

Fall 1-7-2020

Led Down the Garden Path: Cognitive Processing of English Language Idioms

Daniel Alan Vandehey
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds



Part of the [English Language and Literature Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

Vandehey, Daniel Alan, "Led Down the Garden Path: Cognitive Processing of English Language Idioms" (2020). *Dissertations and Theses*. Paper 5355.
<https://doi.org/10.15760/etd.7228>

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Led Down the Garden Path:
Cognitive Processing of English Language Idioms

by

Daniel Alan Vandehey

A thesis submitted in partial satisfaction of the
requirements for the degree of

Master of Science
in
Communication

Thesis Committee:
L. David Ritchie, Chair
Jeffrey D. Robinson
Brian Manata

Portland State University
2019

Abstract

The discourse community surrounding the study of idioms often disagree regarding the proper classification of idioms. Some scholars believe that the figurative meaning of many idioms cannot be determined through the deconstruction of their lexical parts, and therefore should be classified as an irreducible lexical unit. Others believe that even though the figurative meaning of some idioms are obscured, people process these opaque idioms as they would a metaphor, constructing meaning through an integration process that includes the component pieces of the idiom, the surrounding context, and the message receiver's knowledge network. This study forwards research concerning whether idioms should be regarded and studied as lexical units or multiword constructs. Results were inconclusive and did not support the study's hypotheses. Additional analysis of the data and literature surrounding language and conversation suggest the study was confounded by the participants' ability to predict meaning in the audio stimuli before the utterances in the stimuli were complete.

Dedication

I dedicate this work to my wife, Dawn, my son, Shane, and my daughter, Emma. The sacrifices each has made in support of my scholarly pursuits far exceed what I could expect of them in good conscience.

Table of Contents

Abstract.....	i
Dedication.....	ii
List of Tables.....	iv
List of Figures.....	v
American Psychological Association (APA) Title Page.....	1
Introduction.....	2
The Idiom.....	3
The Lexical Representation Hypothesis.....	12
The Construction-Integration / Latent Semantic Analysis Model.....	15
Hypotheses.....	19
Method.....	21
Data Analysis.....	30
Results.....	32
Discussion.....	36
Conclusion.....	47
References.....	50
Appendix A. Experiment Audio-Clip Script and Images with Response Data.....	56
Appendix B. Recruitment Materials.....	68

List of Tables

Table 1: Detail of pseudo-randomization applied to the presentation of image type.....	32
Table 2: Response times grouped by image type and by timing condition.....	35

List of Figures

Figure 1: Diagram of TRE construction.....	30
Figure 2: Response times for determining congruence/incongruence between visual and audio stimuli in the during and after timing-conditions.....	34

Led Down the Garden Path: Cognitive Processing of English Language Idioms

Daniel Alan Vandehey

Portland State University

Led Down the Garden Path: Cognitive Processing of English Language Idioms

Though idioms afford native English speakers myriad short-cuts for conveying messages, they can obscure the message for those who are unfamiliar with, do not understand, or misinterpret their figurative meanings (Kim, 2016). Linguists have been debating the de-constructible nature of the idiom in a decades-long conversation, and yet there is no consensus regarding whether an idiom should be treated as a lexical unit or de-constructible phrase. At the crux of this debate is disagreement over how idioms are stored and recalled in memory (Hillert & Baracas, 2009; Romero-Lauro, Tettamanti, Cappa, & Papagano, 2008). A better understanding of this process could be critical to students of English as a second language (ESL), as current teaching practices assume idioms must be memorized as lexical units in the same way students memorize vocabulary words (Hsieh et al., 2017; Pitzl, 2016; Saberian & Fotovatnia, 2011). If this is not how native speakers learn idiomatic language, the fundamentals of present teaching methodology may be flawed.

This study seeks to examine the current state of idiomatic language research and advance knowledge of the subject through quantitative experimental testing aimed at revealing evidence of the cognitive processes at work during the deconstruction of idioms and the subsequent construction of connotative meaning. This study begins by composing a clear and concise definition of the idiom, as well as examining taxonomies to define and order variations within idiomatic language. This study also explores two competing models designed to deal with the intricacies of idiomatic language. Within the context of each model, this text discusses the ideas behind the model and research supporting its claims. This study then discusses the implications of the current state of the research and

presents hypotheses to guide this research. This text also details the methodology implemented by the researcher to seek evidence of cognitive processes associated with making meaning out of idiomatic language.

The Idiom

This work addresses the concept of the idiom through its role in English language communication within the United States and as the dominant second language and lingua franca around the globe (Maryns, 2015; Mohammad, Yap, Chan, & Wong, 2016; Pitzl, 2016). Study of the idiom suffers from the lack of a universally accepted definition of the concept (Mohammad et al., 2016). Universal definitions are also conspicuously absent for additional multi-word sequences, such as collocations, compounds, and constructions (Saberian & Fotovatnia, 2011), often leading to idioms being mistaken for these concepts or vice-versa. These myriad definitions and other confounding phenomena increase the difficulty associated with idiom identification (Christiansen & Arnon, 2017; Kim, 2016; Mohammad et al., 2016). Much of this difficulty stems from the syntactic and semantic construction of idioms, which are correct and consistent with the literal meaning of these connotative phrases (Cain, Towse, & Knight, 2009; Maryns, 2015; Tzou, Vaid, & Chen, 2017).

Additionally, idioms exist on a spectrum of transparency, with some idioms decomposable to the extent their meaning can be deciphered out of context, and other, non-decomposable idioms whose meanings are opaque without context and often difficult to decipher even when context is available (Cacciari & Glucksberg, 1991; Saberian & Fotovatnia, 2011). In the interest of clarity, this study offers the following explication of the concept of the idiom as it will be engaged in this research.

Definitions

An examination of the literature surrounding the study of idioms reveals that many authors entertain a definition of *idiom* that is largely unique to their own work. One facet consistent through the definitions is that an idiom is a multi-word sequence and never a single word (e.g., see Chetham & Ro, 2011; Pitzl, 2016; Tzou et al., 2017). Even in this respect, the field cannot agree on a unifying name for the *multi-word sequence* (Christiansen & Arnon, 2017), with variations including *multi-word constructs* (Pitzl, 2016), *multiword constructions* (Culicover et al., 2015), and *multi-word units* (Mohammad et al., 2016; Saberian & Fotovatnia, 2011). For the purposes of this research, this text uses the terms *multi-word sequence* or *phrase* interchangeably to refer to this single concept with many labels.

Another common thread through these definitions is the inclusion of both a literal and a figurative meaning presented by the idiom (e.g., see Hsieh, Wu, & Merek, 2017; Mohammad et al., 2016; Saberian & Fotovatnia, 2011). Idioms adhere to rules of construction that leave the phrase open to literal interpretation (Cain et al., 2008; Culicover et al., 2017; Saberian & Fotovatnia, 2011), but often the purpose of the idiom manifests in its figurative interpretation. For example, the idiom *break the ice* (Vulchanova et al., 2011) can be used literally in, “Before Vlad could begin fishing in the frozen lake, he had to break the ice,” or figuratively in, “Michael heard that his new boss likes cats, so he showed her a picture of his own cat to break the ice.” In the first

sentence, Vlad cannot get his line in the water unless he literally breaks a hole in the ice. In the second sentence, Michael uses a picture of his cat as a conversation starter.

The described events in the first and second sentence have nothing in common, yet both make effective use of the words *break the ice*. However, it is only in the second sentence that *break the ice* takes on the characteristics of a multi-word sequence. Scholars differentiate the multi-word sequence from a simple string of words in the sense that the sequence of words takes on a new meaning that is often disconnected from the literal interpretation (e.g., see Cheatham & Ro, 2011; Pitzl, 2016; Tzou et al., 2017). This new meaning cannot be deciphered simply by deconstructing the sequence to ascertain meaning from its components (Culicover et al., 2015).

In addition to multi-word sequences and literal and figurative meaning as features of idioms, the discourse community also agrees that the figurative meaning of an idiom cannot be derived directly from its constituents (Cheatham and Ro, 2011; Mohammad et al., 2016; Saberian & Fotovatnia, 2011; Tzou et al., 2017). In the example, *kick the bucket* (Cheatham & Ro, 2011; Culicover et al., 2017; Kim, 2016), the figurative meaning of the idiom cannot be determined by decomposing the phrase. Literally, this phrase would suggest striking a bucket with one's foot. Figuratively, this phrase means *to die*. Nothing in the verb *kick*, or the noun *bucket*, suggests death as an alternative meaning per se. There are degrees to which the figurative meaning is obscured, which this text addresses further on, but the statement that an idiom cannot be derived directly from its constituents (out of context) is true of all types of idioms.

Idioms are also formulaic and conventionalized (Hsieh et al., 2017; Mohammad et al., 2016; Pitzl, 2016). Culicover et al. (2015) argue that in the example, *kick the bucket*,

neither the verb *kick* nor the noun *bucket* can be replaced with synonyms and maintain the figurative meaning of the idiom. In this regard, the phrase, *kick the pail*, has no figurative meaning, nor does the phrase, *boot the bucket*. This is not true of all idioms, however, as it is permissible to make changes to some forms. The phrase, *she drives me crazy*, can take several forms, such as; *she drives me nuts*, *I drive her bonkers*, and *they drive us bananas* (Culicover et al., 2015). There are also some idioms that reject conventions, such as; *for the most part*, *day in day out*, *time and again*, and *so to speak* (Culicover et al., 2015).

Ritchie and Dyehouse (2008) posit that idioms and other metaphoric language are subject to playful distortions, which can account for a phrase like *boot the pail* in lieu of *kick the bucket*. In this scenario, the objective of the speaker is to make an amusing play on words by breaking with lexical conventions. Glucksberg (2001) acknowledges such playfulness, and cites it as evidence of the compositionality of some idioms, but cautions that in many cases these type of substitutions lead to confusion, adding that he cannot understand why anyone would change the constituents of an idiom, “other than to be cute?” (p. 82). Though playfulness and humor have a place in the discussion of idioms writ large, the present focus recognizes that idioms in their undistorted form follow the conventions of the English language, which often makes them tricky to locate within texts and conversations (Christiansen and Arnon, 2017; Kim, 2016; Mohammad et al., 2016).

Lastly, idioms are commonly used expressions in spoken and written communication (Hung & Nippold, 2014; Saberian & Fotovatnia, 2011; Tzou et al., 2017). As such, a mastery of idioms is considered essential for fluency in the English

language for ESL students (Culicover et al., 2015; Hsieh et al., 2017; Kim, 2016; Pitzl, 2016; Saberian & Fotovatnia, 2011), indicating a person cannot be truly competent in everyday exchanges without a formidable knowledge of idioms and how to use them correctly in their cultural and linguistic context. In addition, Christiansen and Arnon (2017) report that up to 50% of the written text and verbal exchanges produced by native speakers is comprised of multi-word sequences, which include idioms.

Given a reasonable degree of consensus across the discourse, this text defines the concept of the idiom as a commonly used, conventionalized and formulaic multi-word sequence possessing a literal as well as a figurative meaning, and this figurative meaning cannot be derived directly from its constituents.

What an Idiom is Not

With a definition establishing what an idiom is, for the sake of even greater clarity it is important to discuss what an idiom is not. Much of the ambiguity surrounding the identification and classification of idioms in research is hampered by unclear boundaries between idioms and other multi-word sequences (Culicover et al., 2015; Mohammad et al., 2016) such as collocations, compounds, and syntactic constructions.

Collocations. Collocations are multi-word sequences that can be taken literally but adhere to certain combinations expected by English language conventions (Culicover et al., 2015; Hsieh et al., 2017). Examples include; *here and there* (but not *there and here*), *fast and furious* (but not *furious and fast*), and *make a change* (but not *do a change*). Collocations can also include light verb constructions, such as *take a walk* (but not *have/make/do a walk*). Lastly, collocations can include appropriate verb and adjective pairings, such as *become taller*, or *get taller* (but not *turn taller*). Though collocations

share some attributes with idioms, this form of multi-word sequence lacks the ambiguity of idioms, can be deconstructed, and its meaning can be derived from its parts.

Compounds. Compounds are multi-word sequences that are formed as phrases but act with meaning independent of their constructs (Culicover et al., 2015). Examples include; *bat boy* (young male human who tends the bats at a baseball game, not a human/bat hybrid), *tire track* (the imprint left on the ground by a tire rolling over the surface, not a running surface where tires go jogging), *firewood* (wood that is cut and dried for fueling a fire, not wood that perpetually burns for no known reason), and *dog person* (a person who likes dogs, not a human/dog hybrid). In these examples the literal meaning of the multi-word sequence is nonsensical. A *bat boy* has no echo-locating ability and a *dog person* does not scratch behind her ear with her foot. Though compounds do have a figurative meaning, cannot be deconstructed, and are multi-word constructs, they lack the option of being interpreted literally and therefore are not idioms.

Syntactic constructions. All the multi-word sequences this text discusses are constructions, so this category addresses constructions that do not fall under the dominion of idioms, collocations, or compounds. Culicover et al. (2015) report that these are multi-word sequences that render a sentence grammatically incorrect, yet syntactic constructions are used regularly in conversation and text. One example is, “He refuses to work, *no matter* how big the salary” (Culicover et al., 2015, p. 561). This construction omits a verb in the subordinate clause, and could be corrected to read, “He refuses to work; it does not matter how big the salary he is offered.” This type of construction has the potential to be mistaken for an idiom due to its problematic deconstruction, but idioms follow conventions that adhere to syntactic integrity. This does not, however,

prohibit syntactic constructions from producing false positives while translators and scholars work to identify abnormalities in texts and conversations that may isolate idioms (Cain et al., 2008; Christiansen and Arnon, 2017; Kim, 2016; Mohammad et al., 2016).

Taxonomies

Clarification of the definition of an idiom is only a step in identifying all types of idioms. Much like a clear definition of *automobile* could not include every make and model, there are many types of multi-word sequences scholars can confidently classify as idioms. These must be managed through taxonomies, and for the pursuits of this study, there are two taxonomies worthy of examination; Nunberg's taxonomy and Cacciari and Glucksberg's taxonomy (Saberian & Fotovatnia, 2011).

Nunberg's taxonomy. This taxonomy, based on Nunberg (1978), posits three classes of idioms sorted according to the degree of difficulty in deciphering the figurative meaning of the idiom.

Normally decomposable idioms. This first classification deals with what Nunberg regards as normally decomposable idioms. Within these idioms, at least one part of the multi-word sequence is decomposable. In the example, *pop the question* (Saberian & Fotovatnia, 2011), the word *question* can be taken as a metonym for the question, "Will you marry me?" This element of objective truth makes the idiom more susceptible to semantic analysis, as the metonymic meaning maps onto the literal meaning, leaving only part of the idiom with figurative qualities.

Abnormally decomposable idioms. Nunberg refers to this second classification as abnormally decomposable idioms. Within these idioms, at least one part of the multi-word sequence can be identified metaphorically. In the example, *play hard to get*

(Vulchanova et al. 2011), the word *play* can be deciphered as a metaphor for acting or pretending. In this regard, the receiver of the message can decompose the message metaphorically to the extent that the person described is pretending to be something they are not. In this case the described individual is pretending to be *hard to get*, which implies the person is difficult to acquire. In the context of a romantic pursuit, the idiom means the pursued is pretending to be difficult to acquire as a romantic partner, but this idiom can be used in other contexts, such as, “The potential client I am trying to sign is playing hard to get.”

Semantically non-decomposable idioms. Nunberg refers to this third classification as semantically non-decomposable idioms. Within these idioms, there are no indicators of the figurative meaning of the multi-word sequence. In the example, *carry a torch* (Cain et al., 2008), neither the word *carry* nor the word *torch* offers any inclination toward the figurative meaning of the multi-word construct, which refers to latent emotional feelings toward a person after rejection from this same person. In these cases, some argue that the speaker of the language must store the meaning in the lexicon (e.g., see Cain et al., 2009; Christiansen & Arnon, 2017; Culicover et al., 2017) while others argue the receiver must hope for clues within the context of the conversation or text to construct meaning based on semantic rules and their knowledge of the world (Kim, 2016; Kintsch, 1988; 2000; Kintch & Bowles, 2002; Ritchie, 2004).

Cacciari and Glucksberg’s taxonomy. This taxonomy, based on Cacciari and Glucksberg (1991) and later refined in Glucksberg (2001), posits that idioms fall into three general categories; opaque, transparent, and quasi-metaphorical.

Opaque idioms. This category shares qualities with Nunberg’s semantically non-decomposable idioms in that deconstructing the idiom offers no insight into its figurative meaning. In the example, *throw in the towel* (Cain et al., 2008), there are no indications from the action of *throwing* or the object of *towel* to indicate the figurative meaning of this idiom is to give up or stop fighting.

Transparent idioms. This category shares qualities with Nunberg’s classifications of normally and abnormally decomposable idioms. Hung and Nippold describe transparency as “the extent to which the literal and figurative meanings overlap” (2014, p. 209). In the example, *to cross swords with someone* (Cain et al., 2008), the multi-word sequence easily maps the act of initiating a sword fight over the concept of initiating a fight of any other kind.

Quasi-metaphorical idioms. This category contains idioms with literal meanings that map onto a metaphorical construct, or using the literal meaning of the multi-word sequence as a metaphor to describe a referent (Cain et al., 2008). In the example, *give up the ship* (Glucksberg, 2001), the figurative meaning of surrender is connoted metaphorically by a multi-word sequence that denotes a prototypical instance of surrender.

Evaluation of taxonomies. References to the transparency or opacity of idioms from Cacciari and Glucksberg (1991) appear in Cain et al., (2008), Hung and Nippold (2014), Kim (2016), and Pitzl (2016), where quasi-metaphoric idioms make no appearances and afford no examples. Nunberg’s taxonomy was essentially usurped by Cacciari and Glucksberg’s taxonomy, though Nunberg’s taxonomy offers a finer degree of measurement and clearer requirements regarding idioms that Cacciari and

Glucksberg's would simply call transparent. Throughout the literature contributing to this study, the transparency of idioms was portrayed as a spectrum, with transparency and opacity at its poles. Idioms do not always fit neatly into a trichotomy, as Cacciari and Glucksberg (1991) might suggest. Kim (2016) even warns against *deceptive transparency*, a condition where the overlap between literal and figurative meaning is so profound that highly proficient ESL students are unable to identify the existence of an idiom because the literal meaning appears to fit the context.

Though Nunberg's taxonomy offers a finer degree of classification of idioms, with normally and abnormally decomposable idioms occupying the same space as Cacciari and Glucksberg's transparent idioms, the discourse community has overwhelmingly adopted Cacciari and Glucksberg's taxonomy. However, none of the research contributing to this study references Cacciari and Glucksberg (1991), while many made liberal use of the terms *transparent* and *opaque*, though only Saberian and Fotovatnia (2011) mention quasi-metaphoric idioms. It is reasonable to surmise that researchers appreciate the clarity of the *transparent* and *opaque* labels even when they do not credit the taxonomy from which the labels sprang. Cacciari and Glucksberg's taxonomy serves the purposes of this study well, and this text credits their work.

The Lexical Representation Hypothesis

The lexical representation hypothesis (LRH) suggests that idioms are memorized and stored in the human brain in the same fashion as complex and difficult words (Caillies & Butcher, 2007; Rommers, Dijkstra, & Bastiaansen, 2013; Vulchanova, Vulchanova, & Stankova, 2011). This approach argues that one either knows the meaning of the idiom, or does not, just as one knows the definition of a word, or does not. Thus,

the LRH is predicated on the idiomatic multi-word sequence equating to a *lexical unit* (Caillies & Butcher, 2007; Hung & Nippold, 2014; Kim, 2016).

The international linguistics community favors LRH due to the problematic nature of idioms in learning English as a second language (Kim, 2016). Culicover et al. (2015) believe native speakers store and call upon thousands of multiword sequences in the lexicon. As such, a common and reliable measure of English language proficiency among language instructors and students is a student's ability to skillfully interpret and use idioms (Hsieh et al., 2017; Pitzl, 2016; Saberian & Fotovatnia, 2011). Conventional ESL course methods require students to memorize idioms as a lexical unit, as they would memorize a vocabulary word (Hsieh et al., 2017; Pitzl, 2016; Saberian & Fotovatnia, 2011). Some researchers suggest that there are as many multi-word sequences in the lexicon as individual words (approximately 80,000) whose meanings are retrieved in the same way as long words (Glucksberg, 2001; Jackendoff, 1995). However, not all agree, and some language scholars feel that instruction in ESL has not kept pace with the current research and the changing needs of students, creating a deficiency in modern teaching methods (Hsieh et al., 2017; Kim, 2016).

As students struggle with the task of learning thousands of idioms and other multiword sequences, some are concerned an LRH approach is limiting students' ability to naturally acquire new idioms and are calling for expanded research and reform. Saberian and Fotovatnia claim, "Presently, there seems to be no agreed upon model or theory. A better understanding of the way in which idioms are processed and comprehend [sic] by [second language] learners could lead to more effective ways of teaching them to [second language] learners" (2011, p. 1231).

The issue is most salient among ESL students because English continues to be the prevailing second language adopted worldwide (e.g., see Maryns, 2015; Mohammad et al., 2016; Pitzl, 2016). Since English has been deployed as a uniform standard language for international communication (Maryns, 2015), the result is a unique language phenomenon regarded as the *World Englishes* (Maryns, 2015; Mohammad et al., 2016; Pitzl, 2016). The English language has become heterogeneous across varying regions, with examples of unique dialects, such as “Pidgin English,” cropping up all over the planet (Mohammad et al., 2016). Mohammad et al. conducted content analysis research in Nigeria to identify idioms unique to the Nigerian-English dialect. The researchers compared idioms found in local newsprint to known idioms originating in Britain, The United States, and Australia to determine if the idiom had Nigerian origins. Ultimately the research team identified 124 idioms unique to the local dialect.

This type of independent evolution in the language can create barriers to understanding, as in Belgium where Maryns (2015) observed exchanges between a Belgian official and a Nigerian national during an asylum hearing. The Nigerian refugee and Belgian official were attempting to use English as a lingua franca throughout the exchange. As both parties were forced to abandon their native language for a second language, miscommunication was a prevalent issue. Contributing to this problem was the use of idioms unique to each party’s homeland. Pitzl explains that using English as a lingua franca has the benefit of offering both parties a neutral field of play (2016), but if much of the language exchanged is indecipherable the odds of a beneficial outcome for either party are greatly diminished (Maryns, 2015).

Ideally, students of the English language should build skills to acquire idioms as a native speaker would (Saberian & Fotovatnia, 2011). Vulchanova et al. (2011) hypothesizes that native-English-speaking children should be able to determine the meaning of idioms almost as well as adults once they reach the age of ten. However, the researchers found that, aside from early comprehension of idioms referencing the human body (e.g., *get wet to the bones*, or *my skin bristles up*), children progressed in their ability to decipher the meaning of idioms at a predictable rate from age six through adulthood. These results could possibly support the basic premise of LRH, but these findings contradict the hypothesis of the researchers, and the results of other studies.

The Construction-Integration / Latent Semantic Analysis Model

The Construction-Integration Model (CI) considers both the linguistic input and the message receiver's knowledge network as it proceeds through two successive stages; the first stage constructs possible meaning from the linguistic input and the second integrates the constructed meaning with prior knowledge to arrive at a coherent whole (Caillies and Butcher, 2007; Kintsch, 1988). Kintsch (2000) and Kintsch and Bowles (2002) extend the CI model to include Latent Semantic Analysis (LSA), which situates the components of metaphors in a "neighborhood" within a semantic map comprised of 300 to 400 dimensions. Words with more relevant associations are situated closer to the target word within this map than irrelevant words. Target words within the LSA can be analyzed according to their adjacency to other words, and these associations can inform the processing of metaphoric language in CI. Within this CI/LSA model, the processor uses LSA to identify semantically sound options as an input before the possible meanings are blended with prior knowledge to arrive at the correct meaning.

Kintsch (2000) uses the example of the metaphor *my lawyer is a shark*. The LSA neighborhood surrounding shark includes words like *vicious, predatory, aggressive, tenacious, fast swimmer, fish, sharp teeth, leathery skin, and gills*. When engaging CI, integration with one's knowledge base will accentuate attributes that are metaphorically relevant to describing a lawyer. It is unlikely that one's knowledge of lawyers will entertain the comparison of a lawyer to a fast swimming fish with sharp teeth, leathery skin, and gills. It is much more likely the lawyer can be compared to a creature that is vicious, predatory, aggressive and tenacious, so when these attributes are considered and integrated in the context of the knowledge network the irrelevant attributes are suppressed (Kintsch, 2000; Ritchie, 2004). This model becomes more troublesome, however, when the metaphor slips into obscurity and the phrase becomes idiomatic.

Ritchie and Dyehouse (2008) suggest that idioms likely evolved from metaphors which have since lost connection to their original meaning and are now "dead" in the sense that most people have forgotten the original reference (e.g., see also idioms as *frozen metaphors*, Caillies and Butcher, 2007). As an example, Ritchie and Dyehouse use the idiom *happy as a clam*, pointing out that the original metaphor was *happy as a clam at high tide*. This original metaphor includes conditions that one could assume make a clam happy. High tide affords the clam protection from many natural predators and facilitates their ability to feed. The idiom *happy as a clam*, however, does not describe these pleasing (to the clam in question) conditions. The meaning of the reference is not communicated in the phrase itself, but rather in the social and cultural conventions that govern use of the idiom. Kintsch (2000) addresses the limitations of LSA to include the need for an identifiable predicate and argument in the phrase, which is a much easier

requirement to fill when one knows that the high tide is making the clam happy. Since Ritchie and Dyehouse argue that an idiom like *happy as a clam* has no recoverable semantic meaning, it may seem the only way to understand the figurative meaning of the phrase is simply to memorize it.

Caillies and Butcher (2007) hypothesized that opaque idioms like *happy as a clam* would be recalled as a lexical unit, consistent with LRH. Caillies and Butcher conducted an experiment to test participants' processing time of opaque versus transparent idioms with the expectation that opaque idioms would be processed as a unit at the propositional level, and thus return faster processing times than transparent idioms which were expected to engage the more cognitively demanding CI model. Caillies and Butcher discovered that participants activated the figurative meaning of transparent idioms earlier than opaque idioms, indicating the opaque idioms were not recalled as a lexical unit, and required more time to process.

Boulenger, Hauk, and Pulvermuller (2009) conducted experiments using functional Magnetic Resonance Imaging (fMRI) to determine whether a participant's motor cortex was activated when presented with idioms containing body-action verbs. Bergen (2012) describes the study of mirror neurons, and how thinking about a motor action can activate the areas in the brain responsible for corresponding movements without eliciting movement in the body. Boulenger et al. (2009) built on this previous research using the phrases, "He grasped the idea," and "He kicked the habit" in treatment sentences, comparing reactions to these phrases to control sentences using the verbs grasped and kicked in a literal sense. The researchers found significant similarities in motor-cortex stimulation between the treatment and control groups. Also, in the vein of

cognition and neuroscience, two studies included experiments to determine if the brain processes transparent and opaque idioms in the same manner (Hillert and Baracas, 2009; Romero-Lauro, et al., 2008). Both groups concluded that although each type of idiom was processed in different areas of the prefrontal cortex, the size of the cortical networks accessed were statistically comparable.

The discourse community surrounding the study of idiomatic language does not argue over the relevance of CI/LSA, only whether people must engage this model to arrive at a correct figurative meaning for all types of idioms. Though Vulchanova et al. (2011) discovered reliable and steady growth in children regarding learning idioms, it is important to remember that the CI/LSA model will facilitate improved coherence over time as the child gains lexical, semantic, and world knowledge. Results suggesting that children simply memorized idioms as they encountered them could as easily be interpreted to indicate children ages ten years and older possess the framework necessary to construct meaning from the knowledge they possess and improve their ability to construct meaning as their general knowledge of the world around them improves.

The greater cognitive load regarding opaque idioms discovered in Caillies and Butcher (2007), which compliments the comparable size of the cortical networks accessed during exposure to opaque versus transparent idioms (Hillert & Buracas, 2009; Romero-Lauro, Tettamanti, Cappa, & Papagano, 2008) also supports CI/LSA in the sense that idioms situated near the poles of the transparency spectrum require about the same extent of access across cortical networks while affording opaque idioms no cognitive load advantage. Activation of the motor cortex recorded in Boulenger et al. (2009) indicates the participants were not processing the idioms as a lexical unit (which would register no

bodily action in the arms or legs) and were constructing meaning from the components of the multi-word sequence. The sum of these findings indicates idioms of all types are decomposed and analyzed as opposed to simply returning a concise definition as would be expected from a lexical unit.

Boulenger et al. (2009) presents compelling evidence of the cognitive processes activated by the concept of *embodied simulation* (Bergen, 2012). Bergen argues that humans visualize the meaning of messages as they process language and one effect of this simulation is the stimulation of the motor cortex in areas controlling the parts of the body the message describes as being active (e.g., low level stimulation of the motor cortex area controlling forearm muscles when hearing a message about grabbing a hammer). Boulenger et al. (2009) reports that fMRI imaging recorded this type of activity corresponding to idioms with literal meanings that relate to body-part actions (e.g., *grasp the idea, kick the habit*) This discovery lends credence to the argument for the CI/LSA Model. However, Beck (2010) cautions that fMRI scans are incapable of contributing to the understanding of complex thoughts, such as processing language, and that this type of analysis falls outside the scope of effective fMRI research. This assumes fMRI scans are reliable from the start. Ecklund, Nicols, and Knutsson (2016) conducted a series of inter-reliability tests, examining the results of hundreds of fMRI cases, and claim the rate of false positives could be as high as 70%. This, according to Ecklund and his colleagues, calls the results of as many as 40,000 published articles into question.

Hypotheses

With the validity of fMRI research under scrutiny, additional research will benefit from the pursuit of alternative cognitive experimentation. This study conducts a cognitive

experiment that works in concert with embodied simulation experiments recounted in Bergen (2012). The basis of this premise is that people are faster to identify a picture if it is congruent with what they are already envisioning than if it is incongruent. If the CI/LSA model is valid, a person must consider the literal meaning of an idiom before determining the figurative meaning. When hearing the idiom *spill the beans*, the literal meaning conjures images of actual beans spilling, perhaps from a pot, a bowl, or a can. In this example, if the message receiver first considers the literal interpretation of the idiom, she or he will first visualize beans spilling before arriving at a figurative interpretation of the phrase. As such, if asked to confirm whether a photo corresponds to the phrase, this person should be quicker to confirm a picture of someone spilling a bowl full of beans (match/congruent) than a picture depicting the figurative meaning (mismatch/incongruent). If the message receiver processes the idiom as a lexical unit, which only attends to the figurative meaning, she or he will visualize someone divulging confidential information in a conspicuous manner. Perhaps the receiver will visualize a criminal suspect confessing the perpetration of a crime to a police officer. If this is the case, the receiver should be able to confirm a picture depicting a confessing criminal (congruent) faster than a picture of someone literally spilling beans (incongruent).

Another consideration concerns what happens when the figurative interpretation supplants the literal interpretation. Bergen (2012) suggests that the effects of embodied simulation are fleeting, and as one visual construction is replaced by the next, the old construction is swiftly forgotten (see also Christiansen & Chater, 2016). This premise, when applied to the example above, suggests that when the receiver recognizes *spill the beans* as an idiom, the image of spilling beans will be supplanted by the confessing

suspect (or something similar). The transition from literal simulation to figurative simulation should progress as a function of time. Therefore, the longer the person can process the idiom the more likely the figurative simulation will manifest. Based on these considerations, this study forwards the following hypotheses.

- H1: When listening to dialog containing an idiom, participants will be quicker to identify the congruence of the idiom and an image depicting the literal meaning of the idiom (as opposed to the figurative meaning) if the image is viewed during the utterance of the idiom.
- H2: When listening to dialog containing an idiom, participants will be quicker to identify the congruence of the idiom and an image depicting the figurative meaning of the idiom (as opposed to the literal meaning) if the image is viewed shortly after the utterance of the idiom.

Method

Participants

The researcher recruited participants in person, via Desire-2-Learn (D2L) posts, and via email. The study targeted participants based on their enrollment in select undergraduate courses in the Department of Communication at Portland State University (PSU). Instructors for these courses offered potential participants extra-credit in the enrolled courses in exchange for participation. Seventy-five volunteers took part in the experiment, however, data from one participant were removed for a computer/software malfunction, and data from a second participant were removed due to a failure to follow

instructions that resulted in faults across all timed responses. This left an analytical sample of 73 participants ($N = 73$).

The sample was comprised of adults (38 women, 29 men, 6 others, $M_{\text{age}} = 24.55$ years [$SD = 5.20$], age range: 19-47 years.). Most participants were native English speakers ($n = 61$), with the balance identifying as ESL-Fluent ($n = 10$) or ESL-Intermediate ($n = 2$). Participants' education level varied between Associate's degree ($n = 32$), some college ($n = 30$), high school graduate ($n = 5$), Bachelor's degree ($n = 4$), and Doctorate ($n = 2$). As well, ethnicity varied between White/Caucasian ($n = 41$), Mixed Ethnicity ($n = 17$), Asian ($n = 7$), Latinx ($n = 3$), Other ($n = 3$), and Black/African American ($n = 2$).

The researcher provided participants with two copies of the informed consent form (see appendix); one which they signed and returned and the other which they retained for their own records. The researcher collected all signed informed consent forms, then consolidated and randomized these forms and sealed them in a manila envelope where they have remained separate from other collected data.

The researcher kept a log of participants' names for the purpose of reporting participation to Instructors that they may issue extra-credit to their participating students. These logs were transferred to an electronic spreadsheet where participants were listed alphabetically rather than in the chronological order in which each participated. The researcher destroyed the original logs as soon as the information was transferred and generated extra-credit participation lists from the spreadsheet file.

Participants performed the experiment at a computer station where the researcher logged them into the program using a sequential numerical identifier that was

disconnected from any identifying information collected from the participants. At no point through the experiment did the researcher ask the participants for any identifying information.

Components of the Experiment

Timed response exercise (TRE). A TRE is defined here as a single iteration within the experiment that includes an *audio stimulus*, a *visual stimulus*, and a *response* from the participant that is timed and recorded.

Audio stimuli. Each of the twelve TREs begins with an audio stimulus that is a recorded verbal exchange between a female and male actor. Each starts with a simple question and ends with a response (e.g., Question: “Can you try to get along with Ronny tonight?” Response: “Yes, but you know that guy drives me up the wall.”). In six of the twelve TREs, the response ends with an idiomatic phrase (e.g., Question: “Are you going to work things out with Stephanie?” Response: “No, we are both ready to throw in the towel.”). In the remaining six TREs, neither the question nor the response contains an idiomatic phrase (e.g., Question: “Would this shade of purple go with a bright orange?” Response: “You cannot possibly like this for the kitchen.”). All verbal stimuli end with a concrete noun (e.g., puppy, taxi, bucket, beans, train, etc.) to assure that images depicting the literal meaning of the phrase uttered in the responses are easy to identify in both the idiomatic and non-idiomatic phrase conditions.

Bonin, Meot, and Bugaiska (2013) caution against confounding data length variables regarding stimuli and recommend fixed length variables. Therefore, the researcher constructed the audio stimuli in this experiment so that questions and responses are each comprised of twelve syllables, which include all pauses following the

end of an utterance as a syllable (e.g., “I am. [pause] I think I’m going to hit the hay.”). The researcher recorded the verbal exchanges using the audio-capture feature of Sony Vegas Pro 16 media software (www.vegascreativesoftware.com/us/) and then edited each stimulus to a precise five-second duration. The researcher rendered the audio clips as Windows Media Audio (.wma) files for compatibility with the OpenSesame software application (Mathôt, Schreij, & Theeuwes, 2012) the researcher used to construct and conduct the experiment.

Visual stimuli. The experiment used thirty unique images as visual stimuli for this experiment. The researcher cropped each image from an original to create square dimensions that were then rendered to 200 by 200 pixels using GIMP-2.01, an open source image-editing platform (www.gimp.org/docs/). These manipulations created a consistent shape, size, and image quality across the stimuli.

Raw data from a small study conducted by the researcher for a graduate level communication course, which explored the recognizability of English language idioms by adults across varying age ranges, informed the researcher’s selection of images (Vandehey, 2018). In anticipation of recruiting undergraduate students from the PSU Department of Communication, the experiment used images and idiomatic phrases in the experiment that tested well among participants under the age of 30 years. Each image fell in to one of three categories; *figurative image*, *literal image*, and *random image*.

Figurative image. The researcher selected six of the images used in the present experiment for the high recognition rate of the figurative meaning of their respective idioms (Vandehey, 2018). These images depicted actions such as a young girl sleeping for the idiom *hit the hay*, and a criminal suspect making a confession for *spill the beans*.

Literal image. The researcher selected six of the images used in the present experiment for the high recognition rate of the literal meaning of their respective idioms (Vandehey, 2018). These images depicted actions such as a car traveling up the side of a building for “*drives me up the wall,*” and a foot striking a pail for “*kick the bucket.*” The researcher selected six additional images for their strong representation of the constructions used in the non-idiomatic condition (Vandehey, 2018), such as an image of a yellow cab for “I’ll drive a taxi.”

Random image. The researcher selected twelve random images for the experiment based on the absence of any clear overlap in meaning between the images and utterances in the audio stimuli. The experiment assigned each random image to a specific TRE to hold constant the image encountered by every participant in the random image condition for each TRE.

Design of Experiment

The experiment design is an inter-participant congruency test that measures a participant’s response time to stimuli in milliseconds (ms.). The experiment records the time it takes a participant to identify congruence or incongruence between an *audio stimulus* and *visual stimulus*. Audio stimuli consisted of two types, idiomatic and non-idiomatic. The researcher assigned a specific audio stimulus to each TRE so all participants experienced the TREs in the same order. Conditions consisted of two variations on visual stimuli placement (*during audio-stimulus* or *after audio-stimulus*), and image type (*figurative, literal, or random*). Each participant was exposed to twelve total iterations of these TREs. Specifics of the design are detailed below.

Visual stimuli duration. In every condition the image appears on the screen for a duration of precisely 500 ms. (see Bonhage, Mueller, Friederici, & Fiebach, 2015).

Visual stimuli placement. The temporal placement of the image depended on which of the two conditions the experiment assigned to the participant, *during audio-stimulus* or *after audio-stimulus*. Participants assigned to an odd numbered identifier received placements in the opposite order as those with an even numbered identifier (e.g., if participant 101 saw the visual stimulus during the audio stimulus, participant 102 saw the visual stimulus after the same audio stimulus).

During audio-stimulus. In this condition, the image appears at the 4500 ms. mark in the audio-stimulus and remains on the screen through the 5000 ms. mark.

After audio-stimulus. In this condition, the full audio-stimulus plays (5000 ms.), followed by a 1000 ms. pause, and then the image appears on the screen for the duration between the 6000 ms. and 6500 ms. mark.

Procedure. Participants engaged with the experiment through interaction with one of five computer stations staged in a designated laboratory space within the Department of Communication at PSU. Each computer station was preloaded with the OpenSesame 3.2.7 software platform the researcher used to construct the experiment. Four of the five stations used 15.4 inch monitors driving 1920 X 1080 pixel resolution and one station used a 17 inch monitor driving 1920 X 1080 pixel resolution. The researcher adjusted the display resolution on the four smaller monitors to approximate the display size of the visual stimuli on the larger monitor so the difference in size was indistinguishable at a typical viewing distance of 60 cm. Participants listened to audio through Sony model MDR-210 headphones and the researcher adjusted the volume to the

comfort of the individual participant. Participants used built-in QWERTY keyboards and Apple model A1152 computer mice to navigate through the experiment.

The researcher escorted participants into the testing lab and provided brief instructions about basic navigation through the experiment using the mouse and keyboard, staying with each participant long enough to make sure the audio was working correctly and adjusted to an acceptable volume level. The researcher also informed participants they were free to work through the experiment at their own pace as they could control the duration of the pause between each TRE. Before leaving the testing lab, the researcher assured the participants he would be seated outside the door if they needed any assistance.

From this point, the experiment guided the participants through an interactive tutorial with an example TRE to assure they understood the navigation and response characteristics of the experiment before the software started recording their response times. As well, participants were notified that the program was recording their response times and accuracy rates, and that they would be able to see how well they scored at the end of the experiment. These instructions also clearly stated that their accuracy and response scores could in no way negatively impact the amount of course extra-credit they would receive for participation.

The first condition of the experiment regards whether the audio stimuli contain or do not contain an idiomatic phrase. Six TREs used audio stimuli containing an idiomatic phrase and six did not. Rather than present the TREs in a predictable fashion, the researcher randomized the order in which idiomatic and non-idiomatic stimuli appeared. The researcher constructed this sequence and treated all participants to the same

arrangement. The researcher determined this order using a random number generator (google.com), and presented the sequence as; 1) idiomatic, 2) non-idiomatic, 3) non-idiomatic, 4) non-idiomatic, 5) idiomatic, 6) non-idiomatic, 7) idiomatic, 8) non-idiomatic, 9) idiomatic, 10) idiomatic, 11) idiomatic, 12) non-idiomatic (see appendix for specific TRE order and content). This static order was maintained to reduce the number of variables present across the experiment and facilitate more direct comparison within TREs.

Each TRE began with the audio stimulus. Depending on the condition, an image appeared during the last 500 ms. of the audio stimulus, or the image displayed for 500 ms. following the audio stimulus plus a one second pause. The experiment instructed participants to indicate as quickly as they could whether the image was congruent or incongruent with any of the words spoken in the audio-stimulus. Responses from the participants were timed from the instant the image appeared on the screen (see Figure 1).

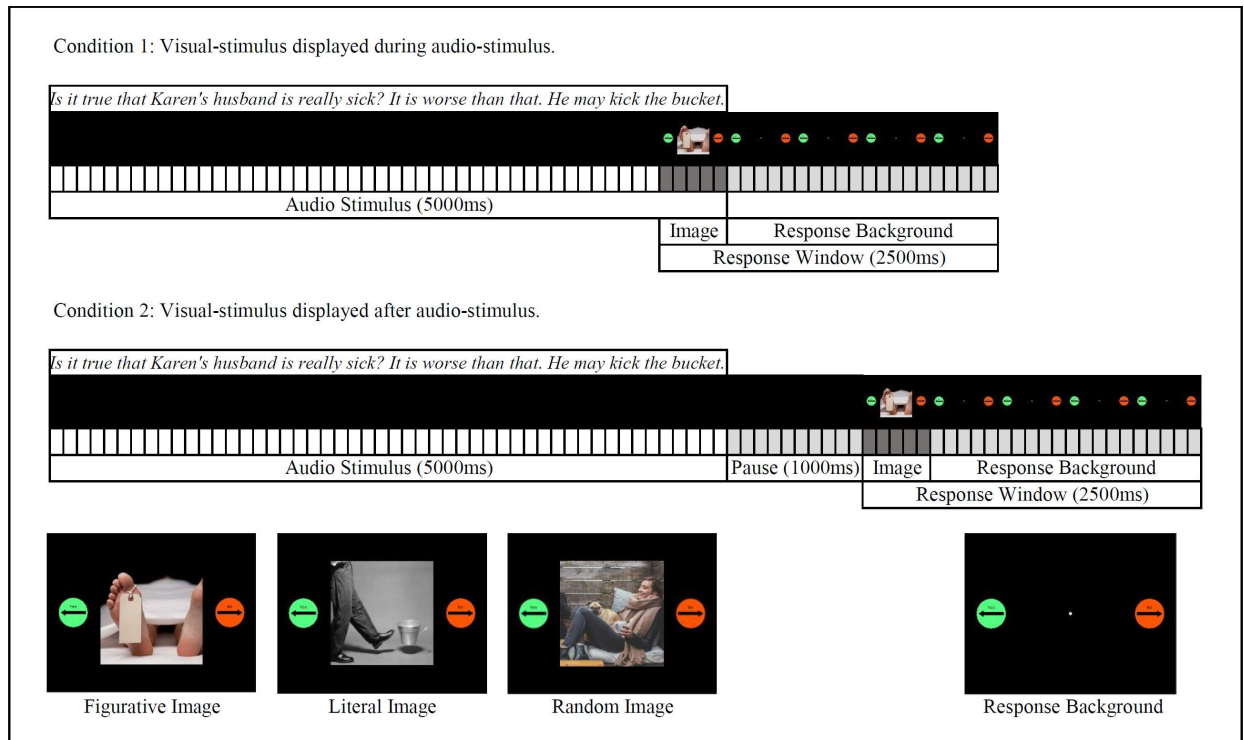
Participants indicated congruency by pressing the left arrow on the computer keyboard, and incongruency by pressing the right arrow. To remind participants of the directionality of the responses, the program displayed a green circle containing the word “Yes” and a left-pointing arrow on the left side of the screen, and a red circle containing the word “No” and a right-pointing arrow on the right side of the screen. These circles appeared on the screen at the same instance as the image and remained on the screen after the image disappeared until the participant registered a response or until the TRE timed out (TREs automatically terminated two seconds after the image vanished from the screen). All imagery appeared over a black background. The visual stimuli were spatially centered on the screen at (0,0) pixels. The green circle was vertically centered and

positioned to the left of center by 384 pixels, (-384,0). The red circle was vertically centered and positioned to the right of center by 384 pixels, (384,0) (see Figure 1).

The second condition activates within each TRE and is the temporal placement of the visual stimulus either *during audio-stimulus* or *after audio-stimulus*. This is a crucial element to the design, as the hypotheses predict that a participant will identify the literal meaning of an idiom before the figurative meaning (see Figure 1).

The third condition also activates within each TRE and is defined by the type of image displayed (figurative depiction, literal depiction, or random image). These distinctions are important to the design as H1 predicts that in the *idiomatic condition* a participant should be able to identify congruence between the audio and visual stimuli faster when a literal image appears in the *during audio-stimulus* condition. H2 predicts that in the *idiomatic condition*, a participant should be able to identify congruence between the audio and visual stimuli faster when a figurative image appears in the *after audio-stimulus* condition. Random images are used here as a control for the purpose of comparing responses between congruent and incongruent stimuli. In the six TREs containing non-idiomatic audio stimuli, there were no images depicting a figurative representation because the utterances in the stimuli had no possible figurative interpretation. In these TREs, participants were treated to either a literal or random image with the random images acting as a control to identify variances in response times across idiomatic and non-idiomatic conditions (see Figure 1).

Figure 1: Diagram of TRE construction.



Once the participants completed the twelve TREs, the program displayed their accuracy rates and average response times. When the participants finished viewing these scores, the program asked them to voluntarily share demographic information. Once the participants completed the demographics section, the program thanked them for their participation and requested that they notify the researcher as they departed. As participants departed, the researcher asked if they had any questions regarding the research and provided answers to those who were curious.

Data Analysis

The researcher designed the experiment to balance participation between the condition of *during audio-stimulus* and *after audio-stimulus*. However, this balance was partially disrupted because 114 of the 876 total responses (13%) were eliminated from the analysis where the participants responded incorrectly to the question of congruency

between the audio and visual stimuli. The researcher eliminated these responses on the assumption that if the participant responded incorrectly it is likely they were unfamiliar with the idiom or did not make the correct assessment of the image. In either case, the response time data provided by incorrect responses had no discernable value, so the researcher removed them. The consequences of this culling of data were some asymmetrical numbers of participants assigned to the conditions of visual stimuli placement and image type in the analysis. This, however, proves inconsequential in comparison to the most serious issue with the data collection process and the corresponding disruption to balanced participation.

During analysis, the researcher learned from an advisor of a crucial flaw in the experiment design that consequently allowed for 909 potential condition assignments through the full course of 12 TREs ($C = (12!/12!(12-6)!)-(6!/2!(6-2)!)$). The design incorporated full-randomization of image type where a pseudo-randomized order (e.g., see Bonhage et al., 2015), like the structure this study uses for phrase type and timing conditions, would have confined possible condition assignments to a total of six. With pseudo-randomization, the initial order of treatments is determined at random, but this order remains fixed thereafter (see example of pseudo-randomization, Table 1). The analyses the researcher performed in efforts to salvage some data revealed that as a result of full-randomization, only two of the 73 participants experienced identical treatments through TRE number five, and then they diverged at TRE number six.

Table 1: Detail of pseudo-randomization applied to the presentation of image type.

TRE Number	Phrase Type	Cond 1			Cond 2			Cond 3			Cond 4			Cond 5			Cond 6				
		Timing	Image Type	Image Type	Image Type	Timing	Image Type	Image Type	Image Type	Timing	Image Type	Image Type	Image Type	Timing	Image Type	Image Type	Image Type	Timing	Image Type	Image Type	Image Type
1	Idiomatic	During	Fig	Lit	Rdm	After	Fig	Lit	Rdm	During	Fig	Lit	Rdm	After	Fig	Lit	Rdm	During	Fig	Lit	Rdm
2	Non-Idiomatic	During	Rdm	Fig	Lit	After	Rdm	Fig	Lit	During	Rdm	Fig	Lit	After	Rdm	Fig	Lit	During	Rdm	Fig	Lit
3	Non-Idiomatic	After	Lit	Rdm	Fig	During	Lit	Rdm	Fig	After	Lit	Rdm	Fig	During	Lit	Rdm	Fig	After	Lit	Rdm	Fig
4	Non-Idiomatic	During	Fig	Lit	Rdm	After	Fig	Lit	Rdm	During	Fig	Lit	Rdm	After	Fig	Lit	Rdm	During	Fig	Lit	Rdm
5	Idiomatic	After	Lit	Rdm	Fig	During	Lit	Rdm	Fig	After	Lit	Rdm	Fig	During	Lit	Rdm	Fig	After	Lit	Rdm	Fig
6	Non-Idiomatic	After	Rdm	Fig	Lit	During	Rdm	Fig	Lit	After	Rdm	Fig	Lit	During	Rdm	Fig	Lit	After	Rdm	Fig	Lit
7	Idiomatic	After	Rdm	Fig	Lit	During	Rdm	Fig	Lit	After	Rdm	Fig	Lit	During	Rdm	Fig	Lit	After	Rdm	Fig	Lit
8	Non-Idiomatic	During	Fig	Lit	Rdm	After	Fig	Lit	Rdm	During	Fig	Lit	Rdm	After	Fig	Lit	Rdm	During	Fig	Lit	Rdm
9	Idiomatic	During	Lit	Rdm	Fig	After	Lit	Rdm	Fig	During	Lit	Rdm	Fig	After	Lit	Rdm	Fig	During	Lit	Rdm	Fig
10	Idiomatic	After	Lit	Rdm	Fig	During	Lit	Rdm	Fig	After	Lit	Rdm	Fig	During	Lit	Rdm	Fig	After	Lit	Rdm	Fig
11	Idiomatic	During	Rdm	Fig	Lit	After	Rdm	Fig	Lit	During	Rdm	Fig	Lit	After	Rdm	Fig	Lit	During	Rdm	Fig	Lit
12	Non-Idiomatic	After	Fig	Lit	Rdm	During	Fig	Lit	Rdm	After	Fig	Lit	Rdm	During	Fig	Lit	Rdm	After	Fig	Lit	Rdm

Areas shaded in gray were pseudo-randomized in the original experiment.

Ultimately, the only salvageable data emerged from TRE number one (TRE1). Analysis demanded the selection of TRE1 because by the end of this short exercise each participant had experienced one of only six possible treatments. By the end of TRE2, each participant had experienced one of 19 possible treatments. As mentioned, by TRE6, each of the 73 participants in the analytical sample had experienced one of 73 possible treatments. As a result, 706 of the remaining 762 responses had to be eliminated, leaving only 56 valid responses in the dataset. The jettison of the bulk of the data strips the study of any power to make claims surrounding the results, however, unreliable as the rejected data were, the overall trend observed across the idiomatic conditions were consistent with the findings isolated to TRE number one.

Results

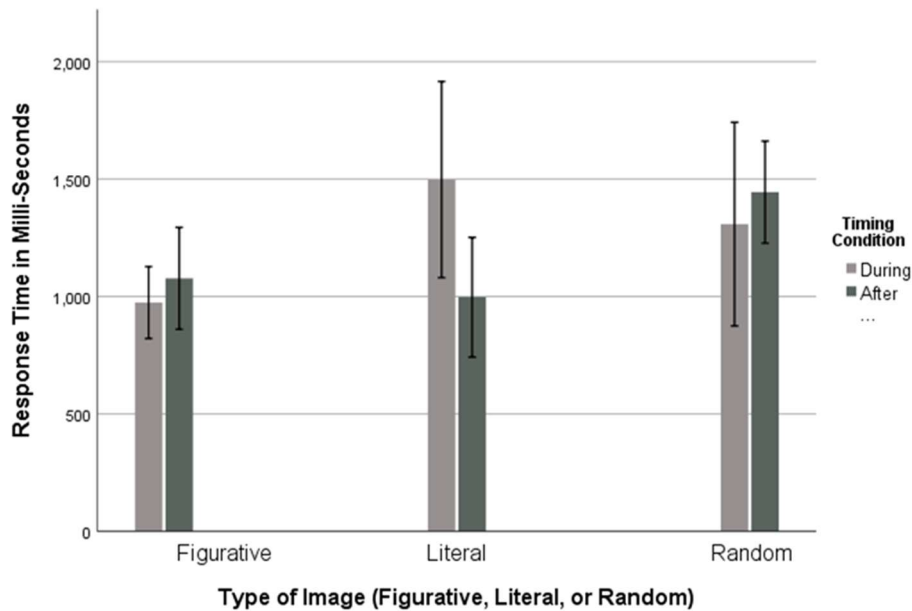
A one-way analysis of variance (ANOVA) indicates that response times recorded for TRE number one constitute a significant main effect for treatment, $F(5, 50) = 4.21$, $p < .05$. Hypothesis one, which predicts that participants will be faster to identify congruence between a literal image and the audio stimulus when the participant views the image during the audio stimulus, was not supported. Participants responded, on average,

501 ms. quicker when they had the opportunity to listen to the entire audio stimulus and take a one second pause before viewing the visual stimulus ($M_{during} = 1498$, $SD = 398$; $M_{after} = 997$, $SD = 332$).

Hypothesis 2, which predicts that participants will be faster to identify congruence between a figurative image and the audio stimulus when the participant views the image after the audio stimulus, was not supported. Participants responded, on average, 103 ms. quicker when viewing the visual stimuli during the audio stimuli ($M_{during} = 974$, $SD = 214$; $M_{after} = 1077$, $SD = 376$).

Random image treatments were included in the experiment as a control to set a baseline between response times registered in the *during* and *after* conditions as random images should demonstrate no advantage in either condition. In the random image treatment, participants identified the incongruence of the image and audio 136 ms. quicker when the image appeared during the audio stimulus rather than after ($M_{during} = 1308$, $SD = 349$; $M_{after} = 1444$, $SD = 342$).

Figure 2: Response times for determining congruence/incongruence between visual and audio stimuli in the *during* and *after* timing-conditions.



The 95% Confidence Intervals (CI) the researcher calculated for each of the six conditions in Figure 2 indicate overlaps in the recorded response times that may render the differences inconsequential. This overlap renders the results inconclusive even in the case of the *literal* image treatments, where it took participants 50.25% longer to identify congruence when the image was viewed during the audio stimulus. Results in the *figurative* and *random* treatments were closer between the *during* and *after* conditions, with participants identifying *figurative* congruence only 9.57% faster and *random* incongruence only 9.42% faster when viewing the image *during* the audio stimulus as opposed to after. This result indicates there is negligible to no advantage in having an extra second and a half to process the audio stimuli prior to viewing the figurative and random visual stimuli.

Table 2: Response times grouped by image type and by timing condition.

Image Placement and Type	N	Response Time (ms.)	Std. Deviation	<i>d.f.</i>	Sig. (2 tailed)	<i>M</i> Difference	95% CI Lower	95% CI Upper
Figurative Image - During	10	974	214	23	$p < .001$	1034	901	1168
Figurative Image - After	14	1077	376					
Literal Image - During	6	1498	398	14	$p < .001$	1197	960	1434
Literal Image - After	9	997	331					
Random Image - During	5	1308	349	16	$p < .001$	1404	1230	1578
Random Image - After	12	1444	342					
Figurative Image - During	10	974	214	20	$p < .001$	1203	1033	1373
Literal Image - During	6	1498	398					
Random Image - During	5	1308	349					
Figurative Image - After	14	1077	376	34	$p < .001$	1182	1047	1318
Literal Image - After	9	997	331					
Random Image - After	12	1444	342					

Post-hoc analysis revealed there were no obvious extenuating circumstances leading to the lengthy response times observed in the condition including a *literal* image viewed *during* the audio stimulus. All participants in this category indicated English as their first language, all had completed at least some college education, and the group demonstrated a 90.2% accuracy rate compared to the aggregate average of 87.0%. However, this group did register overall response times that were 13.8% longer than the average. This sluggishness at the keys may contribute to the 50.25% longer response times between the *during* and *after* conditions, but it does not contribute enough to explain away the gulf between these two data sets.

Another observation that makes data from TRE1 troublesome is that only 56 of the 73 participants (76.7%) were able to accurately identify congruence/incongruence in this first TRE, and the average response time for TRE1 was 1190 ms. (the longest of any TRE) compared to an overall average of 994 ms.. Since the experiment got off to a jittery start, this confounding variable should also be considered when assessing the validity of the data.

Discussion

The study opens this discussion by reiterating that serious problems with the design of the experiment, coupled with low participation numbers, resulted in data that are not strong enough to substantiate claims of any kind. However, if the present and future researchers are to learn and build from the experience of this study, they should consider the patterns in the response times that contradict the predictions in the hypotheses forwarded. These results raise new questions that future studies must address if research in this area is to profit from the errors documented in this study and use these missteps to inform the design of improved research efforts.

The most devastating of these errors was the randomization flaw in the experiment's design. A sounder experiment would be designed with a total of six consistent conditions that carry through the 12 each TREs. This could be accomplished using pseudo-randomization applied to the order in which the image types are viewed (see example, Table 1). The present design fully randomizes which type of image will be viewed in any given TRE, which was the cause of the high number of conditions generated by OpenSesame and also the cause of unbalanced assignment to condition. In the case of TRE1, OpenSesame assigned as few as five participants and as many as 14

participants to a single condition. Had the randomization been level, each condition would have had between nine and ten participants. As mentioned, culling of incorrect responses contributed to this issue. Comparisons were often asymmetrical, weakening the results further.

The chief data of interest in this study are response times and how they differ when identifying congruence between the literal and figurative meaning of the words spoken in the audio stimuli related to when the image is seen, *during* or *after* the audio stimuli. Though this study cannot suggest the data were reliable enough to make claims, certain patterns were consistent enough to suggest the two hypotheses the researcher posited are incorrect. Where the CI/LSA model suggests that it should take longer for a participant to identify congruence when viewing a figurative image during the audio stimulus, the data gathered in this study suggest there is no appreciable difference in the amount of time it takes a participant to identify congruence whether the image is viewed during or after the audio stimulus. Where CI/LSA posits that the participant should still be in the process of constructing the intended meaning of the phrase, the data suggest by the time the image is presented to the participant in the *during* condition (in the last 500 ms. of the audio stimulus) the participant already understands the intended (figurative) meaning of the phrase.

In addition, Caillies and Butcher (2007) and Holsinger (2013) argue that to embrace the figurative meaning of an idiom, a participant must first reject the literal meaning of the phrase. This idea is in concert with the CI/LSA model, which suggests that an idiom like *drives me up the wall* must be decomposed, its parts (drives | me | up | the | wall) compared to meanings in their proximal semantic neighborhoods, recomposed,

and compared to the context around the phrase to determine if the meaning at which the participant arrives is the correct one. Since the figurative meaning of an idiom cannot be determined by examination of its constituents, Caillies and Butcher and Holsinger are reasonable to assert that the literal meaning must be set aside before one can arrive at the figurative meaning if the CI/LSA model does apply to the processing of idiomatic language. This idea was the foundation of the hypotheses for this study, but these hypotheses were not supported by the results. The results are more consistent with Caillies's later work with Declercq where they argue, "Indeed, it appears that listeners or readers can understand the figurative interpretation of metaphors, idioms, irony... without having to first analyze and reject their literal meanings" (Caillies & Declercq, 2011, p. 205).

What the researcher observed suggests it is not the literal meaning that must be set aside to embrace the figurative, but rather that in most cases the figurative must be set aside to embrace the literal. When this theory is applied to the data, the results begin to make sense. If at the 4500 ms. mark (500 ms. before the audio stimulus ends) the participant has already accepted the figurative meaning of the phrase, the participant would then have to push that meaning aside to consider the literal meaning. If the participant views a picture congruent with the figurative meaning of the idiom, the participant, who is already visually simulating that concept, will quickly identify the congruence of an image that depicts the figurative concept. In contrast, if the participant is envisioning the figurative concept, and then views a depiction of the literal concept, it will take time for the participant to reject the figurative concept and embrace the literal. This will theoretically slow the response time as it requires extra steps to reconcile the

image with the words spoken rather than what the participant determines to be the intended meaning of the phrase. However, though this bit of grounded-theory may explain why it takes longer to identify literal congruence, it does not explain why the participant would have already accepted the figurative meaning of the phrase at a point where she or he has not yet heard the phrase in its entirety.

To understand this piece, the researcher looked to the concept of *prediction* (see Bonhage et al., 2015; Christiansen & Chater, 2016; Corps, Pickering, & Gambi, 2018; Ito, Corley, Pickering, Martin, & Nieuwland, 2016). Some research suggests that limitations in human working memory, coupled with the rapid pace of verbal exchange, require the receiver of a message to make predictions regarding what a speaker is going to say. This concept is commonly studied in turn-taking, as the receiver of a message perpetually tries to simultaneously predict the meaning of the speaker and the end of the speaker's turn (Clayman, 2012; Corps et al., 2018). Prediction is such an integral part of communication that people perform it at a subconscious level (Clayman, 2012; Huettig, 2015). All the language the listener is processing must pass through her or his working memory, which is rarely able to accommodate more than seven unique pieces of information at any given time (Caillies & LaSourn-Bissaoui, 2012; Huettig, 2015). Because working memory is so limited people scarcely retain the auditory component of communication for longer than 100 ms., thus humans quickly process strings of words into abstracted *chunks* which are combined and abstracted into higher and higher concepts (Christiansen & Chater, 2016; see also, "conceptual layer," Holsinger, 2013). These chunks commonly extend into predictions, which are formed as the message receiver reads contextual, syntactic, semantic, and phonological cues (Bergen, 2012).

Predictions are seldom so precise as to conjure the exact words to come, but rather manifest more commonly as an abstract conclusion that fits nicely with the overall concept that is forming (Christiansen & Chater, 2016; Holsinger, 2013; Ito et al., 2016). When the actual message that follows does not fit with the prediction, the result is confusion and the receiver must clamber to reevaluate the meaning of the message. Scholars refer to this phenomena as “being led down the garden path,” (Bergen, 2012, p. 103), where a message sender’s initial words lead the receiver to predict what seems like an obvious conclusion, but the full message emerges with a meaning that does not fit the prediction. An example of this is in Groucho Marx’s quip, “Time flies like an arrow. Fruit flies like a banana.” (Bergen, 2012, p. 103). Marx’s clever turn of phrase leads people to assume the term *flies* in the second sentence is used as a verb, as in the first sentence, rather than as a noun. The result for most people is an abstract visualization of some form of fruit hurtling through space. If the sentence ended in *baseball* instead of *banana*, the process of determining meaning is simple. An apple, orange, or peach are similar in shape and size to a baseball, and it stands to reason any of these three would fly like a baseball. People also associate baseballs with flying through the air, so the simile is valid.

Since the concept of banana breaks with any typical prediction about flying fruit, the discrepancy creates a need to reconcile one’s visualization with the words of the message. Though bananas can fly if propelled, this is not the primary function of a banana. In addition, a banana is not like a fruit, it is a fruit, so the simile is invalid. This sentence does not make sense when *flies* is used as a verb, so the receiver of the message must push aside this idea and embrace *flies* as a noun. This example demonstrates how

people make predictions without conscious effort, and how disorienting it can be when the images conjured in one's imagination cannot be reconciled with the actual message.

In this study, the data suggest participants had predicted the meaning of the phrase before they heard it in its entirety. They may have already abstracted the meaning of the verbal exchange and predicted the abstract meaning of the missing piece based on context, syntactic, semantic, and perhaps even phonologic cues. Jackendoff (1995) and Glucksberg (2001) may argue the participants could have completed the sentences word-for-word so long as they were familiar with the idiom as a fixed expression. If these researchers are correct, the implications would favor the LRH model over the CI/LSA model. Jackendoff (1995) posits that English speakers store as many multi-word sequences (approximately 80,000) in the lexicon as they do single words. These multi-word sequences provide short-cuts that aid communication not only in message receipt, but message construction. People attempt to describe abstract concepts by using a word or string of words. Common phrases (such as idioms, compounds, and collocations) act as concise short-cuts that can signify concepts that may otherwise be difficult or time consuming to describe (Glucksberg, 2001; Holsinger, 2013; Jackendoff, 1995).

People can access these stored phrases with great speed, and the more common the phrase the more quickly the person can access it. Jackendoff (1995) points to the classic television gameshow, *Wheel of Fortune* for evidence of this phenomenon. In the program, contestants solve word puzzles to win prizes. The puzzles appear on a large board, and at first are only a series of blanks and spaces. The blanks begin to populate as contestants correctly guess the letters that fill in the blanks, and the first to solve the puzzle wins the round. The solution to the puzzle is often a proper name or a book,

movie, or song title, but the most common solutions are collocations and the second most common are idioms. The idiom is such a common puzzle choice for the producers because they are easy to guess with a minimal number of cues.

Since it was the researcher's aim from the start to source idioms with the highest probability of recognition, the researcher would not be surprised if the idioms selected lend themselves to prediction well. The data, though not up to academic rigor, have enough to say about the thought processes involved in the understanding of idioms to open the door to new inquiry. If the mental imagery does not match the visual stimuli in comparison, research suggests this results in costs to mental resources and time (Bergen, 2012; Bonhage et al., 2015). Therefore, if participants form mental simulations based on their predictions, this should result in quicker responses when the prediction is congruent with the visual stimulus. When the stimulus and simulation are incongruent, the participant is forced to consider a different meaning in much the same way 'garden path' sentences force the reader or listener to consider alternative meanings.

Prediction can account for why there is no advantage in having an extra second and a half to think about the figurative meaning of the phrase. Participants may have predicted the figurative meaning by the 4500 ms. mark, which would also make some sense of why it took them 524 ms. longer to identify the congruence of the literal image. If the participant already had an abstract vision of the concept, she or he would be able to align the image in her or his mind with the image on the screen without needing to hear all the words in the sentence. If the abstract image in the mind of the participant does not align with the image on the screen, the participant would first need to hear all the words in the sentence, which would take an additional 500 ms., to determine congruence or

incongruence. This may be supported by observations in the *after* conditions where participants were able to identify literal congruence 501 ms. faster after hearing the entire sentence than their counterparts in the *during* condition who had yet to hear all the words when the image appeared.

Response times in the random image conditions also align with the idea of prediction. If prediction is moderating, one would expect that the random image would be analyzed for congruence based only on the figurative meaning of the idiom. There are no characteristics associated with the random image conditions that prompt the participant to consider other meanings for the idiomatic phrase. Thus, examination of the image against the figurative meaning of the idiomatic phrase can begin as soon as the participant predicts and conceptualizes the figurative meaning. As in the figurative image condition, having extra time to process the audio stimulus should not provide a significant response-speed advantage to participants, which is consistent with the data salvaged from the study.

Most of what this discussion posits could be spurious, and even if the data were sound, the researcher did not design the experiment to lend internal validity to substantiate any of this post-hoc speculation. The researcher pursued these lines of inquiry because he believes the problems with the study extend beyond the randomization error. The researcher has no delusions that if he were to correct the randomization error and re-administer data collection that the data would suddenly support the hypotheses. If the researcher were to pursue this research further, he would alter and augment his approach significantly based on what he learned.

To begin, given the speed and ease with which the study's participants each completed the experiment, the researcher would expand the number of TREs and include idiomatic phrases using the figurative meaning (as in this study), idiomatic phrases using the literal meaning (which this study neglects), and non-idiomatic phrases (as in this study). By measuring only reactions to stimulus that used the idiomatic phrase in its figurative sense, the study gathered no data that may have been useful in examining prediction as a moderating factor.

Had the study included idiomatic constructions in their literal sense (e.g., "Did you hear John drove into something in the field?", "Yes, Sarah and Ben told me that he *hit the hay*.") Using *hit the hay* in the literal fashion makes this multi-word sequence more difficult to predict. Though there are some contextual cues, such as mention of John driving into something, and that he was in a field when it happened, the phrase *hit the hay* is not typically used in its literal sense. It conforms to syntactic and semantic rules, and makes perfect sense in the context, but it does not constitute a familiar fixed phrase, and thus it is unlikely the literal meaning is stored in the lexicon like the figurative (Holsinger, 2013).

Collecting data on reactions to literal interpretations of idioms would allow for analysis that compares the same multi-word sequences when they take on two different meanings. This data could inform inquiry about how prediction is working, and if prediction is demonstrating the same properties in both conditions. To accomplish an in-depth comparison, it would be ideal to include treatments with audio stimuli that present only the idioms without context (e.g., simply, "hit the hay"). In these cases, the participants would have no cues to help them predict how the phrase ends. Holsinger

(2013) suggests that since the figurative meaning of idioms are stored in the lexicon, the figurative meaning of the idiom is easier to access than the literal. If this is correct, participants should be quicker to identify congruence in figurative images than literal images, but the question remains to what extent quicker reaction times can be attributed to prediction. The answer is likely the advantage recorded in the non-contextual condition subtracted from the advantage recorded in the contextual condition.

Adding more conditions to an experiment already possessing six possible conditions could make administration untenable and would likely leave the study with insufficient power to make claims. In addition, the preparation work for the limited size of the experiment conducted in this study was extensive. Adding TRE's and additional conditions would call for even more preparation. The researcher believes that to mitigate wild variations from one TRE to the next, it is imperative to test audio and visual stimuli extensively through pilot-testing to make sure the clarity and consistency of the utterances are comparable from one recording to the next. This would include elements to assure certain utterances are not more predictable than others. A key factor in this determination might be Glucksberg's (2001) concept of idiom keywords, and whether they appear early or late in the idiom.

In the present study, the idiom used in TRE1, *hit the hay*, has a late keyword, "hay." In this example, "hit" is not the keyword because at the point the participant hears the word "hit" the possible endings are still open to *the hay*, *the sack*, and *the rack*. If context is ignored, and a researcher asks a participant to complete the idiom "hit the _____," responses could expand to include *nail on the head*, and *ball out of the park*. Glucksberg (2001) argues that people consider viable options to end the utterance, and

that these options are narrowed as more words are revealed. Glucksberg posits that the full meaning of the idiom is not activated until the keyword is revealed. If the keyword is revealed early, the full idiom is predicted early. Thus, a reasonable selection criterion to add to the existing criteria is consistent keyword placement. Inconsistent keyword placement could result in shifting prediction away from a moderating effect and toward a mediating effect. Future researchers would need to define and control the role of prediction well to have confidence in the data gathered.

The composition of the idiom is a concern added by Holsinger (2013) who argues that there is strong evidence that idioms are processed structurally at some level, so the idioms used in an experiment such as the one attempted for this study should be structurally similar for continuity sake. Under such analysis, four of the idioms used in this study (*kick the bucket*, *hit the hay*, *spill the beans*, *throw in the towel*) are structurally consistent, each with a clear action and patient presented in the same syntax. However, two of the idioms (*drives me up the wall*, *go cold turkey*) stray from this structure and add to the confounding variability of the experiment. A future study would need to consider consistent structure as an important criterion for idiom selection.

Perhaps of greater importance than the words comprising the idiomatic phrase are the abstracted concepts generated as viable conclusions to the utterances within their surrounding contexts. Bergen (2012) suggests that humans use placeholders for conceptualizations that are a vague, archetypical representation of the concept they signify. This may be demonstrated in the difference between thinking about a *movie star* and thinking about Tom Hanks. Mr. Hanks has a face, voice, and personality that are clear, well known, and easy to visualize for most moviegoers. The concept of a *movie*

star is more abstract and difficult to associate with specific attributes and features. A movie star, therefore, is more likely to be visualized with vague characteristics such as acting talent, glamour, fame, and wealth. Though *movie star* is a broader abstraction than *Tom Hanks*, one would not be surprised in the context of a discussion concerning a movie star to learn that a conversation partner was talking about Tom Hanks. In contrast, one would be surprised to learn a conversation partner was talking about Yo Yo Ma, who is a famous virtuoso cellist but not a movie star.

Similarly, in TRE1, the context may have driven visualization of a vague abstraction of *tired* and the participants were not surprised to see an image of the sleeping girl. The words likely had little influence over their reaction time, as *hit the hay* has a late keyword. At the point the image was presented to the participant, the audio recording was halfway through uttering the word “hit.” From this point, the participant likely would have accepted the endings *the hay*, *the sack*, and *the rack* equally. This may indicate that reaction times were influenced most strongly by context, and the words themselves may have been largely irrelevant in the space of microseconds. The only way a study could make a claim about the influence of context, and the way it drives prediction, is to control for it. It is therefore crucial to control for context in future experiments by adding TREs that include only the idiomatic phrase without context.

Conclusion

The experiment the researcher conducted for this study culminated in no significant answers and a host of new questions. The answers the researcher sought turned out to be tied to the wrong questions, so there is value in the discoveries even if

the work here provides no support to either side of the debate over whether idioms are lexical units or multi-word constructions. Were others to continue this line of research they would be remiss not to pursue a broader understanding of prediction. Prediction likely transcends the words people hear and operates largely in the abstracted concepts listeners form from the context of the conversation. People appear to form predictions not of words, but of concepts that fit the abstracted context. As the listener hears more, the concept is refined, and the list of congruent conclusions is winnowed. If this conclusion fits the description of a fixed phrase, such as an idiom, the listener may be able to predict the exact words of the conclusion before hearing them. The likelihood of such a precise prediction seems to be predicated on both a familiarity with the idiom and the placement of the idiom's keyword.

The foundation of this study presupposes people hear the phrase and then determine whether the idiom is used figuratively or literally based on the context. The limited value of the data gathered here suggests the meaning is determined before the participant hears the phrase. If this is true, then abstracted concepts are driving the latent semantic analysis, and people are comparing neighboring concepts and not lexical units. As in Ito et al. (2016), this indicates the receiver may consider meaning in the absence of form, but never form in the absence of meaning. As such, the concept of being tired can easily be associated with needing rest and needing rest can easily be associated with going to bed. If the participants in this study predicted the verbal exchange in the audio stimulus was leading to a conclusion consistent with going to bed, there are no strong connections between this concept and the word *hit*, the word *the*, or the word *hay*.

However, if the participant is familiar with the idiom *hit the hay*, the connection is strong indeed.

In this way idioms do resemble lexical units, but idioms also depart from lexical units in other aspects. Idioms exist on a continuum of decomposability that spans opaque, transparent and quasi-metaphoric, which is not an attribute of a lexical unit. As such, Glucksberg (2001) has a strong argument behind classifying idioms as fixed phrases, a concept that shares overlap with lexical units but also recognizes what makes each unique. Similarly, many primates share an extraordinary number of attributes with humans, but there is no case to be made that these primates are human. Expanding this analogy, some primates, such as chimpanzees, share more attributes with humans than others, such as lemurs. Idioms also demonstrate this complexity, with some resembling lexical units more than their contemporaries.

Research in this area will benefit from a greater understanding of the differentiation between diverse classifications of idiom and a deeper understanding of how people process and store language at a more general level. As in many debates that span decades, the meaningful truth regarding idiomatic language likely resides somewhere between the opinion that idioms are lexical units and the opinion that idioms are multi-word constructs. The answers are likely as complex as the questions posed in this discussion. Only time and a great deal of ambitious research will tell. Hopefully the debate will continue, and strides will be made toward an understanding of linguistics that benefits all who speak English, and especially those who endeavor to learn English as a second language.

References

- Beck, D.M. (2010). The appeal of the brain in the popular press. *Perspectives on Psychological Science*, 5(6), 762-766. DOI: 10.1177/1745691610388779
- Bergen, B.K. (2012). *Louder than words: The new science of how the mind makes meaning*. New York: Basic Books.
- Bonhage, C.E., Mueller, J.L., Friederici, A.D., & Fiebach, C.J. (2015). Combined eye tracking and fMRI reveals neural basis of linguistic predictions during sentence comprehension. *Cortex*, 68(2015), 33-47. DOI: 10.1016/j.cortex.2015.04.011
- Bonin, P., Meot, A., & Bugaiska, A. (2013) Norms and comprehension times for 305 French idiomatic expressions. *Behavioral Research*, 45. 1259-1271.
doi:10.3758/s13428-013-0331-4
- Boulenger, V., Hauk, O., & Pulvermuller, F. (2009). Grasping ideas with the motor system: Semantic somatotopy in idiom comprehension. *Cerebral Cortex*, 19, 1905-1914. DOI: 10.1093/cercor/bhn217
- Cacciari, C., & Gluksberg, S. (1991). Understanding idiomatic expressions: The contribution of word meanings. In Simpson G.B., (ed.), *Understanding Word and Sentence*. North Holland: Elsevier Science Publishers B.V., 217-240.
- Caillies, S., & Butcher, K. (2007). Processing of idiomatic expressions: Evidence for a new hybrid view. *Metaphor and Symbol*, 22(1), 79-108.
- Cailles, S., & Declercq, C. (2011). Kill the song—steal the show: What does distinguish predicative metaphors from decomposable idioms? *J Psycholinguist Res*, 40, 205-223. DOI 10.1007/s10936-010-9165-8

- Caillies, S., & Le Sourn-Bissaoui, S. (2012). Nondecomposable idiom understanding in children: Recursive theory of mind and working memory. *Canadian Journal of Experimental Psychology*, Advance Online Publication, 1-8. DOI: 10.1037/a0028606
- Cain, K., Towse, A.S., & Knight, R.S. (2009). The development of idiom comprehension: An investigation of semantic and contextual processing skills. *Journal of Experimental Child Psychology*, 102, 280-298. DOI: 10.1016/j.jecp.2008.08.001
- Cheatham, G.A., & Ro, Y.E. (2011). Communication between early educators and parents who speak English as a second language a semantic and pragmatic perspective. *Early Childhood Education Journal*, 39, 249-256. DOI: 10.1007/s10643-011-0467-8
- Christiansen, M.H., & Arnon, I. (2017) More than words: The role of multiword sequences in language learning and use. *Topics in Cognitive Science*, 9, 542-551. DOI: 10.1111/tops.12274
- Christiansen, M.H., & Chater, N. (2016). The now-or-never bottleneck: A Fundamental constraint on language. *Behavioral and Brain Sciences*, 39. 1-72. DOI: 10.1017/S0140525X1500031X
- Clayman, S. (2012). Turn-constructural units and the transition-relevance place. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis*. Wiley-Blackwell.

- Corps, R.E., Pickering, M.J., & Gambi, C. (2019). Predicting turn-ends in discourse context. *Language, Cognition, and Neuroscience*, 34(5), 615-627. DOI 10.1080/23273798.2018.1552008
- Culicover, P.W., Jackendoff, R., & Audring, J. (2017). Multiword constructions in the grammar. *Topics in Cognitive Science*, 9, 552-568. DOI: 10.1111/tops.12255
- Ecklund, A., Nichols, T.E., & Knutsson, H. (2016). Cluster failure: Inflated false positives for fMRI. *Proceedings of the National Academy of Sciences*, 113(33). 1-6. DOI: 10.1073/pnas.1602413113
- Glucksberg, S. (2001). Idioms from metaphors to “just long words”? Retrieved from <https://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195111095.001.0001/acprof-9780195111095-chapter-5>. DOI: 10.1093/acprof:oso/9780195111095.003.0005
- Hillert, D.G., & Buracas, G.T. (2009). The neural substrates of spoken idiom comprehension. *Language and Cognitive Processes*, 24(9), 1370-1391. DOI: 10.1080/01690960903057006
- Holsinger, E. (2013). Representing idioms: Syntactic and contextual effects on idiom processing. *Language and Speech*, 56(3), 373-394. DOI: 10.1177/0023830913484899
- Hsieh, J.S.C., Wu, W.V., & Merek, M.W. (2017). Using the flipped classroom to enhance EFL learning. *Computer Assisted Language Learning*, 30(1-2), 1-21. DOI: 10.1080/09588221.2015.1111910
- Huetig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, 1626, 118-135. DOI: 10.1016/j.brainres.2015.02.014

- Hung, P., & Nippold, M.A. (2014). Idiom understanding in adulthood: Examining age-related differences. *Clinical Linguistics & Phonetics*, 28(3), 208-221. DOI: 10.3109/02699206.2013.850117
- Ito, A., Corley, M., Pickering, M.J., Martin, A.E., & Nieuwland, M.S. (2016). Predicting form and meaning: Evidence from brain potentials. *Journal of Memory and Language*, 86, 157-171. DOI: 10.1016/j.jml.2015.10.007
- Ito, A., Pickering, M.J., & Corley, M. (2018). Investigating the time-course of phonological prediction in native and non-native speakers of English: A visual world eye-tracking study. *Journal of Memory and Language*, 98, 1-11. DOI: 10.1016/j.jml.2017.09.002
- Jackendoff, R. (1995). The boundaries of the lexicon. In M. Everaert, E.J. van der Linden, A. Schenk, & R. Schreuder (Eds.), *Idioms: Structural and psychological perspectives* (pp. 133-165). New York: Lawrence Erlbaum Associates, Inc.
- Kim, C. (2016). L2 learner's recognition of unfamiliar idioms composed of familiar words. *Language Awareness*, 25(1-2), 89-109. DOI: 10.1080/09658406.2015.1122025
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction integration model. *Psychological Review*, 95, 163-182.
- Kintsch, W., & Bowles, A.R. (2002). Metaphor comprehension: What makes a metaphor difficult to understand? *Metaphor and Symbol*, 17, 249-262.
- Maryns, K. (2017). The use of English as ad hoc institutional standard in Belgian asylum interview. *Applied Linguistics*, 38(5), 737-758. DOI: 10.1093/applin/amv061

- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314-324. doi:10.3758/s13428-011-0168-7
- Muhammad, U.A., Yap, N.T., Chan, M.Y. & Wong, B.E. (2016). Identification of Nigerian English idioms: A methodological perspective. *Southern African Linguistics and Applied Language Studies*, *34*(4), 279-288. DOI: 10.2989/16073614.2016.1258999
- Nunberg, G. (1978). *The pragmatics of reference*. Berkeley, CA: Indiana University Linguistics Club.
- Pickering, M.J., & Garrod, S. (2006). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, *11*(3), 105-110. DOI: 10.1016/j.tics.2006.12.002
- Pitzl, M.L. (2016). World Englishes and creative idioms in English as lingua franca. *World Englishes*, *2016*, 293-309. DOI: 10.1111/weng.12196
- Ritchie, L.D. (2004). Lost in “Conceptual Space”: Metaphors of conceptual integration. *Metaphor and Symbol*, *19*(1), 31-50. DOI: 10.1207/S15327868MS1901 2
- Ritchie, L.D., & Dyehouse, V. (2008). Fine as frog’s hair: Three models for the development of meaning in figurative language. *Metaphor and Symbol*, *23*, 85-107.
- Romerso-Lauro, L., Tettamanti, M., Cappa, S.F., & Papagno, C. (2008). Idiom comprehension: A prefrontal task? *Cerebral Cortex*, *18*, 162-170. DOI: 10.1093/cercor/bhm042

- Rommers, J., Dijkstra, T., & Bastiaansen, M. (2013). Context-dependent semantic processing in the human brain: Evidence from idiom comprehension. *Journal of Cognitive Neuroscience*, 25(5), 762-776. DOI: 10.1162/jocn_a_00337
- Saberian, N., & Fotovatnia, Z. (2011). Idiom Taxonomies and Idiom Comprehension: Implications for English Teachers. *Theory and Practice in Language Studies*, 1(9), 1231-1235. DOI: 10.4304/tpls.1.9.1231-1235
- Siefring, J. (Ed.). (2005). *Oxford Dictionary of Idioms* (2nd ed.). New York, NY: Oxford University Press.
- Tzou, Y.Z., Vaid, J., & Chen, H.C. (2017). Does formal training in translation/interpreting affect translation strategy? Evidence from idiom translation. *Bilingualism: Language and Cognition*, 20(3), 632-641. DOI: 10.1017/S1366728915000929
- Vandehey, D. (2018). [Recognizability of idiomatic phrases across age groups]. Unpublished raw data.
- Vulchanova, M., Vulchanova, V., & Stankova, M. (2011). Idiom comprehension in the first language: A development study. *Vigo International Journal of Applied Linguistics*, 8, 207-234.
- Yoshikawa, H. (2008). International intelligibility in world Englishes: Focusing on idiomatic expressions. *International Communication Studies*, XVII (4), 219-226.

Appendix A:

Experiment Audio-Clip Script and Images with Response Data

1.

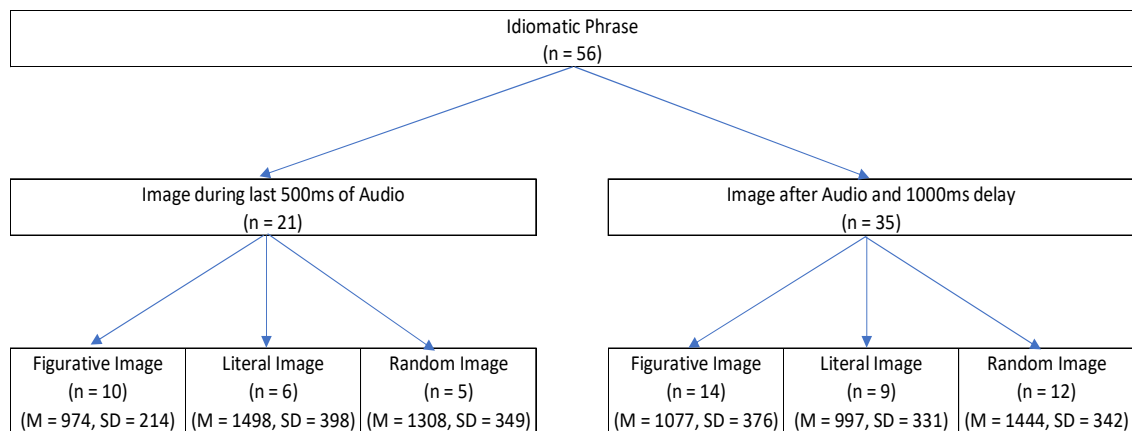
Q: Stacie, are you sleepy after your big game?

A: I am. I think I'm going to *hit the hay*.

Figurative Image

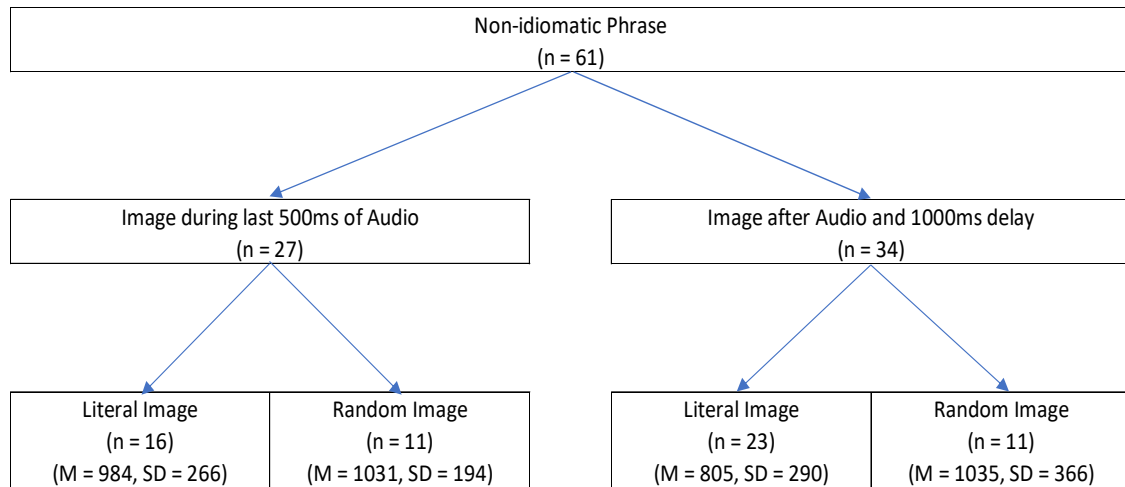
Literal Image

Random Image



2.

Q: How long until you know whether you get the job?

A: It could be a month. Until then I'll drive a taxi.Literal ImageRandom Image

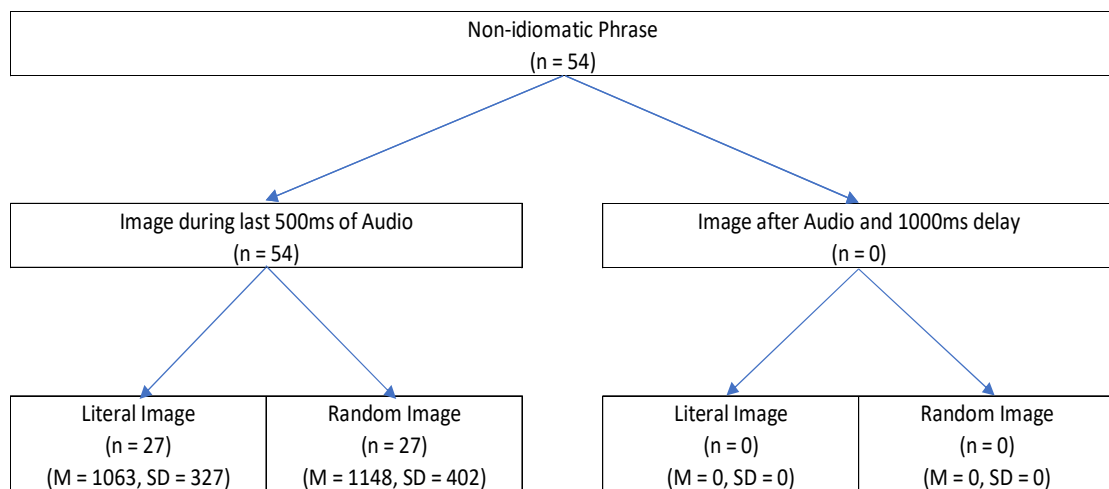
3.

Q: When does the MAX Blue Line arrive at this platform?

A: It won't stop here. This stop is for the green train.

Literal Image

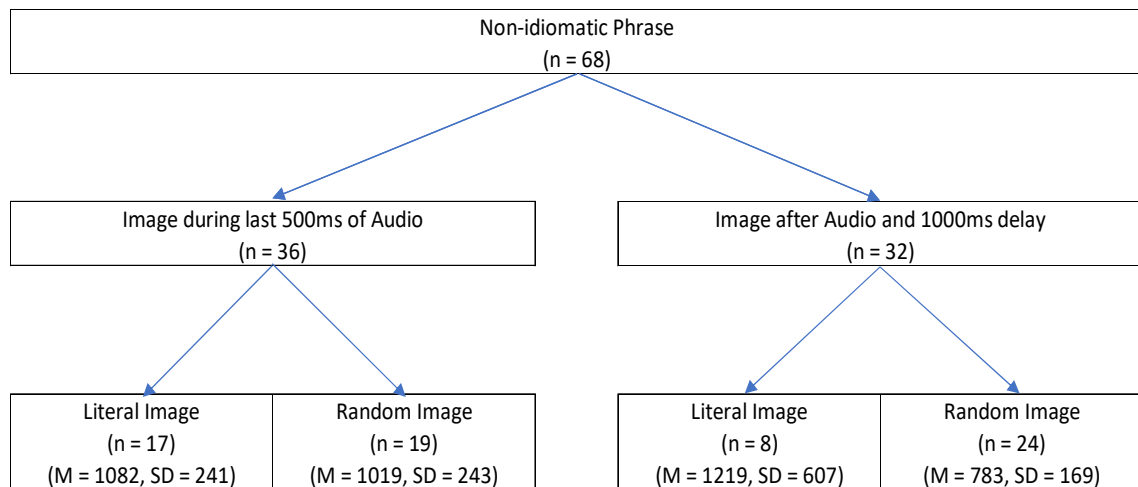
Random Image



Eight participants were exposed to the “Image after Audio and 1000ms delay” condition, but all eight responded to the prompt incorrectly, leaving zero viable samples for this condition.

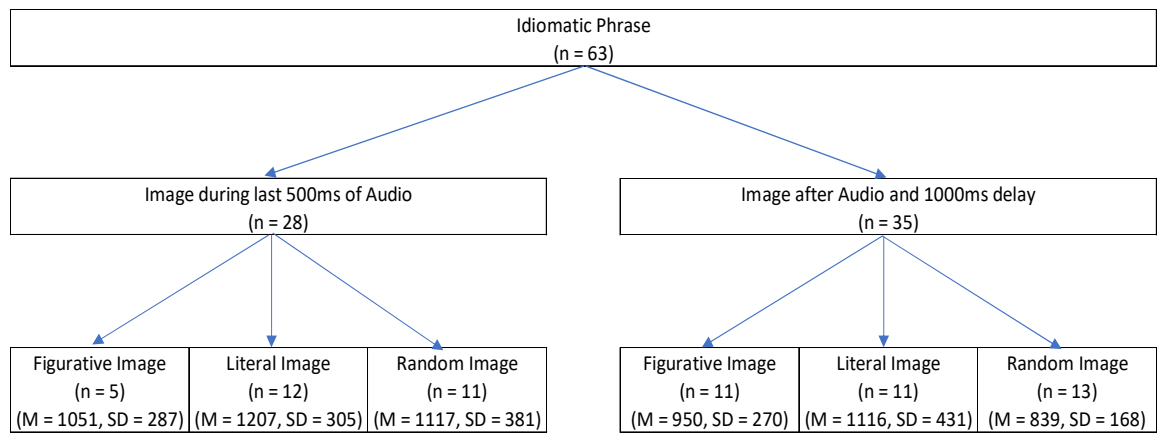
4.

Q: What kind of car did you say your dad drives again?

A: A Subaru, and he hates the driver's seat.Literal ImageRandom Image

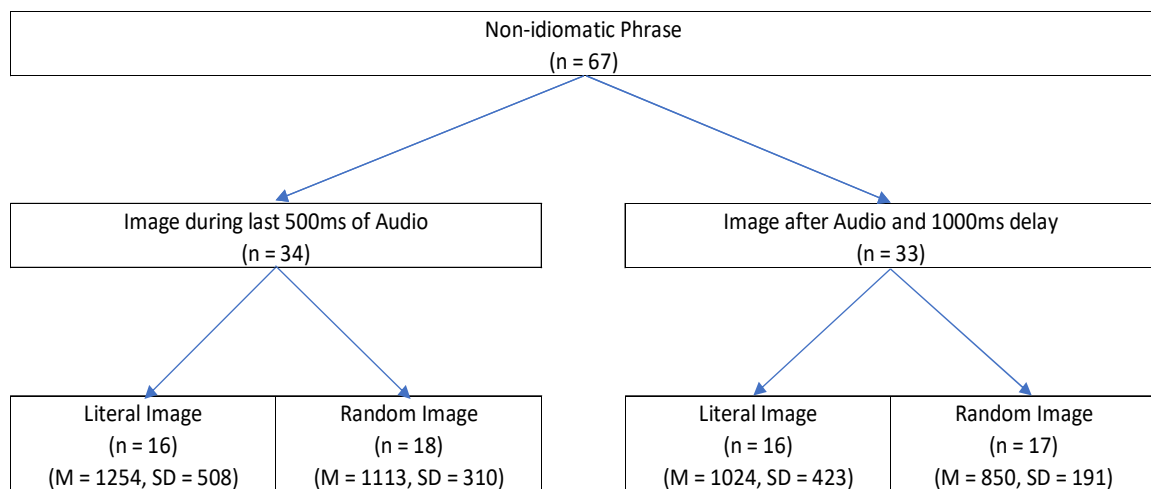
5.

Q: Can you try to get along with Ronny tonight?

A: Yes, but you know that guy *drives me up the wall*.Figurative ImageLiteral ImageRandom Image

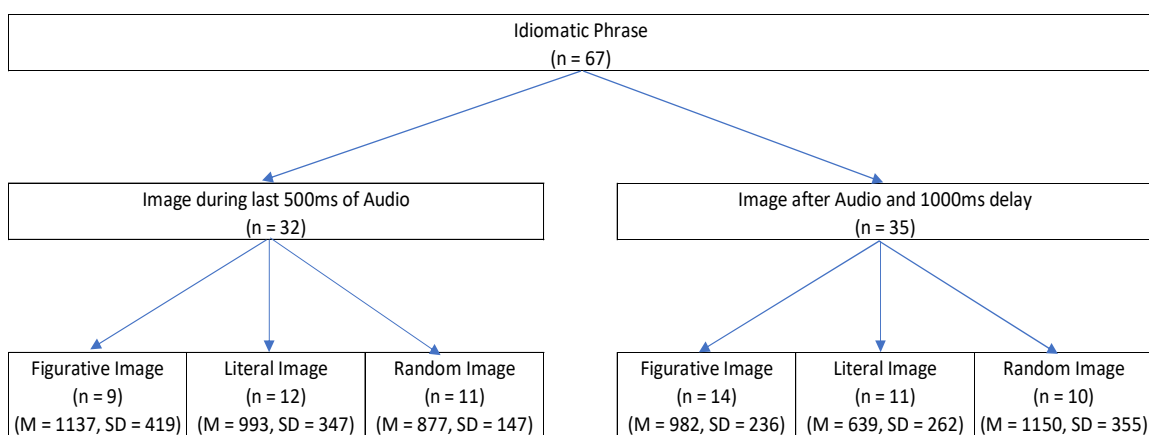
6.

Q: Would this shade of purple go with a bright orange?

A: You cannot possibly like this for the kitchen.Literal ImageRandom Image

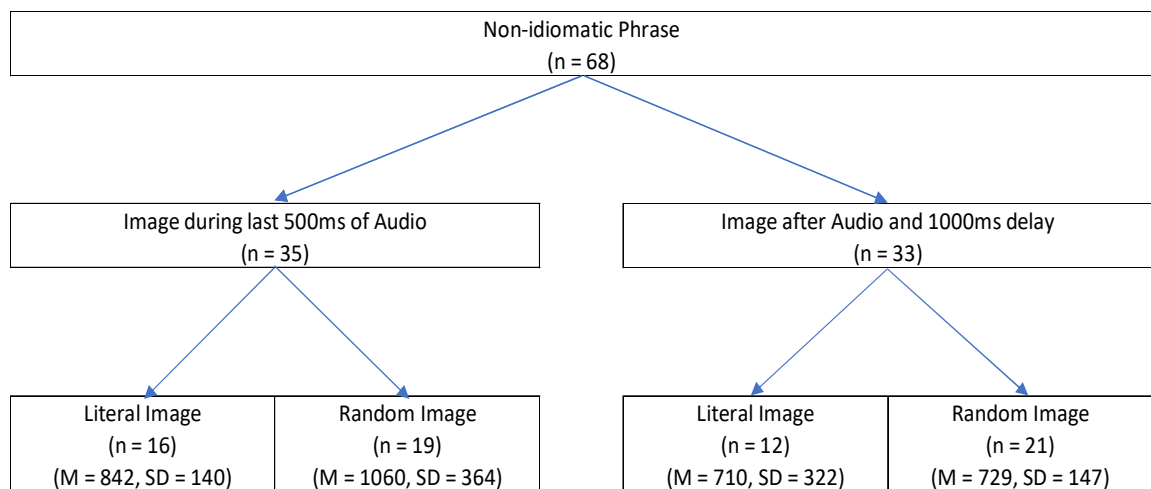
7.

Q: Should we ask the suspect who else is in his gang?

A: We can try, but I doubt he will *spill the beans*.Figurative ImageLiteral ImageRandom Image

8.

Q: How is your mother coping with you being gone?

A: She is fine. She is buying a new puppy.Literal ImageRandom Image

9.

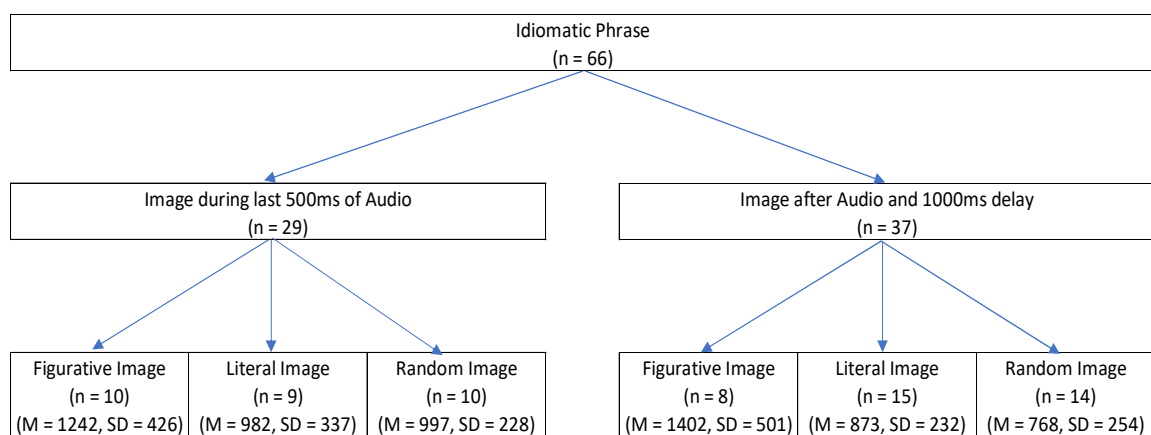
Q: Are you going to work things out with Stephanie?

A: No, we are both ready to *throw in the towel*.

Figurative Image

Literal Image

Random Image



10.

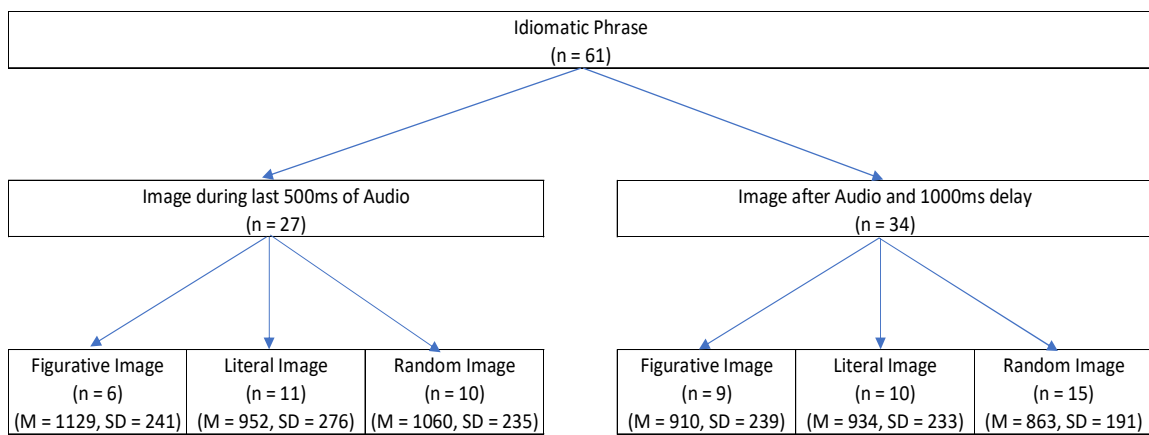
Q: Is it true that Karen's husband is really sick?

A: It is worse than that. He may ***kick the bucket***.

Figurative Image

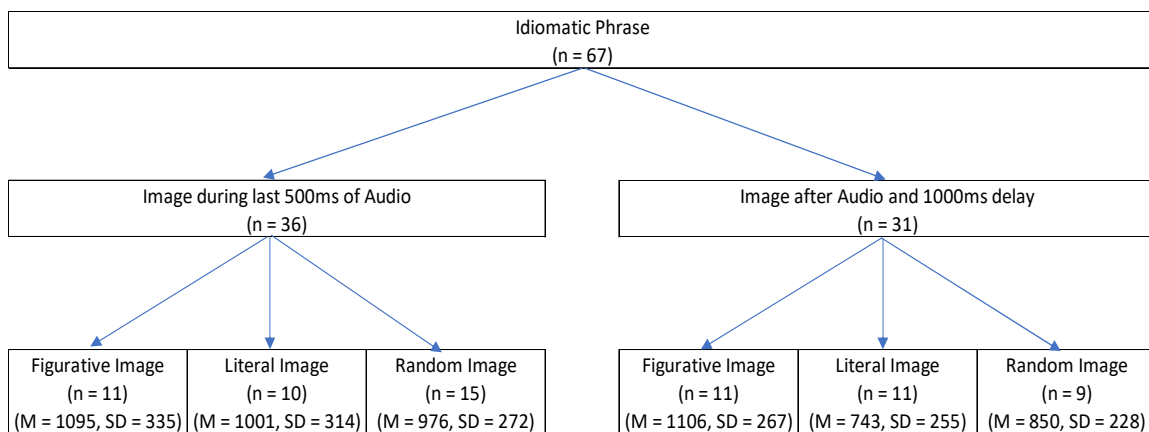
Literal Image

Random Image



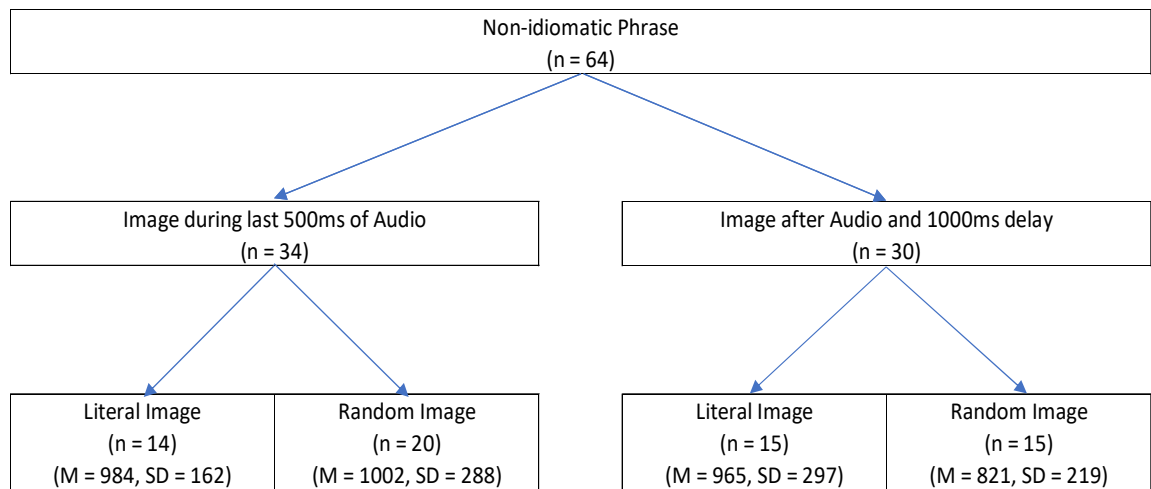
11.

Q: Are you going to try that patch to quit smoking?

A: No, I think maybe I can *go cold turkey*.Figurative ImageLiteral ImageRandom Image

12.

Q: Where have all the chocolate chip cookies gone to?

A: Ask your son. He has lots of crumbs on his shirt.Literal ImageRandom Image

Appendix B. Recruitment Materials

Recruitment Script (Extra Credit)

Thank you for allowing me time to present you with this opportunity. My name is Dan Vandehey, and I am a graduate student in the Communication department conducting a research study on English language and cognition under the supervision of Dr. Ritchie. I would like to invite each of you to participate in this study for an opportunity to earn [number determined by Instructor] points of extra credit in this class.

Participants in this study should be 18 years of age or older. Participation in this study is completely voluntary and your responses will remain confidential. Choosing to participate or choosing to not participate will NOT negatively impact your course grade in any way. If you choose to participate in this study, you will take part in a short, computer-mediated experiment. The experiment should take approximately 5 to 8 minutes.

This opportunity is being offered in [Course (Instructor)], and [Course (Instructor)]. You may only take part in this study once. However, if you: 1) choose not to participate in this study, 2) decide to withdraw early, or 3) have already participated in this study for extra credit in another course, you will have the option to complete an alternative extra credit assignment, which must be requested by 5:00 pm on [day, date], and completed and submitted by 5:00 pm on [day, date]. If you are interested in this option, you need to email me at vdaniel@pdx.edu to request the assignment.

Your classmates should be passing around a sign-up sheet. If you are interested in taking part in the experiment, please clearly print your name and PSU email address on this sheet. In addition, please indicate all times when you could be available to participate on the grid to the right of your name and email address by placing an X in each of the corresponding time slots. I will be scheduling open sessions according to times that accommodate the most people. If you sign up on the sheet, I will send you an email to let you know when those days and times will be.

Information regarding this study will be posted on the D2L home page for this course, and my contact information will be provided there as well.

If you have any questions, feel free to contact me at the email address provided.

Thank you for your time. Your participation is greatly appreciated.

D2L Post (provided to IOR or designated TA)

Hello. My name is Dan Vandehey, and I am a graduate student in the Communication department conducting a research study on English language and cognition under the supervision of Dr. Ritchie. I would like to invite you to participate in this study for an opportunity to earn [number determined by Instructor] points of extra credit in this class.

Participants in this study should be 18 years of age or older. Participation in this study is completely voluntary and your responses will remain confidential. Choosing to participate or choosing to not participate will NOT negatively impact your course grade in any way. If you choose to participate in this study, you will take part in a short, computer-mediated experiment. The experiment should take approximately 5 to 8 minutes.

This opportunity is being offered in [Course (Instructor)], and [Course (Instructor)]. You may only take part in this study once. However, if you: 1) choose not to participate in this study, 2) decide to withdraw early, or 3) have already participated in this study for extra credit in another course, you will have the option to complete an alternative extra credit assignment, which must be requested by 5:00 pm on [day, date], and completed and submitted by 5:00 pm on [day, date]. If you are interested in this option, you need to email me at vdaniel@pdx.edu to request the assignment.

When I visited your classroom your classmates passed around a sign-up sheet. If you were absent at that time or you did not get to sign the sheet, and you are interested in taking part in the experiment, please send me an email to inform me of your interest. Also, please follow the link below to my doodle.com poll to indicate the times you would be available to participate in this short experiment.

[html link to doodle.com poll]

I will be scheduling open sessions according to times that accommodate the most people. If you send me an email, or if you have already signed up on the sheet passed around in class, I will send you an email to let you know when those days and times will be.

If you have any questions, feel free to contact me at vdaniel@pdx.edu.

Thank you for reading. Your participation is greatly appreciated.

Example Email

Subject: Extra Credit Opportunity for COMM[###]!

Hello,

You are receiving this email message because you indicated you would be interested in taking part in my research study about English language and cognition. I want to thank you for this interest, and let you know that open sessions have been scheduled for the study. Please see the attached flyer for information regarding dates, times, and the location. You can show up at any time between the scheduled hours, and you are not required to make an appointment to participate.

If you have any questions that the flyer cannot answer, please feel free to contact me via email.

I hope to see you there!

Daniel A. Vandehey

Graduate Student/Teaching Assistant

Department of Communication

Portland State University

vdaniel@pdx.edu

Example Flyer (see next page)

Extra Credit!

You are invited to participate in a research study conducted by graduate student Daniel Vandehey under the supervision of Professor L. David Ritchie regarding the English language and cognition.

As a reminder, your instructor has agreed to award extra credit toward this class (COMM[###]) for participation in the study.

Although it is possible that some of the images or passages in the study may contain language or refer to ideas that you find disturbing, we do not anticipate that participation in this study entails any significant risk of harm.

Participation is completely voluntary, and you may contact Dan Vandehey (vdaniel@pdx.edu) to receive an alternative assignment if you wish to earn extra credit without participating in the study. Alternative assignments must be arranged before 5:00 pm on [day, date], and completed assignments must be submitted by 5:00 pm on [day, date]. Both options are voluntary, and you will have the option to withdraw from the study at any point you choose.

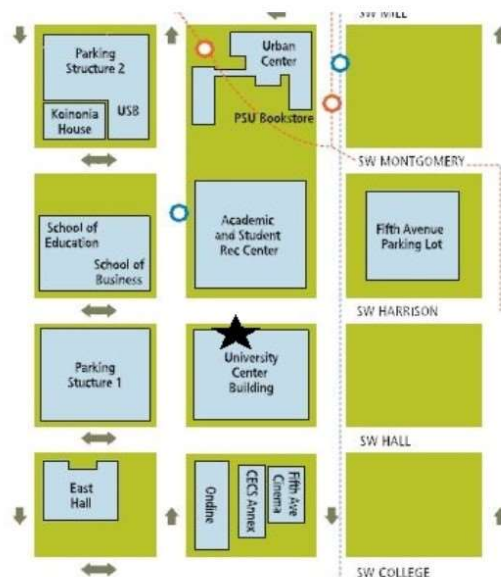
All responses will be submitted anonymously, meaning your personal information will be recorded separately from you responses.

The survey will be administered between [times, day, date] in the Department of Communication Offices in UCB440 (see map and directions).

The survey should take approximately 5-8 minutes.

If you choose to participate, you will need to report to the Department of Communication Offices (UCB440) and check in with personnel at the front desk.

The Department of Communication offices are located on the fourth floor of the University Center Building. Take the elevator in the north entrance (indicated by the star on the map). When the doors open on the fourth floor, you will be facing the entrance to the department offices.



Informed Consent

You are invited to participate in a laboratory-based, computer-mediated experiment regarding the English language and its related cognitive processes. This is part of a research project administrated by Daniel Vandehey, a graduate student of Communication at Portland State University, under the supervision of Professor L. David Ritchie, PhD.

COMMITMENT AND DURATION

The experiment should take approximately 5-8 minutes to complete. The experiment will play a total of 12 each short, two-sentence dialogs accompanied by a brief viewing of an image. The experiment is designed to register how quickly you can determine if the image depicts the meaning of the short dialog. After you have completed the experiment today, there will be no need for any form of follow-up. The researchers will not initiate contact with you following your completion of this one-time activity, but this does not prohibit you from contacting the researchers if you have any residual questions or concerns.

PARTICIPATION

Your participation in this experiment is voluntary. You may refuse to take part in the research or exit the experiment at any time.

If you wished to participate to receive extra credit for a PSU Communication course: If you choose to exit the experiment you may opt to complete an alternative assignment. If the assignment is completed and submitted by the deadline indicated, your name will be added to the list of participants we will compile for your instructor.

BENEFITS

You may not receive any direct benefit from taking part in this study (beyond receiving extra credit for a course if this applies to you.) However, it is our hope that through your participation the study may help to increase knowledge that might help others in the future.

RISKS

Experiments like this one have been known to require great concentration and may cause mental fatigue. This experiment is designed to allow you to move at your own pace. We encourage you to find and maintain a pace that is comfortable. This experiment also contains images, and though the researchers have tried to use images that are benign and appropriate, if you find any imagery disturbing please remember you have the option of leaving the experiment at any time. Aside from these two issues, there are no foreseeable risks involved in participating in this study other than those encountered in day-to-day life.

CONFIDENTIALITY

You will not be asked for any information regarding your identity within the experiment itself. Responses will be submitted electronically and will be completely anonymous.

This form and any other identifying information (such as your name on a list to receive extra credit) cannot be cross-referenced to the data you provide during the experiment. Any information you provide will be maintained in accordance with federal law and PSU Institutional Review Board guidelines until such time as the records can be legally destroyed/erased.

CONTACT

If you have questions at any time about the study or the procedures, you may contact the research supervisor, Professor L. David Ritchie, PhD via email at cgrd@pdx.edu.

If you feel you have not been treated according to the descriptions in this form, or that your rights as a participant in research have not been honored during the course of this project, or you have any questions, concerns, or complaints that you wish to address to someone other than the administrator or research supervisor, you may contact the Portland State University Institutional Review Board at :

PSU Institutional Review Board
 Research Integrity, Research & Strategic Partnerships
 1600 SW 4th Ave., Richard and Maurine Newburger Center, Ste. 620
 Portland, OR 97201
 (503) 725-2227 or 1 (877) 480-4400

CONSENT:

You are making a decision whether to participate in this study. Your signature below indicates that you have read the information provided (or the information was read to you). By signing this consent form, you are not waiving any of your legal rights as a research participant.

You have had an opportunity to ask questions and all questions have been answered to your satisfaction. By signing this consent form, you agree to participate in this study. A copy of this consent form will be provided to you.

Name of Adult Subject (print)

Signature of Adult Subject

Date