognition, when presented with a list of possible meanings for the MWUs, participants chose the distractor options, and even the “I don’t know” option, more often than they chose the literal interpretation. They most often chose the figurative interpretations, the correct answer – as hypothesized – but in this second part of the test they were not misled any longer by possible literal interpretations very often. Perhaps the participants were clued in enough to the goal of the task that they knew the literal meaning could not be the correct answer. Another possibility is that perhaps the literal interpretation was too heinous to be considered a possible right answer, as was probable in the case of the MWU “have the floor” – the literal interpretation provided was “take the floor home with him.”

An informal look at the data indicated that neither length of stay in the US nor amount of time spent studying English, whether participants began studying as children in their home countries or as adults, seemed to be salient variables in their scores. It is possible that social factors – such as relationships with native English speakers – or interest in the culture of English-speaking countries – like watching movies, TV shows, reading books, etc. – could play a factor. Other factors within each student would have to be looked at to find the reason, but this is beyond the scope of the background survey conducted here.

## Chapter 6: Conclusion

Even in this brief study, it is clear that vocabulary knowledge is so much more than the form-meaning connection of single words. While the individual words within the figurative MWUs looked at in this study contribute to their meaning in some ways, as discussed within the field of conceptual metaphor, the full figurative meaning cannot be known simply from the parts. The participants within this study understood that to some extent, as most of them refused to choose the literal interpretation of these MWUs in the Passive Recognition section of the test. It would appear, then, that deceptive transparency (Laufer, 1989), was not as much of an issue as originally hypothesized. However, the figurative MWUs looked at in this study were still largely unknown to these participants at a level where they could understand them without assistance – the Passive Recall level. Yet, because such figurative MWUs are a communicatively powerful tool used frequently by both highly proficient and native speakers of English, it can be assumed that such MWUs create obstacles for learners at a passive level in the real world – understanding them when they encounter them in a reading passage or in spoken English.

One of the most important findings from this study was that these ESL learners have a complex understanding of MWUs. Though their scores, on average, were quite low on the Passive Recall test, their scores, on average, nearly doubled on the Passive Recognition test, showing some shallow level of knowledge of the meaning of these MWUs. Another important finding was that the semantic function of figurative MWUs seems to have some influence on whether they are easily understood by language learners. This can be seen in the fact that these participants, on average, scored higher on test items that included the action MWUs as opposed to situation MWUs during the Passive Recall test.

## **Limitations**

One of the main limitations of this study is the conflation of syntax and semantics as variables in the MWUs tested. A goal of the test, to answer the second research question, was to see if the semantic nature of the MWU was a factor in the accuracy of participant answers. It is admittedly unclear whether it was the syntactic structure of the MWUs or the semantic function of them which most influenced the participants correct or incorrect answers on the test instrument, as the syntactic structure of the MWUs was not something that was able to be isolated for. This is because the majority of idioms that express an action take the form of a verb phrase and those that describe a situation take the form of a prepositional phrase. There was a clear difference in participant accuracy as it relates to these two types of MWUs, but it is unclear if that is a result of the semantics or the syntax. All factors that were able to be held consistent for the MWUs, however, were, such as length of the MWU and the words in the MWU (those within the first 2,000 most commonly used English words).

A second limitation could be that the filler items took a variety of syntactic forms and performed a variety of semantic functions. The filler items were not considered in the final analysis for this reason. They were meant only to aid the participants and relieve any test fatigue.

In addition, another limitation is the sample size. To have a clearer picture of the depth of knowledge of second language learners, a larger sample size is necessary. It would also be interesting to see what the results would be if this study was conducted with groups of students from across several different language backgrounds as well.

## **Pedagogical Implications**

The results from this study confirm what others – such as Martinez and Murphy (2011) and Kim (2016) – have already been saying: MWUs are particularly difficult for second

language learners. However, their understanding of MWUs is complex – they seem to have some basic understanding of certain MWUs but not a deep, more productive knowledge of many. As Chart 3 shows above, even when L2 learners cannot provide the meaning of an MWU on their own, they are able to infer the correct meaning of an MWU when presented with possible options. The role of the teacher, then, is to help the students deepen their knowledge so that it becomes more productive – so that when they encounter such MWUs in a conversation or in a text they are reading (which is essentially a passive recall task) they can comprehend the input.

Because the difference in levels of accuracy – 0.000 – is statistically significant between both parts of the test – Passive Recall and Passive Recognition – for all MWUs in this study, we can expect this to be true for all Arabic L1 speakers of English, not limited to these participants. Therefore, what language teachers need to know is that explicit instruction in and explanation of MWUs in class is of vital importance. Language learners may have a shallower level of knowledge about these MWUs simply from exposure to them, as in incidental learning. However, their inability to put that understanding into words shows a gap in their knowledge that language teachers can fill. Because MWUs are so commonly used and so important for proficient use, language teachers cannot overlook this aspect of language instruction in their classrooms. In addition, language learners need to be provided with opportunities to practice using MWUs in productive ways, not merely in recognition type tasks.

## References

- Alali, F.A. & Schmitt, N. (2012). Teaching formulaic sequences: the same as or different from teaching single words? *TESOL Journal*, 3,2, 153-180.
- Ellis, N. C. (1996). Sequencing in SLA: Phonological Memory, Chunking, and Points of Order. *Studies in Second Language Acquisition*, 18, 91–126.
- Elmqvist, E.A. (2014) “A Comparison of Non-Native Speakers’ Depth of Knowledge of Frozen Idioms to that of Native Speakers of English.” (Unpublished master’s thesis). St. Cloud State University, St. Cloud, MN, USA.
- Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text*, 20, 29–62.
- Fernando, C. (1996). *Idioms and Idiomaticity*. Oxford: Oxford University Press.
- Folse, K. S. (2011). Myth 1: In learning another language, vocabulary is not as important as grammar or other areas . In *Vocabulary Myths* (pp. 19–34). Ann Arbor: University of Michigan Press.
- Gibbs, R. W., Bogdanovich, J. M., Sikes, J. R., & Barr, D. J. (1997). Metaphor in Idiom Comprehension. *Journal of Memory and Language*, 37, 141–154.
- Jones, M., & Haywood, S. (2004). Facilitating the acquisition of formulaic sequences. In *Formulaic Sequences* (Vol. 9, pp. 269–292). Philadelphia, PA: John Benjamins Publishing Company.
- Kim, Choonkyong (2016) L2 learners’ recognition of unfamiliar idioms composed of familiar words, *Language Awareness*, 25:1-2, 89-109, DOI: 10.1080/09658416.2015.1122025
- Lakoff, G., & Johnson, M. (1980a). The Metaphorical Structure of the Human Conceptual System. *Cognitive Science*, 4, 195–208.

- Lakoff, G., & Johnson, M. (1980b). Conceptual Metaphors in Everyday Language. *The Journal of Philosophy*, 78(8), 453–486.
- Laufer, B. (1989). A Factor of Difficulty in Vocabulary Learning: Deceptive Transparency. *AILA Review*, 6, 10-20.
- Laufer, B., & Goldstein, Z. (2004). Testing Vocabulary Knowledge: Size, Strength, and Computer Adaptiveness. *Language Learning*, 54(3), 399–436.
- Lin, P. M. (2012). Sound Evidence: The Missing Piece of the Jigsaw in Formulaic Language Research. *Applied Linguistics*, 33(3), 342-347.
- Liu, D. (2008). *Idioms: Description, Comprehension, Acquisition, and Pedagogy*. New York, NY: Routledge.
- Martinez, R., & Murphy, V. A. (2011). Effects of Frequency and Idiomaticity on Second Language Reading Comprehension. *TESOL Quarterly*, 45(2), 267-290.
- Moon, R. (1998). *Fixed Expressions and Idioms in English*. Oxford: Clarendon Press.
- Nation, I. S. P. (2013). Finding and learning multiword units. In *Learning Vocabulary in Another Language* (pp. 479-513). Cambridge, UK: Cambridge University Press.
- Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. New York: Oxford University Press.
- Pawley, A., & Snyder, F. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In *Language & Communication* (pp. 186-225). London: Longman.
- Swinney, D. A., & Cutler, A. (1979). The Access and Processing of Idiomatic Expressions. *Journal of Verbal Learning and Verbal Behavior*, 18, 523–534.
- Web VocabProfile Classic. (n.d.). Retrieved from <https://www.lex tutor.ca/vp/eng/>.
- Wray, A. (2000). *Formulaic Sequences in Second Language Teaching: Principle and Practice*.

*Applied Linguistics*, 21(4), 463–489

Wray, A. (2002). *Formulaic Language and the Lexicon*. New York, NY: Cambridge University Press.



## Appendix A: IRB Approval



### Institutional Review Board (IRB)

720 4th Avenue South AS 210, St. Cloud, MN 56301-4498

**Name:** Eleanor Jefferson  
**Email:** ejefferson@stcloudstate.edu

### IRB PROTOCOL DETERMINATION: **Exempt Review**

**Project Title:** Language Users Interpretation of English Expressions

**Advisor:** Choonyong Kim

The Institutional Review Board has reviewed your protocol to conduct research involving human subjects. Your project has been: **APPROVED**

Please note the following important information concerning IRB projects:

- The principal investigator assumes the responsibilities for the protection of participants in this project. Any adverse events must be reported to the IRB as soon as possible (ex. research related injuries, harmful outcomes, significant withdrawal of subject population, etc.).

- For expedited or full board review, the principal investigator must submit a Continuing Review/Final Report form in advance of the expiration date indicated on this letter to report conclusion of the research or request an extension.

- Exempt review only requires the submission of a Continuing Review/Final Report form in advance of the expiration date indicated in this letter if an extension of time is needed.

- Approved consent forms display the official IRB stamp which documents approval and expiration dates. If a renewal is requested and approved, new consent forms will be officially stamped and reflect the new approval and expiration dates.

- The principal investigator must seek approval for any changes to the study (ex. research design, consent process, survey/interview instruments, funding source, etc.). The IRB reserves the right to review the research at any time.

If we can be of further assistance, feel free to contact the IRB at 320-308-4932 or email ResearchNow@stcloudstate.edu and please reference the SCSU IRB number when corresponding.

**IRB Chair:**

Dr. Benjamin Witts  
 Associate Professor- Applied Behavior Analysis  
 Department of Community Psychology, Counseling, and Family Therapy

**IRB Institutional Official:**

Dr. Latha Ramakrishnan  
 Interim Associate Provost for Research  
 Dean of Graduate Studies

#### OFFICE USE ONLY

<b>SCSU IRB#</b> 1940 - 2502	<b>Type:</b> Exempt Review	<b>Today's Date:</b> 12/19/2019
<b>1st Year Approval Date:</b> 12/19/2019	<b>2nd Year Approval Date:</b>	<b>3rd Year Approval Date:</b>
<b>1st Year Expiration Date:</b>	<b>2nd Year Expiration Date:</b>	<b>3rd Year Expiration Date:</b>

## Appendix B: Pilot Test

Directions:

- Put an X in front of the expressions listed below that you know an idiomatic meaning for (a meaning different from the literal understanding of these words).
- Write that idiomatic meaning for the expression in the space.
- You may write in English or in your first language. If you do not understand the meaning, you can leave both spaces blank.

Example:

1.  Ask for trouble \_\_\_\_\_ to act in a way that is dangerous or will cause a problem \_\_\_\_\_
2.  Brush up on (something) \_\_\_\_\_
3.  Hit the books \_\_\_\_\_
4.  Kick a habit \_\_\_\_\_
5.  Beat the clock \_\_\_\_\_
6.  See the light \_\_\_\_\_
7.  Test the waters \_\_\_\_\_
8.  Pull some strings \_\_\_\_\_
9.  Spill the beans \_\_\_\_\_
10.  Beat around the bush \_\_\_\_\_
11.  Break the bank \_\_\_\_\_
12.  Ask for the moon \_\_\_\_\_
13.  Drop the ball \_\_\_\_\_
14.  Turn a blind eye \_\_\_\_\_
15.  Bite your tongue \_\_\_\_\_
16.  Show your true colors \_\_\_\_\_
17.  Tell it like it is \_\_\_\_\_

18. \_\_\_ Have the floor \_\_\_\_\_
19. \_\_\_ In a tight spot \_\_\_\_\_
20. \_\_\_ On top of the world \_\_\_\_\_
21. \_\_\_ On the spot \_\_\_\_\_
22. \_\_\_ On the fence \_\_\_\_\_
23. \_\_\_ In hot water \_\_\_\_\_
24. \_\_\_ Out of hand \_\_\_\_\_
25. \_\_\_ Out of my hands \_\_\_\_\_
26. \_\_\_ Out on a limb \_\_\_\_\_
27. \_\_\_ On thin ice \_\_\_\_\_
28. \_\_\_ Out of the woods \_\_\_\_\_
29. \_\_\_ In deep water \_\_\_\_\_
30. \_\_\_ On shaky ground \_\_\_\_\_
31. \_\_\_ On the wrong track \_\_\_\_\_
32. \_\_\_ On the same page \_\_\_\_\_
33. \_\_\_ In over your head \_\_\_\_\_
34. \_\_\_ On pins and needles \_\_\_\_\_
35. \_\_\_ A piece of cake \_\_\_\_\_
36. \_\_\_ Follow your heart \_\_\_\_\_
37. \_\_\_ Save the day \_\_\_\_\_
38. \_\_\_ Cost an arm and a leg \_\_\_\_\_
39. \_\_\_ Kill two birds with one stone \_\_\_\_\_
40. \_\_\_ Clear as mud \_\_\_\_\_

### Appendix C: Informed Consent Form

- The information collected in this study will be used for academic research only.
- This is NOT a test. You will NOT be given a grade for your results.
- Your participation will NOT affect your grades in any way.
- The results will only be used for academic research only.
- Your name will NOT be used after the data have been recorded for analysis.
- The results of the research may be published or presented. However, your name will NEVER be used.
- Your participation is **voluntary**. You may withdraw at any time.
- Your relationship with the researcher or the university will not be affected even if you decide NOT to do this.

If you give your permission to use the data for research, please sign your name.

Signature: \_\_\_\_\_ Date: \_\_\_\_\_

### Appendix D: Background Survey

1. Name: \_\_\_\_\_
2. Home Country: \_\_\_\_\_
3. Gender:        Male ( )                      Female ( )
4. Age: \_\_\_\_\_ years old.
5. At what age did you start studying English? \_\_\_\_\_ years old.
6. How long have you been in the US? (Or any other English-speaking countries):  
\_\_\_\_\_ years \_\_\_\_\_ months

For questions, please contact  
Eleanor Jefferson  
[ejefferson@stcloudstate.edu](mailto:ejefferson@stcloudstate.edu)

Or

Dr. Choon Kim  
Department of English  
[ckim@stcloudstate.edu](mailto:ckim@stcloudstate.edu)

### Appendix E: The Test Instrument: Parts 1 & 2

Name: \_\_\_\_\_

### Part 1 – Passive Recall

Directions: The **underlined bold phrase** in the sentences below are idioms. Do you know the meaning of these idioms? If so, write the meaning in the space provided. You may write in English or in Arabic.

1. Alice told me that she is **in over her head**.

\_\_\_\_\_

2. Their relationship has always been **on shaky ground**.

\_\_\_\_\_

3. You have to **beat the clock** if you want to win.

\_\_\_\_\_

4. Sarah said she would need to **brush up on** her Spanish.

\_\_\_\_\_

5. They realized that they were **on the wrong track**.

\_\_\_\_\_

6. Caroline is really good at **telling it like it is**.

\_\_\_\_\_

7. Please try not to **drop the ball** this time.

\_\_\_\_\_

8. I will need to **pull some strings** to make this happen.

\_\_\_\_\_

9. John knew he was **in hot water** after their conversation.

\_\_\_\_\_

10. You need to be careful, you are **on thin ice.**

---

11. She **showed her true colors** when she said that today.

---

12. Sam was allowed to **have the floor** during the meeting.

---

13. I am glad to see that we are **on the same page.**

---

14. There are a lot of good ways to **kick a habit.**

---

15. Nathan told Amy that she was **asking for the moon.**

---

16. John and Sam want to **test the waters** this week.

---

17. Jenny hoped that Dan would **see the light** and agree with her.

---

18. It would be better if you did not **beat around the bush.**

---

19. They told us that Blake is not **out of the woods** yet.

---

20. The tests for this class are **a piece of cake.**

---

21. My advice is that you should **follow your heart.**

---

22. Ben has been **on the fence** all week.

---

23. I am sorry but this is **out of my hands.**

---

24. She put me **on the spot** this morning.

---

25. If you do that you are going to **break the bank.**

---

26. Andrea accidentally **spilled the beans** yesterday.

---

27. I have been **on pins and needles** all day.

---

28. James was glad he **killed two birds with one stone** today.

---

29. Chelsea really went **out on a limb** in her presentation.

---

30. Mark said he is going to **hit the books** tonight.

---



Name: \_\_\_\_\_

### Part 2 – Passive Recognition

Directions: The **underlined bold phrase** in the sentences below are idioms. Circle the correct meaning of these idioms from the list provided.

1. Nathan told Amy that she was **asking for the moon.**
  - a. Asking for something that is impossible.
  - b. Asking to look at the moon.
  - c. Asking for him to get her the moon.
  - d. I don't know
  
2. Sarah said she would need to **brush up on** her Spanish.
  - a. Forget about
  - b. Sweep dirt off of
  - c. Review
  - d. I don't know
  
3. James was glad he **killed two birds with one stone** today.
  - a. Did what he needed to
  - b. Solved two problems with one solution.
  - c. Hit two birds with the same stone and killed them.
  - d. I don't know
  
4. Please try not to **drop the ball** this time.
  - a. Fail and disappoint people
  - b. Let go of the ball
  - c. Play games
  - d. I don't know

5. It would be better if you did not **beat around the bush**.
  - a. Ask too many questions
  - b. Delay saying the most important part
  - c. Hit the bush on every side
  - d. I don't know
  
6. My advice is that you should **follow your heart**.
  - a. Do what your feelings tell you.
  - b. Walk to the beat of your heart.
  - c. Not do anything.
  - d. I don't know
  
7. They realized that they were **on the wrong track**.
  - a. Making bad decisions
  - b. On the wrong side of town
  - c. Standing on the wrong part of the road
  - d. I don't know
  
8. John knew he was **in hot water** after their conversation
  - a. In trouble
  - b. Confused
  - c. Inside a pool of hot water
  - d. I don't know
  
9. Chelsea really went **out on a limb** in her presentation.
  - a. Took a risk and tried something new
  - b. Spoke very clearly
  - c. Climbed out on a tree branch

- d. I don't know
10. If you do that, you are going to **break the bank.**
- a. Damage the building that stores money
  - b. Hurt yourself
  - c. Spend all your money
  - d. I don't know
11. Their relationship has always been **on shaky ground.**
- a. Fun
  - b. Difficult; not peaceful
  - c. Standing on ground that is moving
  - d. I don't know
12. You need to be careful, you are **on thin ice.**
- a. On a thin piece of ice that could break
  - b. About to offend someone
  - c. Standing too far away
  - d. I don't know
13. Sam was allowed to **have the floor** during the meeting.
- a. Dance
  - b. Share his ideas
  - c. Take the floor home with him
  - d. I don't know
14. Alice told me that she is **in over her head.**
- a. Very busy and anxious

- b. Inside a room with things above her
  - c. Having a good time
  - d. I don't know
15. I am sorry but this is **out of my hands.**
- a. Not something I am holding
  - b. Something I do not understand
  - c. Out of my control
  - d. I don't know
16. Jenny hoped that Dan would **see the light** and agree with her.
- a. Notice the lamp
  - b. Realize the truth
  - c. Stop talking
  - d. I don't know
17. I am glad to see that we are **on the same page**
- a. At the same place.
  - b. Looking at the same page of the book.
  - c. Agreeing with each other.
  - d. I don't know
18. John and Sam want to **test the waters** this week.
- a. Start their work
  - b. Try something and see how it goes
  - c. Feel the temperature of the water
  - d. I don't know

19. They told us that Blake is not **out of the woods** yet.
- Safe
  - Out of the trees
  - Here
  - I don't know
20. Andrea accidentally **spilled the beans** yesterday.
- Fell over something
  - Told the secret
  - Dropped a can of beans
  - I don't know
21. I will need to **pull some strings** to make this happen.
- Move some strings around
  - Use my connections to powerful people
  - Work hard
  - I don't know
22. There are a lot of good ways to **kick a habit**.
- Try something new
  - Stop a bad habit
  - Strike a habit with your foot
  - I don't know
23. She **showed her true colors** when she said that today.
- Showed who she really is
  - Showed everyone her paint colors
  - Showed how much she doesn't know

- d. I don't know
24. The tests for this class are **a piece of cake**
- a. A sweet I can eat
  - b. Enjoyable
  - c. Very easy
  - d. I don't know
25. I have been **on pins and needles** all day.
- a. Tired
  - b. Sitting on pins and needles
  - c. Worried
  - d. I don't know
26. She put me **on the spot** this morning.
- a. At the right place
  - b. In the center of everyone's attention
  - c. On top of the mark on the ground
  - d. I don't know
27. You have to **beat the clock** if you want to win.
- a. Hit the clock
  - b. Work faster than the time limit
  - c. Try very hard
  - d. I don't know
28. Ben has been **on the fence** all week.
- a. Feeling sick

- b. Sitting on the fence
  - c. Undecided between two things
  - d. I don't know
29. Caroline is really good at **telling it like it is.**
- a. Not asking questions
  - b. Expressing her honest opinions
  - c. Saying things correctly
  - d. I don't know
30. Mark said he is going to **hit the books** tonight.
- a. Strike the book with his hand
  - b. Study hard
  - c. Read the books
  - d. I don't know