Are Figurative Tropes Unique? An Eye Tracking Comparison of Metaphors, Similes, and Idioms

BY

SPENCER J. CAMPBELL
B.A. Augustana College, 2010
M.A. University of Illinois at Chicago, 2014

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Psychology
in the Graduate College of the
University of Illinois at Chicago, 2017

Chicago, Illinois

Defense Committee:
Gary Raney, Chair and Advisor
Susan Goldman, Psychology
Jennifer Wiley, Psychology
Kara Morgan-Short, Psychology
Jane Ashby, Central Michigan University

This thesis is dedicated to my loving family, without whom it would never have been accomplished.

ACKNOWLEDGMENTS

I would like to thank my thesis committee—Gary Raney, Susan Goldman, Jennifer Wiley, Kara Morgan-Short, and Jane Ashby—for their support. Their guidance throughout this process was invaluable, and I cannot thank them enough.

In addition, I would like to thank Gary Raney for taking a chance on me as an aspiring graduate student. You allowed me to explore my interests in new and exciting ways, and were always there alongside me. Lastly, I'd like to thank Dan Corts for inspiring me to study figurative language so many years ago. Thank you, Dan, for helping lead me to this path and shaping me into the researcher I am today.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (continued)

SUMMARY

Many models of figurative language processing have been proposed that emphasize the uniqueness of a given trope. For example, metaphor research typically falls into support of either the Career of Metaphor (Bowdle & Gentner, 2005) or Categorization (Glucksberg, McGlone, & Manfredi, 1997) models of metaphor processing. Because of this, more efforts have been made to show the differences in tropes rather than the similarities. However, throughout the decades of figurative language research, a consistent finding has been the importance of familiarity in how a trope is processed.

The goal of this study was to explore an alternative model of figurative language processing that I have proposed, called the Figurative Funnel. This model predicts that figurative language processing occurs either through meaning construction or direct access of meaning. Which processing style is used is based on the familiarity of the figurative phrase. Unfamiliar tropes use meaning construction, but the style of this construction can vary based on the trope. As tropes become more familiar, they shift to a direct access of the meaning process, unifying how all figurative tropes are processed.

Two eye-tracking experiments were conducted to determine the validity of the Figurative Funnel model for explaining how metaphors, idioms, and similes are processed. Patterns of eye movements were analyzed to determine if unique processing strategies were being used for unfamiliar metaphors across the three tropes. In addition, I examined whether there was a shift in eye movement patterns as tropes became more familiar, supporting a shift in processing style from meaning construction to direct access. Results showed some familiarity effects in reading time and fixation count for metaphors and idioms, but no familiarity effects for similes. In addition, metaphors and similes appeared to use a very similar processing styles regardless of

familiarity. Familiar and unfamiliar idioms did have different patterns of eye movements. Overall, there was little support for the figurative funnel for metaphors and similes, but there was support based on idioms. Implications for other models of figurative language processing are discussed.

Introduction

Figurative language is often considered a unique aspect of language because it provides us with an opportunity to convey ideas or connections that may be difficult to comprehend at a purely literal level (Colston, 2005; Gentner & Wolff, 1999; Gibbs, 1994; Gibbs, 2002; Gibbs, Leggitt, & Turner, 2002; Giora & Fein, 1999). Figurative language goes beyond the literal interpretation of the words and phrases used to convey an alternative or special meaning. For example, if I told you "*My lawyer is a shark*," you would understand that my attorney is not actually an animal of the ocean, but that my lawyer is especially aggressive and tenacious. Because there are dozens of figurative forms, also called tropes, such as metaphors, idioms, and similes, creating a model that can explain all figurative language use is difficult. Instead, researchers have attempted to create models for each individual trope, generally treating each as a unique experience from all other figurative tropes, despite a number of characteristics that many of them share (Bowdle & Gentner, 2005; Glucksberg, 2001; Titone & Connine, 1999).

The research described here cataloged and compared eye movements when reading three figurative tropes: metaphors, similes, and idioms. Current theories make predictions about how the tropes are processed and how much time will be needed to process familiar and unfamiliar versions of the tropes. The models do not make predictions about patterns or durations of eye movements, so these data provided by the present study are very useful in assessing current theories of figurative language processing. This project provides a direct comparison of processing strategies by comparing patterns of eye movements for these tropes. This is essential information as very little direct comparison between tropes currently exists. The most common comparison made in past research has been between metaphors and similes, due to their similar syntactic and semantic structure (Gibbs, 2001; Glucksberg, 2001; Glucksberg & Haught, 2006;

Johnson, 1996; Miller, 1976; Ashby et al., 2017; Shibata, Toyomura, Motoyama, Itoh, Kawabata, & Abe, 2012). However, very few comparisons are made between other tropes, instead treating the tropes as entirely unique from one another.

I have proposed (Campbell, 2015) a model that is generalizable across tropes that accounts for how figurative language is processed based on the most important process factor: familiarity (Blasko & Connine, 1993; Bowdle & Gentner, 2005; Campbell, 2014; Frisson & Pickering, 1999; Pickering & Frisson, 2001). Research shows that the meaning of familiar tropes is accessed largely automatically, whereas the meaning of unfamiliar tropes must be constructed. What follows is a review of these two processes used for figurative language comprehension, a description of the importance of familiarity, and a description of a new model for understanding how the two processes are related.

Two Varieties of Processing

Based on my own research and the current literature, I proposed that figurative language is processed in one of two ways: Meaning Construction, and Direct Access of Meaning (Campbell, 2016). Before describing when these two processes are used, I will begin with a description of these two processes, provide evidence for the processes, and provide examples of how they apply to the three tropes examined in this project (metaphors, similes, and idioms).

Meaning Construction

Meaning construction is the process by which the intended meaning of the figurative phrase is not readily apparent and it must be created in some manner. How this occurs is different across models and tropes. In metaphors, the Career of Metaphor model predicts that meaning construction is necessary when processing unfamiliar metaphors (Bowdle & Gentner, 2005). For unfamiliar metaphors, literal characteristics of the target (the first word in the

metaphor) and vehicle words (the last word in the metaphor) are activated and aligned with one another. This involves matching similar characteristics of the target and vehicle words to generate an inferred meaning of the metaphor. For example, in the metaphor a *lawyer is a lighthouse*, the reader could activate and match the features of *guidance* or *assistance* between the target (*lawyer*) and vehicle (*lighthouse*) words in order to understand how the lawyer is acting as a source of guidance for the subject of the story. According to the Career of Metaphor model, familiar metaphors have their meaning accessed directly using the figurative meaning of the vehicle word. The vehicle acts as a prototypical member of a category (such as *shark* representing *aggression* or a *predatory* nature, or a *lighthouse* acting as a *guide*), and the characteristics of that category are then asserted onto the target word.

Similes, based on their structure, invite the reader to go through a process of meaning construction when comparing the "likeness" of the compared objects. For example, in the simile *the lawyer is like a shark*, the features of *aggressiveness* and *tenacity* are activated for both *lawyers* and *sharks*, and so they are matched with one another in order to determine the lawyer is particularly tenacious in the courtroom. In these instances, readers use the same comparison process that is described in the Career of Metaphor model (Bowdle & Gentner, 2005). In fact, research shows that unfamiliar metaphors are preferred in simile form, meaning readers give higher preference ratings to as well as read phrases faster that have the word "like" added to them to make them similes rather than metaphors, as the word like more clearly signals to the reader to use a comparison to understand the trope (Bowdle & Gentner, 2005; Gentner & Wolff, 1999). Due to their nature, similes always have their meaning constructed (Bowdle & Gentner, 2005; Glucksberg & Haught, 2006). The interpretations made of similes tend to be focused on the more tangible properties that can be compared between the target and vehicle words, as

opposed to metaphors in which the reader is more focused on emergent properties of the vehicle word (Glucksberg & Haught, 2006; Hasson, Estes, & Glucksberg, 2001).

According to the compositional approach for processing idioms proposed by Nunberg et al. (1994), all idioms, regardless of familiarity, require meaning construction. This type of meaning construction is very different from the comparison approach that is necessary for understanding unfamiliar metaphors and for similes (Bowdle & Gentner, 2005; Gentner & Wolff, 1999). Nunberg et al. (1994) claim that meaning construction is required because the literal meaning of the component words of an idiom always contribute to its figurative meaning to some degree. The level of necessary construction is dependent on the idiom's compositionality, which refers to the degree to which the meaning of the individual words contribute to the overall meaning of the phrase. For example, the idiom *shoot the breeze* is considered to have low compositionality because the act of shooting, and the presence of breeze has very little to do with having a casual conversation. An example of a highly compositional idiom would be *speak your mind*, because the subject is literally speaking the thoughts that they are currently having in their mind.

Other researchers (Swinney & Cutler, 1974; Titone & Connine, 1999) claim that some idioms are nondecomposable, in which the literal meaning of their component words do not add any meaning at all. For these idioms, meaning is accessed as if it were a very long word string. Nunberg et al. (1994) claim that nondecomposable idioms, which are assumed to be held in the lexicon as a single unit, require some meaning construction based on the transparency of the verb-phrase. For instance, while *sawing logs* has no obvious connection with sleep, the sound associated with sawing can be linked to snoring.

According to the hybrid model of idiom processing proposed by Titone and Connine

(1999), meaning construction is necessary for unfamiliar idioms, but not familiar idioms. The meaning of unfamiliar idioms is constructed based on the compositionality of the idiom's component words, as well as contextual cues (i.e. the transparency of idiom's use). The literal meaning of the words in the idiom provide guidance to the intended meaning. Consider the situation in which you are just learning the meaning of the idiom *kick the bucket*. The meaning of the idiom would be constructed through a combination of contextual cues (used in a situation where someone has died) as well as the literal implication of the word *kick*, which refers to a fast action (the person died quickly, as opposed to a long, dragged out death).

Direct Meaning Access

The second general process found across many models of figurative language processing is direct access. This is the process in which the meaning of the figurative word or phrase is accessed directly from the lexicon, therefore, no meaning construction occurs.

In the case of metaphors, according to the categorization model proposed by Glucksberg et al. (1983), the meaning of the vehicle word in a metaphor is accessed directly and in parallel with the literal meaning. This figurative meaning is asserted onto the target word, allowing the reader to understand the metaphor automatically. Because the vehicle word of the metaphor acts as an exemplar of a category, its intended figurative meaning is accessed quickly and directly. Glucksberg claims that this process is used for all metaphors, so long as they are used aptly (Glucksberg, 2008). Aptness describes whether the metaphor is being used appropriately based on the context in which it appears. A category for the vehicle cannot be accessed if the reader is unclear whether they should activate it. For example, one would not walk up to a family member and say *death is a thief* in order to convey the passing of a loved one. Instead, the speaker may use that metaphor in a context in which they have described the loved one struggling with an

illness, and that *death is a thief* in that it has taken them away. Some studies have demonstrated the importance of aptness, and even shown how it may be more important than familiarity (Jones & Estes, 2006; Roncero, de Almeida, Martin, & de Caro, 2016). The Career of Metaphor model also proposes that direct access of a metaphor's meaning is possible when the metaphor is highly familiar to the reader (Bowdle & Gentner, 2005). Familiar metaphors use the same categorical assertion process described by Glucksberg et al.'s (1997) categorization model.

While it may be the case that highly familiar metaphors are accessed directly because of ease of activation of the multiple meanings of the vehicle, it could be argued that some level of meaning construction is required when the categorical properties are asserted onto the target word. That is, when the figurative meaning of *shark* is asserted onto *my lawyer* in the metaphor *My lawyer is a shark*, the reader is taking characteristics from the vehicle word and applying them to the target word, constructing a new, nonliteral meaning for the figurative phrase. The most important part of this process is the direct access of the figurative meaning of the vehicle though, making it an inherently different process than the meaning construction by comparison process used with unfamiliar metaphors described in the Career of Metaphor (Bowdle & Gentner, 2005).

Going a step further, some metaphoric phrases and words become so familiarized that they lose their intended literal meaning entirely, becoming "dead" metaphors (Bowdle & Gentner, 2005). An example of this would be the word *blockbuster*. The original definition of this word referred to a bomb meant to destroy a city block. Now, however, it is almost exclusively used to describe a smashing success, typically referring to a movie. In these cases, the figurative meaning is directly accessed from the lexicon, and with no interference from a literal meaning. For words such as *blockbuster*, no construction is needed or even possible

because the "success" meaning is the only meaning stored in memory. In some sense, "success" is the literal meaning because no alternative exists.

For idioms, the noncompositional model proposes that all idioms are accessed directly from the lexicon (Swinney & Cutler, 1979). This is due to the fact that idioms are considered to be similar to long, single-word units, and so they are stored as a single unit with a single meaning. This model makes the assumption that all idioms behave this way, and that direct access begins after reading the first word in the idiom. Swinney and Cutler make the claim that this occurs for both familiar and unfamiliar idioms. It is unclear however, how this would be possible for unfamiliar idioms. One can't help but ask how is a meaning directly accessed if it has never been encountered before? While this direct approach may be appropriate for idioms the reader is familiar with, the noncompositional model is insufficient for explaining how unfamiliar and novel idioms are processed.

The hybrid model of Titone and Connine (1999) also claims that familiar idioms are accessed directly. They claim that familiar idioms are not dependent on the decomposability of the idiom's component words, and instead assume that with enough exposure, an idiom is stored as a single unit in the lexicon. For both familiar and unfamiliar idioms, some level of literal access will occur. However, Titone and Connine claim that this is only completely necessary for unfamiliar idioms, and that access of the literal meaning of the words will end in familiar idioms before the reader has finished reading it, instead switching to directly accessing the meaning of the phrase as a whole.

The Role of Familiarity

I will now describe if, and how models of figurative language processing account for familiarity effects. Some models directly account for familiarity effects, whereas some models

do not do so adequately, or at all. I will also try to provide some insight into why this might be the case.

The two major models of metaphor processing are Gluckberg's categorization model (1991; 2001; 2008) and Bowdle and Gentner's career of metaphor model (1997; 2005). Glucksberg focuses on the aptness of the metaphor rather than the familiarity. Glucksberg makes the assumption that regardless of familiarity, so long as a metaphor is used aptly in the context in which it is provided, it will always use the categorical assertion process (Glucksberg, 2003; Glucksberg, McGlone, & Manfredi, 1997). The vehicle word acts as the exemplar of a category (*shark* representing the concepts of aggression and a predatory nature), and the characteristics of that category are asserted onto the target word, making it a new member of the category. Glucksberg (2008) has pointed out that the categorization model cannot account for novel metaphors that are inappropriately used (i.e., no context or an ill-fitting context). Such instances are not representative of how we naturally use metaphor however, and so he still strongly makes the claim that aptness is far more relevant in metaphor processing that familiarity. However, Glucksberg (2008) has ceded the point that truly unfamiliar metaphors may need an entirely different process, especially when inappropriately used. In those instances, a feature matching comparison processing may be necessary.

The Career of Metaphor model clearly describes how the familiarity of a metaphor will impact how it is processed (Bowdle & Gentner, 1997, 2005; Gentner & Bowdle, 2008). Unfamiliar metaphors behave the same as literal comparisons and similes, and thus use a relatively resource heavy process of feature comparison and matching. Familiar metaphors use the more easily and rapidly completed process of categorical activation and assertion described in the categorization model (Glucksberg, 2003, 2008). Beyond describing how familiarity affects

metaphor processing, Bowdle and Gentner describe a trajectory, the "career", which a metaphor will follow over its lifetime as it moves from novelty to high familiarity. In doing so, Bowdle and Gentner have provided a map that clearly shows how important familiarity is to metaphor processing. The Career of Metaphor model has continued to receive support from cognitive modeling (Utsumi, 2011), divided visual field assessment (Chettih, Durgin, & Grodner, 2012), as well as conventionality testing (Gokcesu, 2008).

Models of idiom processing also fall into two camps, those that claim familiarity is an important factor and those that do not account for it. Noncompositional models, and specifically the lexical representation model, propose that all idioms, regardless of familiarity, are accessed directly from the lexicon (Swinney & Cutler, 1978). Figurative and literal meaning access of the phrase occurs in parallel, but figurative meaning access is completed more quickly. How novel idioms can possibly be stored in the lexicon prior to learning their meaning is unclear. It may be the case that the lexical representation model is appropriate for familiar and even less familiar idioms, but truly novel idioms cannot be accounted for by this model.

The compositional (Gibbs & Nayak, 1989; Nunberg, 1994) and hybrid models (Titone & Connine, 1999) of idiom processing do make claims about how familiarity affects processing. Nunberg and colleagues (1994) claim that highly familiar idioms are much easier to construct a meaning for, as the meaning has been constructed and encountered before. If an idiom has been encountered a number of times, then its meaning has been constructed a number of times as well. Because of this, any future construction is made easier by the familiarity of the idiom. This can get to the point where construction is no longer necessary at all because the meaning is stored in memory after multiple exposures and then accessed directly for highly familiar idioms (Titone & Connine, 1999).

It is easy to make parallels between the hybrid model of Titone and Connine (1999) and the Career of Metaphor model of Bowdle and Gentner (2005). Both make the claim that familiar and unfamiliar versions of their tropes use different processes, and that how an individual will process the trope changes with time and exposure. For the idiom hybrid model, unfamiliar idioms follow the processes described in the compositional approach. The meaning of the idiom is constructed based on how the literal component words contribute to the figurative meaning of the phrase. Familiar idioms, on the other hand, are processed in the same way described by the noncompositional account. Familiar idioms have their meaning accessed directly due to the fact that they have been encountered so many times that they are stored as a single unit in the lexicon. Thus, the hybrid model and the Career of Metaphor model propose meaning construction for unfamiliar phrases and meaning retrieval for familiar phrases.

The Funnel of Figurative Tropes

Reviewing and evaluating the processes described in the models as well as how they may or may not account for familiarity effects has led me to propose a preliminary model of figurative language processing that can be applied across figurative tropes. I refer to this model as the Figurative Funnel. The model focuses on what I believe to be the single most important factor that influences how figurative language is processed, whether a figurative phrase is familiar or unfamiliar.

Consider if you will a horizontal line representing the continuum of novel figurative phrases on the left side and highly familiar phrases on the right side. By arranging the models described above on this continuum based on what process they rely on (meaning construction or direct access), a very interesting funnel pattern emerges. Figure 1 illustrates the model. At the unfamiliar end, there is a scattering of models that all propose that the meaning must be

constructed, regardless of trope type. What is important to note is that *how* meaning is constructed vastly differs between the models. This is why they form the wide end of the funnel shape--there is a wide range of processes. The process of matching features of the target and vehicle words in a metaphor is very different than piecing together the meaning of an idiom based on the literal meaning of its component words as well as contextual cues. While the processes themselves are very different, they have the same underlying assumption that the meaning of the figurative phrase must be constructed.

As figurative phrases become more familiar through repeated exposure over time, how they are processed changes as well. Specifically, at the high familiarity end of the continuum all of the models are based on directly accessing the meaning of the figurative phrase. Unlike the unfamiliar end, how the models explain direct access is very similar--there is a narrow range of processes. Consider a dead metaphor such as *blockbuster*. In this case, the figurative meaning is accessed directly from the lexicon and does not require the reader to construct the meaning. Similarly, highly familiar idioms are accessed directly from the lexicon, whether by default (Swinney & Cutler, 1978) or because they have been used so often that their idiomatic key word is at or near the beginning of the phrase (Titone & Connine, 1999). The key word is the point in which the reader recognizes the specific idiom and begins to automatically access the phrase's meaning.

Note that some models fall between the two ends of the continuum. Categorization, for instance, is based on the direct access of the figurative meaning of the vehicle word, but this meaning must be incorporated with the target word before the metaphor is fully understood. This process may be done very quickly for familiar metaphors, to the point where it seems as though the meaning of the metaphor is accessed directly, and it may be a slower process for unfamiliar

metaphors, as the reader must use contextual cues to create the figurative category for the vehicle in the first place.

All models of figurative language processing can be placed along the familiarity-processing style continuum. In this way, the focus is on the degree of meaning construction or meaning retrieval. Rather than treating each trope as a uniquely different class of figurative language dependent on unique processing styles, I believe the push should be to demonstrate the underlying processing similarities and differences. Familiarity is constantly pushed to the forefront of empirical research (Blasko & Connine, 1993; Bowdle & Gentner, 2005; Connine et al., 1992; Cronk & Schweigert, 1992; Nippold & Taylor, 2002; Schweigert & Moates, 1988; Titone & Connine, 1999), and so it seems natural to use it as the unifying factor of a general figurative language processing model. Based on the literature I have read and reviewed, a general model of figurative language processing would predict the following:

- Across figurative tropes, unfamiliar phrases will have their meaning constructed. How this is done will vary greatly across the tropes. As figurative phrases become more familiar their meaning will be accessed more directly. Direct access will lead to more similar processing across figurative tropes.

The assumption that tropes behave more similarly as they shift from unfamiliar to familiar is why a figurative language "funnel" shape is appropriate.

Eye-Tracking and Figurative Language Research

The goal of this study was to explore the hypothesis that processing strategies across tropes, as reflected by eye movement behavior, will become more similar as familiarity of the phrases increases (for an overview of the relevant eye movement measures, see Appendix A). I compared eye movement patterns for metaphors, similes, and idioms. Doing so allowed me to

create a catalog of eye-movement patterns for each trope, a task that has not yet been done in prior research. Such a catalog allowed me to make comparisons of how the tropes are processed as a function of familiarity. Very little empirical research has been done using eye-tracking to examine how figurative language is processed. There have been a handful of studies that have explored irony and sarcasm (Baptista, Macedo, & Boggio, 2015; Filik, Leuthold, Wallington, & Page, 2014; Filik & Moxey, 2010; Olkoniemi, Ranta, & Kaakinen, 2015) and metonymy (Frisson & Pickering, 1999). However, very few studies exist exploring the tropes covered in this paper using eye tracking.

The most beneficial aspect of using eye tracking is that it provides data on the patterns of eye movements as well as reading times for individual words, rather than just overall reading time. Eye movement patterns can help distinguish how various tropes might be processed differently from one another, a conclusion that cannot be made from reading time data alone. Several eye movement measures were evaluated as part of this research. These include reading time and fixation counts, which are important measures of processing difficulty (reading time and number of fixations increase as processing difficulty increases). In addition, the number of regressions out of individual words, into individual words, as well as into the preceding context were evaluated.

While the previous models of figurative language processing may not make predictions specifically about eye movements, some inferences can be made of what sorts of patterns might emerge based on those theories. For example, if a model predicts more difficult processing for an unfamiliar version of a trope compared to familiar, I would expect to see increased reading time for the phrase as a whole, as well as the individual words, increased fixation counts, more refixations, and more regressions to and from the phrase and its individual words. As another

example, if a model predicts that a comparison process takes place, then I would expect to see an increased number in regressions to and from the words being compared, creating longer reading times for those words as well and greater fixation counts, when compared to tropes that do not use a comparison process.

I (Campbell, 2014) previously found consistent familiarity and context effects in eye-tracking measures when reading metaphors. Metaphors were presented in short, four-sentence passages that either strongly supported the meaning of the figurative meaning of the metaphor or moderately supported the figurative meaning. A strongly supportive passage had two sentences that primed the reader for the figurative meaning of both the target and vehicle words. The moderately supportive context only primed the reader for the vehicle word. I found shorter (faster) total reading times for familiar metaphors than unfamiliar metaphors as a whole (the entire phrase) as well as for the individual target and vehicle words in the metaphors. This pattern also was found for total fixation counts at the phrase and word level, with fewer fixations for familiar than unfamiliar metaphors. In addition, I found that metaphors presented within strongly supportive contexts had faster reading times and required fewer fixations and regressions into the phrase than metaphors presented in moderately supportive contexts. I found a consistent "stair step" pattern in terms of processing difficulty. Specifically, familiar metaphors in strong contexts were the easiest to process in terms of reading times (fastest) and fixation counts (smallest), followed by "step up" in difficulty for familiar metaphors in moderate contexts (longer reading times and more fixations), another step up in difficulty for unfamiliar metaphors in strong contexts, and a final step up in difficulty unfamiliar metaphors in moderate contexts (longest reading times and the most fixations). Of particular interest is familiarity had a much larger effect on reading times and fixation counts than did context strength. This is consistent

with the idea that familiarity is the dominant factor in terms of how a metaphor is processed. The same basic pattern was found for several other measures, such as the number of regressions from the vehicle word to the target word and regression loops. The pattern of results for first pass and second pass times were inconsistent at the phrase level. This makes sense for first pass times given that readers might not realize they are reading a metaphor until they get to the last word. As such, I will focus on total reading times in the present study.

Metaphors and similes have been explored using eye-tracking by a small number of other researchers. For example, Ashby, Roncero, de Almeida, & Agauas (2017) used eye-tracking to compare how metaphors and similes are processed. Metaphors and similes were presented in single sentences, such as *Mary believes that liquor is (like) a crutch because it is used in difficult situations*. The first half of the sentence created a context for the trope (*liquor is [like] a crutch*) and the second half explained the meaning of the trope. Similes were created by adding the word *like* to the metaphors. They found longer first pass reading times for unfamiliar metaphors compared to familiar metaphors and similes. Additionally, they found that participants regressed to the vehicle word more often for metaphors than similes. These findings imply extra processing in which additional cognitive resources are required for metaphors compared to similes, and this is especially the case with unfamiliar metaphors as the reader attempts to consolidate the phrase within the context. It is important to note that because an explanatory clause was at the end of each trial, it makes sense that there were more regressions back to the vehicle in order to consolidate the explanation with their own interpretation (or lack thereof). In addition, the context prior to the metaphor was minimal, and existed to present the trope in a more natural way than purely isolation. Because of this, my previous work and the work presented are not a replication of the work of Ashby and colleagues, and therefore direct comparisons between the

studies is difficult. Blasko and Briihl (1997) tracked eye movements during reading of metaphors in a very short (one sentence) context and found similar findings to Ashby et al., with shorter first pass reading times for familiar metaphors compared to unfamiliar metaphors. My prior work (Campbell, 2014) and the current study used a supportive context prior to the trope. It will be interesting to see if I find the same pattern of results found by Ashby et al. for metaphors and similes when the tropes follow strongly supportive contexts.

Very little research on idioms has made use of eye-tracking methodologies. Of note is Titone and Connine's (1999) work in which they describe their hybrid model of idiom processing. They used eye tracking to explore how the decomposability of an idiom impacts processing in an attempt to determine when compositional and noncompositional theories are appropriate. Idioms were presented in single sentence contexts that were biased towards the literal or figurative meanings. Importantly, the biasing context could appear before or after an idiom. Nondecomposable idioms were read slower than decomposable idioms when the context appeared before the idiom than when it appeared afterwards. Decomposable idioms showed no difference in reading times regardless of position of the context. The authors took this as evidence that the literal components of the idioms are always activated, and that this is beneficial for decomposable idioms but a detriment for nondecomposable idioms in that there was a conflict between the literal meaning of the component words and the figurative meaning of the phrase, causing a slowdown in reading. If the individual component words of a nondecomposable idiom do not directly help with its interpretation, and access of their literal meaning will interfere with the processing of the figurative meaning of the phrase as a whole.

Design Rationale

While little research has been completed using eye tracking methods to examine

figurative language processing, the few studies described above show potential for uncovering the different processing strategies used for different tropes. In order to begin evaluating the Figurative Funnel model, two experiments were conducted to examine differences in eye movement patterns between familiar and unfamiliar versions of metaphors and similes (Experiment 1), and metaphors and idioms (Experiment 2).

Metaphors and similes were selected because they are the most common comparison made between tropes, in the rare instances when such comparisons occur. In addition, metaphors have dominated the figurative language literature for the last 30 years, and so their inclusion was an obvious choice. Metaphors and similes also share a basic structure [X is (like) a Y], which facilitates comparison because each trope includes a target and vehicle. Idioms were selected for comparison to metaphors because of their structural differences. These differences may result in different patterns of eye movements, and these patterns might provide insight into how idioms are processed. In addition, idioms are more culturally entrenched, much more so than metaphors and similes (Buckingham, 2006). Because of this, they might be even more susceptible to familiarity effects, allowing me to better understand the idiom's place along the Figurative Funnel. In addition, idioms might rely on contextual information more than metaphors.

The focus of the study was on two key observations. First, differences in patterns within the tropes. One of the primary principles of the Figurative Funnel, and similar models that incorporate familiarity effects, is that the process used to understand a trope will change with increased familiarity. Therefore, it will be important to observe different patterns for how a given trope is processed when it is unfamiliar compared to when it is familiar. The second observation was changes in pattern within the tropes. These across-trope comparisons were the basis for the shape of the Figurative Funnel. If unfamiliar versions of tropes demonstrate different eye

movement patterns across tropes, it would support the notion that tropes are inherently processed differently when they are unfamiliar versus familiar. Comparisons of familiar tropes would then hopefully show similar patterns between tropes, if the Figurative Funnel is indeed correct.

Predictions

Below are the predictions for the study. The predictions were tailored to explore the Figurative Funnel model, rather than attempting to provide support for the many other models that currently exist in the literature.

Hypothesis 1. I predict a consistent familiarity effect for all tropes, such that unfamiliar tropes will be more difficult to process than familiar tropes. This difference should be seen in all measures that reflect processing difficulty, such as longer reading times for unfamiliar than familiar tropes, a larger number of fixations for unfamiliar than familiar tropes, and a larger number of regressions and regression loops for unfamiliar than familiar tropes.

Hypothesis 2. The first major assumption of the Figurative Funnel is that, within a given trope, there will be a shift in processing style from meaning construction to direct access of meaning. Therefore, I predict that there will be a difference in eye movement patterns for the unfamiliar and familiar versions of each trope. Differences in reading time alone may be indicators of differences in processing style, but I believe that fixation and regression patterns are even more valuable when looking for different processing patterns. For example, unfamiliar metaphors might require the reader to construct the meaning by matching features of the targets and vehicles. This might produce a pattern in which there is an equal number of fixations and equal reading times for target and vehicle words, and possibly regressions from the vehicle to the target words as well. There might be a shift in eye movement patterns for familiar metaphors that would be reflected by (a) less reading time on the entire phrase, (b) more time spent on the

vehicle word than the target, and (c) few regressions between the vehicle and target words. This would reflect the reader directly accessing the category of the vehicle and processing the metaphor, emphasizing the importance of the vehicle over the target. This would indicate a different process is being used than the comparison process proposed for unfamiliar metaphors. It is not feasible to attempt to predict every possible pattern ahead of time. Instead, I evaluated differences within each trope to determine if they reflect a shift in processing style from meaning construction to direct access.

Hypothesis 3. The second major assumption of the Figurative Funnel is that unfamiliar tropes may use a variety of different meaning construction processes in order to be comprehended. Therefore, I predict that there will be differences in the eye movement patterns between the unfamiliar tropes when they use different processes. Differences in reading time are an excellent measure of processing difficulty, and might reflect different processes being used, but I believe that the patterns found in fixations and regressions will better reflect distinct processes. For example, unfamiliar metaphors might require the reader to construct the meaning by matching features of the targets and vehicles. This may produce a pattern in which there is an equal number of fixations and equal reading times for those words, and possibly regressions from the vehicle words to the target words as well. Unfamiliar similes use the same feature matching process, and therefore I would not expect to find different patterns between unfamiliar metaphors and similes. In contrast, an unfamiliar idiom, using context clues to construct meaning instead of having its meaning constructed from the meaning of individual words. As a result, unfamiliar idioms are more likely to have regressions into the preceding context and subsequent regressions from the context following the idiom in order to create a meaning for the idiom based on the narrative. Once again there are many different patterns that could emerge. The key

prediction is that I expect different patterns of eye movements between the unfamiliar versus familiar tropes if they use a different for meaning construction.

Method

Experiment 1

Participants

Participants were 40 UIC subject pool students. Participants were required to have attended English-speaking schools for at least 10 years. This restriction was used to ensure that the familiar tropes were in fact familiar to the participants, and that a language barrier was not causing reading difficulty for the participants.

Design

The experiment used a 2 (Familiarity: familiar vs. unfamiliar) x 2 (Trope: metaphor vs. simile) mixed design, with trope as the between-subjects factor and familiarity being the within-subjects factor.

Materials

The stimuli consisted of 20 high familiar and 20 low familiar metaphors, and 20 high familiar and 20 low familiar similes presented within four sentence passages. All metaphors used the "X is a Y" format. Metaphors and their contexts (passages) were selected from pre-existing stimuli developed by Campbell (2014). The metaphors used by Campbell were taken from familiarity norms created by Katz et al. (1988) and re-normed using UIC students (Campbell & Raney, 2016). Campbell and Raney confirmed that the familiar and unfamiliar metaphors included in this study were generally rated as familiar and unfamiliar, respectively, by UIC students. The UIC student familiarity ratings were also highly correlated with Katz et al.'s ratings. The 20 metaphors were selected to be low familiarity (bottom 25% of familiarity ratings) or high familiarity (upper 25% of familiarity ratings) based on Katz et al. Familiarity was measured on a scale from 1-7, with low familiarity metaphors averaging 2.2 and high familiarity

metaphors averaging 4.6.

The target words were approximately matched for word length in low (M = 6.8 characters) and high (M = 6.1 characters) familiarity metaphors. Likewise, the vehicle words were approximately matched for word length in low (M = 6.3 characters) and high (M = 6.8 characters) familiarity metaphors. Using the Corpus of Contemporary American English (Davies, 2016), the target words were also approximately matched for frequency of occurrence in low (M = 36.5 uses per million) and high familiarity (M = 62.1 uses per million) metaphors. Likewise, vehicle words were approximately matched for frequency of occurrence in low (M = 6.1 uses per million) and high (M = 10.5 uses per million) familiarity metaphors. Due to inherent differences between target and vehicles in metaphors (e.g., vehicles can represent figurative categories), vehicle words have lower word frequencies than target words. The average of the target words for low and high familiarity metaphors were both considered moderately frequent, while the average of the vehicle words was considered low frequency.

Similes were created by adding the word "like" to the metaphor stimuli so that all similes followed the "X is *like* a Y" format. The produced 20 low and 20 high familiarity similes. As the similes were created from the selected metaphors described above, the target and vehicle words for familiar and unfamiliar similes were matched on word length and frequency across tropes. I assumed that adding *like* to a high familiarity metaphor would result in a high familiarity simile, and that adding *like* to a low familiarity metaphor would result in a low familiarity simile.

The tropes were placed within short stories as a means of providing context. Sample stimuli can be found in Appendix B. All contexts (sentences preceding the tropes) were previously normed by Campbell (2014) to confirm the contexts supported the intended figurative meanings of the figurative phrases. In all stories, the figurative trope was used appropriately.

Each story included four sentences. The first two sentences clearly supported the figurative meaning of the trope. In the case of the metaphors and similes, this was done by devoting a sentence each to the target and vehicle words. In the examples in Appendix B, for the metaphor *the mind is a sponge*, the first sentence supports the intended figurative meaning of the target word *mind* by using the word *learning* in an appropriate manner, and the second sentence is intended to support the figurative meaning of vehicle word *sponge* by using the word *absorb* in an appropriate manner.

Forty filler passages were included that contained no figurative language so that readers were unlikely to become aware of the consistent "X is (like) a Y" metaphor or simile that appeared in the third sentence of the passages. These filler passages had the same four sentence structure of the figurative passages.

Half of the passages (40 total) were followed by a multiple-choice question with four alternatives to ensure participants were reading for comprehension. The multiple-choice questions did not refer to the figurative phrases.

The breakdown of the stimuli used based on trope condition for Experiment 1 was as follows:

Metaphor Condition

- 20 passages with familiar metaphors
- 20 passages with unfamiliar metaphors
- 40 fillers passages with no figurative language

Simile Condition

- 20 passages with familiar similes
- 20 passages with unfamiliar similes

- 40 filler passages with no figurative language

Procedure

The experiment was run using an Eyelink 1000 eye tracker. Passages were presented on a 17" flat panel monitor in 20 pt black Arial font, double-spaced, on a white background. The visual angle of the widest line of text was 25.2˚of visual angle. To put this in perspective, there are 64 characters and spaces in the widest line, which means one horizontal degree of visual angle equals approximately 2.5 character positions for that line. Visual angle is not constant for each character because Arial is a proportional font. Double spacing creates approximately 2.5˚ vertically between lines.

Participants were run one at a time. Prior to reading the passages, the participant was randomly assigned into one of the trope conditions (metaphor or simile). The eye tracker was calibrated to ensure accurate measurement of eye position. Participants were required to have an average error of less than 0.5˚ of visual angle before beginning the experiment. After being told how the eye-tracker works and how to answer comprehension questions, participants read two practice passages in order to become comfortable with the procedure. Afterwards, participants read the 80 passages for the study. Passages were presented in a random order. After completing the reading portion of the experiment, participants completed a vocabulary quiz and language history questionnaire.

The English vocabulary quiz (developed by Raney) has been used in several prior studies (Minkoff & Raney, 2000; Therriault & Raney, 2007) and is moderately correlated with comprehension ability (r = 0.40 to 0.52 in past studies). The vocabulary test included 30 items with five alternatives. Prior research indicates the average score is approximately 15-16 items correct (the test is designed to be difficult). The Language History Questionnaire (LHQ) was

used to collect information about participants' language background, such as their native language, years of experience with English, and self-reported proficiency ratings in English and other languages. See Appendix C for the vocabulary quiz and LHQ. These tests were used to ensure that participants in each trope condition were similar overall (i.e., similar vocabulary quiz scores, similar number of native and non-native English speakers). For the metaphor condition, the average vocabulary score was 15.6, and 13 of the 20 participants were bilingual. For the simile condition, the average vocabulary score was 16.3, and 14 of the 20 participants were bilingual. After completing the study, participants were debriefed on the purpose of the experiment and thanked for their time.

Experiment 2

Participants

Participants were 40 UIC subject pool students. Participants were required to have attended English-speaking schools for at least 10 years. This restriction was used to ensure that the familiar tropes were in fact familiar to the participants, and that a language barrier is not causing an increase in processing difficulty for any of the participants.

Design

Experiment 2 followed the same design as Experiment 1, but used idioms as the second trope rather than similes. Therefore, it was a 2 (Familiarity: familiar vs. unfamiliar) x 2 (Trope: metaphor vs. idiom) mixed design, with trope as the between-subjects factor and familiarity being the within-subjects factor.

Materials

The stimuli consisted of 20 high familiar and 20 low familiar metaphors, and 20 high familiar and 20 low familiar idioms presented within four sentence passages. All metaphors used

the "X is a Y" format. The metaphors and passages used in Experiment 1 were used for Experiment 2, and so had the same criteria for selection, matching, and passage creation.

The idioms were all noncompositional. This means that the individual words do not strongly contribute to the reader's understanding of the idiom. The rationale behind using noncompositional idioms is two-fold. First, if compositional idioms were selected, then idiom processing would be more similar to metaphor and simile processing because the literal meanings of the words comprising the idioms could be used to help determine the meaning. As a result, potential differences between the tropes would be weakened. Second, because the meaning of noncompositional idioms cannot be determined based on the literal meanings of their words, they produce stronger familiarity effects than compositional idioms (Titone & Connine, 1999). Idioms were taken from pre-normed idioms that were used to examine familiarity effects in past research (Miller, 2014). The idioms were selected to be low familiarity (bottom 25% of familiarity ratings) or high familiarity (upper 25% of familiarity ratings). Familiarity was measured on a scale from 1-7, with low familiarity idioms averaging a rating of 2.7, and high familiarity idioms averaging a rating of 6.3. Idiom length was also matched between low (M = 3.9 words) and high (M = 3.7 words) familiarity idioms, and the frequency of the last words in low familiarity (M = 88 words per million) idioms was matched with the frequency of the last words in high familiarity idioms (M = 113 words per million).

The tropes were placed within four sentence stories as a means of providing context. Sample stimuli can be found in Appendix B. The idiom passages were modified from previous research (Miller, 2014), reducing the sentence count from five to four sentences in order to match the metaphor condition. In order to reduce the sentence count, the second sentence from the previous passages was removed. The second sentence was chosen because it acted as a filler

sentence, linking the important story points in the first and third sentence. Therefore, the story was reduced to two context sentences that strongly supported the figurative meaning; the same structure used in the metaphor passages. All contexts were considered strong contexts based on previous research (Miller, 2014). That is to say, they support the intended figurative meaning of the figurative phrase. In all stories, the figurative trope was used appropriately. Each story included four sentences. The first two sentences clearly supported the figurative meaning of the trope. In the case of the idioms, the passages had two sentences that focused on the figurative meaning of the idiom rather than any possible literal interpretations. This differs from the metaphor condition in that the sentences were not priming particular words of the idiom. This was due to the noncompositional nature of the idioms, and so an alternative approach to providing contextual support was needed. In the example seen in Appendix B, for the idiom *have a ball*, this is done by establishing that a party is being held with the subject's closest friends, supporting the notion that they will enjoy themselves very much.

Forty filler passages were included that contained no figurative language so that readers did not become aware of the consistent figurative phrase (metaphor or idiom) that appeared in the third sentence of the passages. These filler passages have the same four sentence structure of the figurative passages.

Half of the passages (40 total) were followed by a multiple-choice question to ensure participants were reading for comprehension.

The breakdown of the stimuli used based on trope condition for Experiment 2 is as follows:

Metaphor Condition

- 20 passages with familiar metaphors

- 20 passages with unfamiliar metaphors
- 40 fillers passages with no figurative language

Idiom Condition

- 20 passages with familiar idioms
- 20 passages with unfamiliar idioms
- 40 filler passages with no figurative language

Procedure

The same procedure was followed in Experiment 2 as was used in Experiment 1. Participants were randomly assigned to one of the two trope conditions (metaphor or idiom), then read 80 passages using the same calibration protocol as described above. Participants also completed the LHQ and vocabulary test components after completing the reading portion of the experiment, and were debriefed after finishing the experiment. For the metaphor condition, the average vocabulary score was 16.7 and 14 of the 20 participants were bilingual. For the idiom condition, the average vocabulary score was 16.1 and 14 of the 20 participants were bilingual.

Results

Explanation of Analyses

Analyses were conducted using subject means (F1) and item means (F2). To simplify presentation, complete statistics for each analysis are presented in Tables 11-14. Within the text, I simply refer to comparisons as significant ($p < .05$) or non-significant, with the notation *p1* used for reporting significance levels for the subject analyses, and *p2* used for reporting the item analyses. Means from the subject analyses are presented in the text when describing the results. All subject and item means for Experiment 1 can be found in Tables 1-5. Subject and item means for Experiment 2 can be found in Tables 6-10. The measures reported here are total reading time, total fixation count, regression in count, regression out count, and regression loop count (phrase level only, as a regression loop on an individual word is just a refixation). Total reading time was measured as the sum of all fixations (in ms) on the word or phrase. Fixation counts were measured as an average. For example, if the average number fixations on first words in metaphors was 2.0, that would mean that there was an average of two fixations on first words for each metaphor. Regressions were measured as summed counts. For example, if there was an average number of 5.0 regressions to context for familiar metaphors, that means there were five regressions to the context done in the possible 20 familiar metaphors. See Appendix A for definitions of these measures.

Subjects were removed if they did not complete the experiment or did not have data for every condition in which they participated. Raw eye movement data were cleaned using a three-step process in which fixations less than 80 ms and more than 1200 ms were removed from the data set. In addition, fixations less than 40 ms that were adjacent to one another were combined in order to adjust for fixations that were broken up by blinks or other interruptions. A trial was

removed if the participant did not complete the passage. Typically, this occurred when a participant felt the need to take a break during the experiment. Less than 5% of the trials were removed, as well as less than 5% of subjects.

For each experiment, I performed phrase-level analyses (the entire trope) followed by word-level analyses (data for each word individually). Because the tropes use a different number of words, with metaphors always having four words, similes always have five words, and idioms varying, the data for the word-level analyses were categorized as either the First Word, the Middle Words, or the Last Word. The Middle Words are all words that are not either the first or last word of the phrase. For metaphors, the Middle Words are always "is a". For similes, the Middle Words are always "is like a". For idioms, the Middle Words vary depending on the phrase. The data used for the Middle Words are the average of the words, not the sum of the words, and so the means will not have exaggerated reading times or fixation counts compared to the First and Last Words.

The phrase analyses for reading time and fixation count are redundant in terms of patterns of significance to the analyses presented at the word level, but they are presented anyway to illustrate the total amount of processing time and fixations required to read entire phrases. For example, because total time is the sum of the time on first, middle, and last words, the pattern of results will be same for phrase- and word-level analyses, but the numbers for total time will be three times larger than the numbers for word-level reading times. Three analyses unique to the phrase-level are also presented. The first two are regressions out of the figurative phrase and into the context, and regressions out of the final sentence of the passage back into the figurative phrase. These measures reflect the need for participants to use contextual clues to comprehend the phrase and assimilate its meaning into the story. The third unique phrase-level analysis is the

regression loop, which can only occur at the phrase level.

For each measure at the phrase-level, the by-subject analysis is a 2 (Familiarity: Familiar, Unfamiliar) X 2 (Trope: Metaphor, Simile) mixed ANOVA, with familiarity being the within-subjects factor and trope being the between-subjects factor. The by-items analysis is a 2 (Familiarity: Familiar, Unfamiliar) X 2 (Trope: Metaphor, Simile) mixed ANOVA, with both familiarity and trope being between-items factors.

For each word level analysis, for the subject means a 2 (Familiarity: Familiar, Unfamiliar) X 2 (Trope: Metaphor, Simile) X 3 (Word: First, Middle, Last) mixed ANOVA was run, with familiarity and word being within-subjects factors, and trope being between-subjects. For the item means a 2 (Familiarity: Familiar, Unfamiliar) X 2 (Trope: Metaphor, Simile) X 3 (Word: First, Middle, Last) mixed ANOVA was run, with word being a within-item factor, and familiarity and trope being between-items factors. The Word X Trope interaction is presented here, but follow-ups were not done in favor of focusing on the three way, Familiarity X Trope X Word interaction. Similarly, follow-ups for Word X Trope interaction were not done in favor of the three-way interaction. This interaction was most crucial to answering the hypotheses laid out and providing support for the Figurative Funnel. Because of this, planned comparisons of the three-way interaction are always presented. While a Bonferroni correction is typical when performing multiple comparisons on a single dataset, this was not done as the research is largely exploratory and thus I am looking for initial effect.

Experiment 1

Reading Time Analyses

Total Reading Time for the Phrase. The main effect of familiarity was significant by subject ($p1 < .01$), but not by item ($p2 = ns$). The reading time means for familiar tropes (M =

978 ms) were 90 ms faster than unfamiliar tropes (M = 1068 ms). The main effect of trope was not significant by subject (*p1 = ns*), but was significant by item (*p2* < .05). The reading time means for metaphors (M = 1023 ms) were 69 ms longer than similes (M = 954 ms), although as noted this difference was only significant by items. Means are presented in Table 1 and the pattern of means is illustrated in Figure 2.

Total Reading Time for Individual Words. The main effect of word was significant by subject and by item (*p1* < .01, *p2* < .01). Pairwise comparisons showed that more time was spent on First (M = 355 ms) and Last (M = 357 ms) than Middle words (M = 277) (*p1's* < .05), and there was no difference between First and Last words (*p1 = ns*). The Word X Familiarity interaction was significant by subject (*p1* < .05) but not by item (*p2 = ns*). The Word X Trope interaction was not significant by subject (*p1 = ns*) or by item (*p2 = ns*). The Familiarity X Trope interaction was significant by subject (*p1* < .05) but not by item (*p2 = ns*). Follow up simple effects tests showed that words in familiar metaphors (M = 326 ms) were read 20 ms faster (*p1* < .05) than words in unfamiliar metaphors (M = 356 ms). There was no difference in reading time for words in familiar similes (M = 317 ms) and unfamiliar similes (M = 319 ms) (*p1 = ns*). Means are presented in Table 2 and the pattern of means is illustrated in Figure 3.

The Word X Familiarity X Trope interaction was not significant by subject or by item (*p1 = ns*, *p2 = ns*). Follow up simple effects tests were conducted because the interaction involves planned comparisons between conditions needed to test the hypotheses. For the metaphor condition, there was no significant difference (*p1 = ns*) in reading times on first words between familiar metaphors (M = 379 ms) and unfamiliar metaphors (M = 380 ms). There was a significant difference (*p1* < .01) in reading times on middle words between familiar metaphors (M = 257 ms) and unfamiliar metaphors (M = 301 ms), such that middle words in familiar

metaphors were 44 ms faster than in unfamiliar metaphors. There was a significant difference ($p1 < .01$) in reading time on last words between familiar metaphors (M = 342 ms) and unfamiliar metaphors (M = 387 ms), such that last words in familiar metaphors were 45 ms faster than in unfamiliar metaphors. These comparisons reflect longer processing time on words at the end of unfamiliar metaphors than familiar metaphors. For words within familiar metaphors, there was a significant difference ($p1 < .01$) in reading time between first words (M = 379 ms) and middle words (M = 257 ms), such that first words were read 122 ms slower than middle words. There was a significant difference ($p1 < .01$) in reading time between middle words (M = 257 ms) and last words (M = 342 ms) in familiar metaphors, such that last words were read 85 ms slower than middle words. There was a significant difference ($p1 < .05$) in reading time between first words (M = 379 ms) and last words (M = 342 ms) in familiar metaphors, such that first words were read 37 ms slower than last words. For words within unfamiliar metaphors, there was a significant difference ($p1 < .05$) in reading time between first words (M = 380 ms) and middle words (M = 301 ms), such that first words were read 79 ms slower than middle words. There was a significant difference ($p1 < .01$) in reading time between middle words (M = 301 ms) and last words (M = 387 ms) in unfamiliar metaphors, such that last words were read 86 ms slower than middle words. There was no difference ($p1 = ns$) in reading time between first words (M = 380 ms) and last words (M = 387 ms) in unfamiliar metaphors. These comparisons also reflect increased processing time on final words for unfamiliar metaphors.

For the simile condition, there was no difference ($p1 = ns$) in reading times on first words between familiar similes (M = 333 ms) and unfamiliar similes (M = 327 ms). There was no difference ($p1 = ns$) in reading time on middle words between familiar similes (M = 276 ms) and unfamiliar similes (M = 276 ms). There was no difference ($p1 = ns$) in reading time on last words

between familiar similes (M = 343 ms) and unfamiliar similes (M = 354 ms). These comparisons reflect no processing time differences between familiar and unfamiliar similes. For words within familiar similes, there was a significant difference ($p1 < .01$) in reading time between first words (M = 333 ms) and middle words (M = 276 ms) such that first words were read 57 ms slower than middle words. There was a significant difference ($p1 < .01$) in reading time between middle words (M = 276 ms) and last words (M = 343 ms) in familiar similes, such that last words were read 67 ms slower than middle words. There was no difference ($p1 = ns$) in reading time between first words (M = 333 ms) and last words (M = 343 ms) in familiar similes. For words within unfamiliar similes, there was a significant difference ($p1 < .01$) in reading time between first words (M = 327 ms) and middle words (M = 276 ms), such that first words were read 51 ms slower than middle words. There was a significant difference ($p1 < .01$) in reading time between middle words (M = 276 ms) and last words (M = 354 ms) such that last words were read 78 ms slower than middle words in unfamiliar similes. There was a significant difference ($p1 < .05$) in reading time between first words (M = 327 ms) and last words (M = 354 ms) in unfamiliar similes, such that last words were read 27 ms slower than first words.

Fixation Count Analyses

Fixation Count for the Phrase. The main effect of familiarity was not significant by subject or item ($p1 = ns$, $p2 = ns$). This reflects no difference in the number of fixations on familiar tropes (M = 3.3) and unfamiliar tropes (M = 3.4). There was no main effect of trope by subject or item ($p1 = ns$, $p2 = ns$), such that there was no difference in the number of fixations on metaphors (M = 3.4) and similes (M = 3.3). Means are presented in Table 1 and the pattern of means is illustrated in Figure 4.

Fixation Count for the Words. The main effect of word was significant by subject and by

item (*p1* < .01, *p2* < .01). Pairwise comparisons showed that more fixations were made on First (M = 1.4) and Last words (M = 1.3) than Middle words (M = .64) (*p1's* < .05), and that there were more fixations on First words than Last words. The Word X Familiarity interaction was not significant by subject or by item (*p1 = ns*, *p2 = ns*). The Word X Trope interaction was marginally significant by subject (*p1* = .10) but not by item (*p2 = ns*). The Familiarity X Trope interaction was significant by subject (*p1* < .05) but not by item (*p2 = ns*). Follow up simple effects tests showed that there was no difference in fixation count on words in familiar metaphors (M = 1.1) than words in unfamiliar metaphors (M = 1.2) (*p1 = ns*). There was no difference in fixation count on words in familiar similes (M = 1.1) and unfamiliar similes (1.1) (*p1 = ns*). Means are presented in Table 3 and the pattern of means is illustrated in Figure 5.

The Word X Familiarity X Trope interaction was not significant by subject or by item (*p1 = ns*, *p2 = ns*). Follow up simple effects tests were conducted because the interaction involves planned comparisons needed to test the hypotheses. Starting with the metaphor condition, there was no difference (*p1 = ns*) in fixation count on first words between familiar metaphors (M = 1.5) and unfamiliar metaphors (M = 1.6). There was no significant difference (*p1 = ns*) in fixation count on middle words between familiar metaphors (M = .57) and unfamiliar metaphors (M = .66). There was no significant difference (*p1 = ns*) in fixation count on last words between familiar metaphors (M = 1.2) and unfamiliar metaphors (M = 1.3). In essence, there were no differences in fixation counts as a function of familiarity. For words within familiar metaphors, there was a significant difference (*p1* < .01) in fixation count between first words (M = 1.5) and middle words (M = .57), such that first words had 0.93 more fixations than middle words. There was a significant difference in fixation count between middle words (M = .57) and last words (M = 1.2) in familiar metaphors, such that last words had 0.63 more fixations than middle words.

There was a significant ($p1 < .05$) difference in fixation count between first words (M = 1.5) and last words (M = 1.2) in familiar metaphors, such that first words had 0.3 more fixations than last words. For words within unfamiliar metaphors, there was a significant difference ($p1 < .01$) in fixation count between first words (M = 1.6) and middle words (M = .66), such that first words had 0.94 more fixations than middle words. There was a significant difference ($p1 < .01$) in fixation count between middle words (M = .66) and last words (M = 1.3) in unfamiliar metaphors, such that last words had 0.64 more fixations than middle words. There was a significant difference ($p1 < .01$) in fixation count between first words (M = 1.6) and last words (M = 1.3) in unfamiliar metaphors, such that first words had 0.3 more fixations than last words. The larger number of fixations on first words could reflects regressions to these words. This will be discussed in more detail after presenting the regression results.

For the simile condition, there was no significant difference ($p1 = ns$) in fixation count on first words between familiar similes (M = 1.3) and unfamiliar similes (M = 1.3). There was no significant difference ($p1 = ns$) in fixation count on middle words between familiar similes (M = .66) and unfamiliar similes (M = .66). There was no significant difference ($p1 = ns$) in fixation count on last words between familiar similes (M = 1.3) and unfamiliar similes (M = 1.3). For words within familiar similes, there was a significant difference ($p1 < .01$) in fixation count between first words (M = 1.3) and middle words (M = .66), such that first words had 0.64 more fixations than middle words. There was a significant difference ($p1 < .01$) in fixation count between middle words (M = .66) and last words (M = 1.3) in familiar similes, such that last words had 0.64 more fixations than middle words. There was no difference ($p1 = ns$) in fixation count between first words (M = 1.3) and last words (M = 1.3) in familiar similes. For words within unfamiliar similes, there was a significant difference ($p1 < .01$) in fixation count between

first words (M = 1.3) and middle words (M = .66), such that first words had 0.64 more fixations than middle words. There was a significant difference ($p1 < .01$) in fixation count between middle words (M = .66) and last words (M = 1.3) in unfamiliar similes, such that last words had 0.64 more fixations than middle words. There was no difference ($p1 = ns$) in fixation count between first words (M = 1.3) and last words (M = 1.3) in unfamiliar similes. In essence, there were no reliable differences in fixation counts as a function of familiarity or between words within similes.

Regression Analyses

Regressions from the Trope to the Context. The main effect of familiarity was not significant by subject or by item ($p1 = ns$, $p2 = ns$). The main effect of trope was not significant by subject or by item ($p1 = ns$, $p2 = ns$). The Familiarity X Trope interaction was not significant by subject or by item ($p1 = ns$, $p2 = ns$). Means are presented in Table 1 and the pattern of means is illustrated in Figure 6.

Regressions from the Final Sentence to the Trope. The main effect of familiarity was not significant by subject or by item ($p1 = ns$, $p2 = ns$). The main effect of trope was not significant by subject or by item ($p1 = ns$, $p2 = ns$). The Familiarity X Trope interaction was not significant by subject or by item ($p1 = ns$, $p2 = ns$). Means are presented in Table 1 and the pattern of means is illustrated in Figure 7.

Regressions from Last Word Region to First Word Region. The main effect of familiarity was not significant by subject or by item ($p1 = ns$, $p2 = ns$). The main effect of trope was not significant by subject or by item ($p1 = ns$, $p2 = ns$). The Familiarity X Trope interaction was not significant by subject or by item ($p1 = ns$, $p2 = ns$). Means are presented in Table 1 and the pattern of means is illustrated in Figure 8.

Regression Loops Count. The main effect of familiarity was not significant by subject or by item (*p1 = ns*, *p2 = ns*). The main effect of trope was not significant by subject or by item (*p1 = ns*, *p2 = ns*). The Familiarity X Trope interaction was not significant by subject or by item (*p1 = ns*, *p2 = ns*). Means are presented in Table 1 and the pattern of means is illustrated in Figure 9.

Regressions In Count to the Words. The main effect of word was significant by subject and by item (*p1* < .01, *p2* < .01). Pairwise comparisons showed that there were more regressions into First (M = 4.9) and Middle words (M = 7.0) than Last words (M = 1.1) (*p1's* < .05), and there were more regressions into Middle words than First words (*p1* < .05). There was no Word X Familiarity interaction by subject (*p1 = ns*) but it was significant by item (*p2* < .01). The Word X Trope interaction was significant by subject (*p1* < .01) but not by item (*p2 = ns*). The Familiarity X Trope interaction was marginally significant by subject (*p1* = .08) but not by item (*p2 = ns*). Follow up simple effects tests showed that there was no difference in regressions in to words between familiar metaphors (M = 4.0) and unfamiliar metaphors (M = 4.6) (*p1 = ns*). There was no difference in regressions in to words between familiar similes (M = 4.6) and unfamiliar similes (M = 4.3). Means are presented in Table 4 and the pattern of means is illustrated in Figure 10.

The Word X Familiarity X Trope interaction was not significant by subject or by item (*p1 = ns*, *p2 = ns*). Follow up simple effects tests were conducted because the interaction involves planned comparisons between conditions needed to test the hypotheses. Starting with the metaphor condition, there was no difference (*p1 = ns*) in regression in count on first words between familiar metaphors (M = 5.6) and unfamiliar metaphors (M = 5.9). There was no difference (*p1 = ns*) in regression in count on middle words between familiar metaphors (M = 5.5) and unfamiliar metaphors (M = 6.4). There was no difference (*p1 = ns*) in regression count

on last words between familiar metaphors (M = .90) and unfamiliar metaphors (M = 1.4). For words within familiar metaphors, there was no difference (*p1 = ns*) in regression in count between first words (M = 5.6) and middle words (M = 5.5). There was a significant difference (*p1* < .01) in regression in count between middle words (M = 5.5) and last words (M = .90) in familiar metaphors, such that middle words had 4.6 more regressions in than last words. There was a significant difference (*p1* < .01) in regression in count between first words (M = 5.6) and last words (M = .90) in familiar metaphors, such that first words had 4.7 more regressions in than last words. For words within unfamiliar metaphors, there was no difference (*p1 = ns*) in regression in count between first words (M = 5.9) and middle words (M = 6.4). There was a significant difference (*p1* < .01) in regression in count between middle words (M = 6.4) and last words (M = 1.4) in unfamiliar metaphors, such that middle words had 5.0 more regressions in than last words. There was a significant difference (*p1* < .01) in regression in count between first words (M = 5.9) and last words (M = 1.4) in unfamiliar metaphors, such that first words had 4.5 more regressions in than last words. The larger number of regressions for first words might explain the larger total number of fixations for first words noted above.

For the simile condition, there was no significant difference (*p1 = ns*) in regression in count on first words between familiar similes (M = 4.6) and unfamiliar similes (M = 3.7). There was no significant difference (*p1 = ns*) in regression in count on middle words between familiar similes (M = 7.8) and unfamiliar similes (M = 8.5). There was no significant difference (*p1 = ns*) in regression in count on last words between familiar similes (M = 1.5) and unfamiliar similes (M = .85). For words within familiar similes, there was a significant difference (*p1* < .01) in regression in count between first words (M = 4.6) and middle words (M = 7.8), such that middle words had 3.2 more regressions in than first words. There was a significant difference (*p1* < .01)

in regression in count between middle words (M = 7.8) and last words (M = 1.5) in familiar similes, such that middle words had 6.3 more regressions in than last words. There was a significant difference (*p1* < .01) in regression in count between first words (M = 4.6) and last words (M = 1.5) in familiar similes, such that first words had 3.1 more regressions in than last words. For words within unfamiliar similes, there was a significant difference (*p1* < .01) in regression in count between first words (M = 3.7) and middle words (M = 8.5), such that middle words had 4.8 more regressions in than first words. There was a significant difference (*p1* < .01) in regression in count between middle words (M = 8.5) and last words (M = .85) in unfamiliar similes, such that middle words had 7.65 more regressions in than last words. There was a significant difference (*p1* < .01) in regression in count between first words (M = 3.7) and last words (M = .85) in unfamiliar similes, such that first words had 2.85 more regressions in than last words. The larger number of regressions to first words than other words is one of the few reliable differences found for similes.

Regressions Out Count of the Words. The main effect of word was not significant by subject (*p1 = ns*) but was significant by item (*p2* < .05). The Word X Familiarity interaction was not significant by subject (*p1 = ns*) but was by item (*p2* < .05). The Word X Trope interaction was not significant by subject or by item (*p1 = ns*, *p2 = ns*). There was no Familiarity X Trope interaction by subject or by item (*p1 = ns*, *p2 = ns*). Means are presented in Table 5 and the pattern of means is illustrated in Figure 11.

The Word X Familiarity X Trope interaction was not significant by subject or by item (*p1 = ns*, *p2 = ns*). Follow up simple effects tests were conducted because the interaction involves planned comparisons between conditions needed to test the hypotheses. Starting with the metaphor condition, there was no difference (*p1 = ns*) in regression out count of first words

between familiar metaphors (M = 4.0) and unfamiliar metaphors (M = 3.9). There was no difference (*p1 = ns*) in regression out count of middle words between familiar metaphors (M = 5.1) and unfamiliar metaphors (M = 5.7). There was no difference (*p1 = ns*) in regression out count of last words between familiar metaphors (M = 6.2) and unfamiliar metaphors (M = 5.9). For words within familiar metaphors, there was a significant difference (*p1* < .05) in regression out count of first words (M = 4.0) and middle words (M = 5.1), such that middle words had 1.1 more regressions out than first words. There was a significant difference (*p1 = ns*) in regression out for middle words (M = 5.1) and last words (M = 6.2) in familiar metaphors, such that last words had 1.1 more regressions out than middle words. There was a significant difference (*p1* < .01) in regression out count between first words (M = 4.0) and last words (M = 6.2) in familiar metaphors, such that last words had 2.2 more regressions out than first words. For words within unfamiliar metaphors, there was a significant difference (*p1* < .01) in regression out count between first words (M = 3.9) and middle words (M = 5.7), such that middle words had 1.8 more regressions out than first words. There was no significant difference (*p1 = ns*) in regression out count between middle words (M = 5.7) and last words (M = 5.9) in unfamiliar metaphors. There was a significant difference (*p1* < .01) in regression out count between first words (M = 3.9) and last words (M = 4.6) in unfamiliar metaphors, such that last words had 0.7 more regressions out than first words. The larger number of regressions out of the last words is consistent with the idea that readers do not need to make regressions until reaching the end of a phrase, which is when they are most likely to realize they are reading a metaphor.

For the simile condition, there was no difference (*p1 = ns*) in regression out count of first words between familiar similes (M = 4.6) and unfamiliar similes (M = 5.2). There was no difference (*p1 = ns*) in regression out count of middle words between familiar similes (M = 6.0)

and unfamiliar (M = 5.9). There was no significant difference (*p1 = ns*) in regression out count of last words between familiar similes (M = 5.1) and unfamiliar similes (M = 4.5). For words within familiar similes, there was a significant difference (*p1* < .05) in regression out count between first words (M = 4.6) and middle words (M = 6.0), such that middle words had 1.4 more regressions out than first words. There was no significant difference (*p1 = ns*) in regression out count between middle words (M = 6.0) and last words (M = 5.1) in familiar similes. There was no significant difference (*p1 = ns*) in regression out count between first words (M = 4.6) and last words (M = 5.1) in familiar similes. For words within unfamiliar similes, there was no significant difference (*p1 = ns*) in regression out count between first words (M = 5.2) and middle words (M = 5.9). There was a significant difference (*p1* < .05) in regression out count between middle words (M = 5.9) and last words (M = 4.5) in unfamiliar similes, such that middle words had 1.4 more regressions out than last words. There was no significant difference (*p1 = ns*) in regression out count between first words (M = 5.2) and last words (M = 4.5) in unfamiliar similes.

In summary, reliable familiarity effects were found for total time and fixation count at the word level, as well as regressions to context and from the last sentence. Middle and last words were read slower than first words consistently. However, no metaphor familiarity effects were found at the phrase level for fixation count as well as regression loops. For similes, there were no differences found as a function of familiarity. In addition, there were equal reading times and fixation counts between first and last words in similes.

Experiment 2

Reading Time Analyses

Total Reading Time for the Phrase. The main effect of familiarity was significant by

subject (*p1* < .01) and marginally significant by item (*p2* = .10). The mean reading times for familiar tropes (M = 973 ms) were 106 ms faster than unfamiliar tropes (M = 1079 ms). The main effect of trope type was significant by subject (*p1* < .01) and by item (*p2* < .01), such that metaphors (M = 1123 ms) were read 193 ms slower than idioms (M = 930 ms). Means are presented in Table 6 and the pattern of means is illustrated in Figure 12.

Total Reading Time for the Words. The main effect of word was significant by subject and by item (*p1* < .01, *p2* < .01). Pairwise comparisons showed that more time was spent on First (M = 370 ms) and Last words (M = 355 ms) than Middle words (M = 301) (*p1's* < .05), and there was no difference between First and Last words (*p1* = *ns*). There was no Word X Familiarity interaction by subject or by item (*p1* = *ns*, *p2* = *ns*). The Word X Trope interaction was significant by subject and by item (*p1* < .01, *p2* < .01). There was no Familiarity X Trope interaction by subject or by item (*p1* = *ns*, *p2* = *ns*). Means are presented in Table 7 and the pattern of means is illustrated in Figure 13.

The Word X Familiarity X Trope interaction was not significant by subject or by item (*p1* = *ns*, *p2* = *ns*). Follow up simple effects tests were conducted because the interaction involves planned comparisons needed to test the hypotheses. Starting with the metaphor condition, there was a significant difference (*p1* < .05) in reading times on first words between familiar metaphors (M = 397 ms) and unfamiliar metaphors (M = 423 ms), such that first words in familiar metaphors were read 26 ms faster than in unfamiliar metaphors. There was no significant difference (*p1* = *ns*) in reading times on middle words between familiar metaphors (M = 300 ms) and unfamiliar metaphors (M = 320 ms). There was a significant difference (*p1* <) in reading time on last words between familiar metaphors (M = 390 ms) and unfamiliar metaphors (M = 423 ms), such that last words in unfamiliar metaphors were read 33 ms slower than in

familiar metaphors. For words within familiar metaphors, there was a significant difference ($p1 < .01$) in reading time between first words (M = 397 ms) and middle words (M = 300 ms), such that first words were read 97 ms slower than middle words. There was a significant difference ($p1 < .01$) in reading time between middle words (M = 300 ms) and last words (M = 390 ms) in familiar metaphors, such that last words were read 90 ms slower than middle words. There was no difference ($p1 = ns$) in reading time between first words (M = 397 ms) and last words (M = 390 ms) in familiar metaphors. For words within unfamiliar metaphors, there was a significant difference ($p1 < .01$) in reading time between first words (M = 423 ms) and middle words (M = 320 ms), such that first words were read 103 ms slower than middle words. There was a significant difference ($p1 < .01$) in reading time between middle words (M = 320 ms) and last words (M = 416 ms) in unfamiliar metaphors, such that last words were read 96 ms slower than middle words. There was no difference ($p1 = ns$) in reading time between first words (M = 423 ms) and last words (M = 416 ms) in unfamiliar metaphors.

For the idiom condition, there was a significant difference ($p1 < .05$) in reading time on first words between familiar idioms (M = 313 ms) and unfamiliar idioms (M = 347 ms), such that first words in familiar idioms were 34 ms faster than in unfamiliar idioms. There was a significant difference ($p1 < .05$) in reading time on middle words between familiar idioms (M = 273 ms) and unfamiliar idioms (M = 313 ms), such that middle words in familiar idioms were 40 ms faster than in unfamiliar idioms. There was a significant difference ($p1 < .05$) in reading time on last words between familiar idioms (M = 273 ms) and unfamiliar idioms (M = 341 ms), such that last words in familiar idioms were 68 ms faster than in unfamiliar idioms. For words within familiar idioms, there was a significant difference ($p1 < .01$) in reading time between first words (M = 313 ms) and middle words (M = 273 ms), such that first words were read 40 ms slower

than middle words. There was no difference (*p1 = ns*) in reading time between middle words (M = 273 ms) and last words (M = 273 ms) in familiar idioms. There was a significant difference (*p1* < .01) in reading time between first words (M = 313 ms) and last words (M = 273 ms) in familiar idioms, such that first words were 40 ms faster than last words. For words within unfamiliar idioms, there was a significant difference (*p1* < .05) in reading time between first words (M = 347 ms) and middle words (M = 313 ms), such that first words were read 34 ms slower than middle words. There was a significant difference (*p1* < .05) in reading time between middle words (M = 313 ms) and last words (341 ms) in unfamiliar idioms, such that last words were read 28 ms slower than middle words. There was no difference (*p1 = ns*) in reading time between first words (M = 347 ms) and last words (M = 341 ms) in unfamiliar idioms.

Fixation Count Analyses

Fixation Count for the Phrase. The main effect of familiarity was significant by subject (*p1* < .01), and marginally significant by item (*p2* = .06). The mean number of fixations for familiar tropes (M = 3.3) was 0.3 less than unfamiliar tropes (M = 3.6). The main effect of trope was not significant by subject (*p1 = ns*) or by item (*p2 = ns*). The number of fixations on metaphors (M = 3.8) did not reliably differ from idioms (M = 3.1). Means are presented in Table 6 and the pattern of means is illustrated in Figure 14.

Fixation Count for the Words. The main effect of word was significant by subject and by item (*p1* < .01, *p2* < .01). Pairwise comparisons showed that more fixations were made on First (M = 1.4) and Last words (M = 1.3) than on Middle words (M = .76) (*p1's* < .05), and there were no differences between First and Last words (*p1 = ns*). There was a significant Word X Familiarity interaction by subject (*p1* < .05), but not by item (*p2 = ns*). The Word X Trope interaction was significant by subject and by item (*p1* < .01, *p2* < .01). There was no Familiarity

X Trope interaction by subject or by item (*p1 = ns*, *p2 = ns*). Means are presented in Table 8 and the pattern of means is illustrated in Figure 15.

The Word X Familiarity X Trope interaction was marginally significant by subject (*p1* = .11) but not by item (*p2 = ns*). Follow up simple effects tests were conducted because the interaction involves planned comparisons needed to test the hypotheses. Starting with the metaphor condition, there was no difference (*p1 = ns*) in fixation count on first words between familiar metaphors (M = 1.7) and unfamiliar metaphors (M = 1.7). There was a significant difference (*p1* < .05) in fixation count on middle words between familiar metaphors (M = .64) and unfamiliar metaphors (M = .78), such that middle words in familiar metaphors had 0.14 fewer fixations than in unfamiliar metaphors. There was no difference (*p1 = ns*) in fixation count on last words between familiar metaphors (M = 1.4) and unfamiliar metaphors (M = 1.5). For words within familiar metaphors, there was a significant difference (*p1* < .01) in fixation count between first words (M = 1.7) and middle words (M = .64), such that first words had 1.06 more fixations than middle words. There was a significant difference (*p1* < .01) in fixation count between middle words (M = .64) and last words (M = 1.4) in familiar metaphors, such that last words had 0.76 more fixations than middle words. There was a significant difference (*p1* < .05) in fixation count on first words (M = 1.7) and last words (M = 1.4) in familiar metaphors, such that last words had 0.3 more fixations than first words. For words within unfamiliar metaphors, there was a significant difference (*p1* < .01) in fixation count between first words (M = 1.7) and middle words (M = .78), such that first words had 0.92 more fixations than middle words. There was a significant difference (*p1* < .01) in fixation count between middle words (M = .78) and last words (M = 1.5) in unfamiliar metaphors, such that last words had 0.72 more fixations than middle words. There was a significant difference (*p1* < .05) in fixation count on first words (M =

1.7) and last words (M = 1.5) in unfamiliar metaphors, such that last first words had 0.3 more fixations than last words.

For the idiom condition, there were no differences (*p1* = *ns*) in fixation count on first words between familiar idioms (M = 1.2) and unfamiliar idioms (M = 1.2). There was a significant difference (*p1* < .01) in fixation count on middle words between familiar idioms (M = .74) and unfamiliar idioms (M = .89), such that middle words in familiar idioms had 0.15 fewer fixations than unfamiliar idioms. There was a significant difference (*p1* < .05) in fixation count on last words between familiar idioms (M = .96) and unfamiliar idioms (M = 1.2), such that last words in familiar idioms had 0.24 fewer fixations than unfamiliar idioms. For words within familiar idioms, there was a significant difference (*p1* < .01) in fixation count between first words (M = 1.2) and middle words (M = .74), such that first words had 0.46 more fixations than middle words. There was a significant difference (*p1* < .01) in fixation count between middle words (M = .74) and last words (M = .96) in familiar idioms, such that last words had 0.22 more fixations than middle words. There was a significant difference (*p1* < .01) in fixation count between first words (M = 1.2) and last words (M = .96) in familiar idioms, such that first words had 0.24 more fixations than last words. For words within unfamiliar idioms, there was a significant difference (*p1* < .01) in fixation count between first words (M = 1.2) and middle words (M = .89), such that first words had 0.31 more fixations than middle words. There was a significant difference (*p1* < .01) in fixation count between middle words (M = .89) and last words (M = 1.2) in unfamiliar idioms, such that last words had 0.31 more fixations than middle words. There was no difference (*p1* = *ns*) in fixation count between first words (M = 1.2) and last words (M = 1.2) in unfamiliar idioms.

Regression Analyses

Regressions from the Trope to the Context. The main effect of familiarity was not significant by subject or by item (*p1 = ns*, *p2 = ns*). There was no difference in regressions out of the trope to the context in familiar tropes (M = 6.1) and unfamiliar tropes (M = 6.1). The main effect of trope was significant by subject and by item (*p1* < .01, *p2* < .01) such that there were more regressions from metaphors to the context (M = 8.7) than idioms to the context (M = 3.6). The Familiarity X Trope interaction was not significant by subject or by item (*p1 = ns*, *p2 = ns*). Means are presented in Table 6 and the pattern of means is illustrated in Figure 16.

Regressions from the Final Sentence to the Trope. The main effect of familiarity was not significant by subject or by item (*p1 = ns*, *p2 = ns*). There was no difference in regressions from the last sentence to the trope in familiar tropes (M = 2.6) and unfamiliar tropes (M = 2.2). The main effect of trope was not significant by subject (*p1 = ns*), but was marginally significant by item (*p2* = .08). There were fewer regressions from the last sentence to the trope for metaphors (M = 2.0) than idioms (M = 2.8), but this difference was only marginal by item. The Familiarity X Trope interaction was not significant by subject or by item (*p1 = ns*, *p2 = ns*). Means are presented in Table 6 and the pattern of means is illustrated in Figure 17.

Regressions from Last Word Region to First Word Region. The main effect of familiarity was marginally significant by subject (*p1* = .10) but not by item (*p2 = ns*), such that there were more regressions from the last word region to the first word region in unfamiliar tropes (M = 4.0) than in familiar tropes (M = 3.4). The main effect of trope was significant by subject and by item (*p1* < .01, *p2* < .01), such that there were more regressions from the last word region to the first word region in metaphors (M = 4.3) than in idioms (M = 3.2). The Familiarity X Trope interaction was not significant by subject or by item (*p1 = ns*, *p2 = ns*). Means are presented in Table 6 and the pattern of means is illustrated in Figure 18.

Regression Loops Count. The main effect of familiarity was not significant by subject or by item (*p1 = ns*, *p2 = ns*). The main effect of trope was significant by subject and by item (*p1* < .01, *p2* < .01), such that there were more regression loops in metaphors (M = 3.0) than in idioms (M = 2.0). The Familiarity X Trope interaction was not significant by subject or by item (*p1 = ns*, *p2 = ns*). Means are presented in Table 6 and the pattern of means is illustrated in Figure 19.

Regressions In Count to the Words. The main effect of word was significant by subject and by item (*p1* < .01, *p2* < .01). Pairwise comparisons showed that there were more regressions into First (M = 5.6) and Middle words (M = 8.9) than Last words (M = 2.0) (*p1's* < .05), and there were marginally more regressions into Middle words than First words (*p1* = .12). There was no Word X Familiarity interaction by subject or by item (*p1 = ns*, *p2 = ns*). The Word X Trope interaction was significant by subject and by item (*p1* < .01, *p2* < .01). The Familiarity X Trope interaction was marginally significant by subject (*p1* = .08). Follow up simple effects tests showed that there was no difference in regression in count for words in familiar metaphors (M = 5.0) than words in unfamiliar metaphors (M = 5.8). There was a significant difference (*p1* < .05) in regression in count between words in familiar idioms (M = 3.3) and words in unfamiliar idioms (M = 5.1), such that words in unfamiliar idioms had 1.8 more regressions in than words in familiar idioms. Means are presented in Table 9 and the pattern of means is illustrated in Figure 20.

The Word X Familiarity X Trope interaction was not significant by subject or by item (*p1 = ns*, *p2 = ns*). Follow up simple effects tests were conducted because the interaction involves planned comparisons needed to test the hypotheses. Starting with the metaphor condition, there was a significant difference (*p1* < .05) in regression in count of first words between familiar metaphors (M = 6.3) and unfamiliar metaphors (M = 7.4), such that first words in unfamiliar

metaphors had 1.1 more regressions in than unfamiliar metaphors. There was a significant difference ($p1 < .05$) in regression in count of middle words between familiar metaphors (M = 6.9) and unfamiliar metaphors (M = 8.7), such that middle words in unfamiliar metaphors had 1.8 more regressions in than middle words in familiar metaphors. There was no difference (*p1 = ns*) in regression in count of last words between familiar metaphors (M = 1.7) and unfamiliar metaphors (M = 1.4). For words within familiar metaphors, there was no significant difference (*p1 = ns*) in regression in count between first words (M = 6.3) and middle words (M = 6.9). There was a significant difference ($p1 < .01$) in regression in count between middle words (M = 6.9) and last words (M = 1.7) in familiar metaphors, such that middle words had 5.2 more regressions in than last words. There was a significant difference ($p1 < .05$) in regression in count between first words (M = 6.3) and last words (M = 1.7) in familiar metaphors, such that first words had 4.6 more regressions in than last words. For words within unfamiliar metaphors, there was a significant difference ($p1 < .05$) in regression in count between first words (M = 7.4) and middle words (M = 8.7), such that middle words had 1.3 more regressions in than first words. There was a significant difference ($p1 < .01$) in regression in count between middle words (M = 8.7) and last words (M = 1.4) in unfamiliar metaphors, such that middle words had 7.3 more regressions in than last words. There was a significant difference ($p1 < .05$) in regression in count between first words (M = 7.4) and last words (M = 1.4) in unfamiliar metaphors, such that first words had 6.0 more regressions in than last words.

For the idiom condition, there was a significant difference ($p1 < .05$) in regression in count of first words between familiar idioms (M = 3.8) and unfamiliar idioms (M = 4.9), such that first words in unfamiliar idioms had 1.1 more regressions in than familiar idioms. There was a significant difference ($p1 < .05$) in regression in count of middle words between familiar

idioms (M = 4.5) and unfamiliar idioms (M = 7.5), such that middle words in unfamiliar idioms had 3.0 more regressions in than familiar idioms. There was a significant difference ($p1 < .05$) in regression in count of last words between familiar idioms (M = 1.8) and unfamiliar idioms (M = 3.0), such that last words in unfamiliar idioms had 1.2 more regressions in than familiar idioms. For words within familiar idioms, there was no significant difference ($p1 = ns$) in regression in count between first words (M = 3.8) and middle words (M = 4.5). There was a significant difference ($p1 < .05$) in regression in count between middle words (M = 4.5) and last words (M = 1.8) in familiar idioms, such that middle words had 2.7 more regressions in than last words. There was a significant difference ($p1 < .05$) in regression in count between first words (M = 3.8) and last words (M = 1.8) in familiar idioms, such that first words had 2.0 more regressions in than last words. For words within unfamiliar idioms, there was a significant difference ($p1 < .05$) in regression in count between first words (M = 4.9) and middle words (M = 7.5), such that middle words had 2.6 more regressions in than first words. There was a significant difference ($p1 < .05$) in regression in count between middle words (M = 7.5) and last words (M = 3.0) in unfamiliar idioms, such that middle words had 4.5 more regressions in than last words. There was a significant difference ($p1 < .05$) in regression in count between first words (M = 4.9) and last words (M = 3.0) in unfamiliar idioms, such that first words had 1.9 more regressions in than last words.

Regressions Out Count of the Words. The main effect of word was significant by subject and by item ($p1 < .05$, $p2 < .05$). Pairwise comparisons showed that more regressions were made out of Last words (M = 5.6) and Middle words (M = 5.7) than first words (M = 4.0) ($p1's < .05$), and there was no difference between last and middle words ($p1 = ns$). The Word X Familiarity interaction was significant by subject ($p1 < .05$) and by item ($p2 < .05$). The Word X Trope

interaction was significant by subject ($p1 < .01$) but was marginal by item ($p2 = .06$). There was no Familiarity X Trope interaction by subject or by item ($p1 = ns$, $p2 = ns$). Means are presented in Table 10 and the pattern of means is illustrated in Figure 21.

The Word X Familiarity X Trope interaction was not significant by subject or by item ($p1 = ns$, $p2 = ns$). Follow up simple effects tests were conducted because the interaction involves planned comparisons needed to test the hypotheses. Starting with the metaphor condition, there was no significant difference ($p1 = ns$) in regression out count of first words between familiar metaphors (M = 5.6) and unfamiliar metaphors (M = 4.8). There was no significant difference ($p1 = ns$) in regression out count of middle words between familiar metaphors (M = 5.7) and unfamiliar metaphors (M = 6.8). There was no significant difference ($p1 = ns$) in regression out count of last words between familiar metaphors (M = 8.0) and unfamiliar metaphors (M = 8.3). For words within familiar metaphors, there was no significant difference ($p1 = ns$) in regression out count between first words (M = 5.6) and middle words (M = 5.7). There was a significant difference ($p1 < .05$) in regression out count between middle words (M = 5.7) and last words (M = 8.0) in familiar metaphors, such that last words had 2.3 more regressions out than middle words. There was a significant difference ($p1 < .05$) in regression out count between first words (M = 5.6) and last words (M = 8.0) in familiar metaphors, such that last words had 2.4 more regressions out than first words. For words within unfamiliar metaphors, there was a significant difference ($p1 < .05$) in regression out count between first words (M = 4.8) and middle words (M = 6.8), such that middle words had 2.0 more regressions out than first words. There was a significant difference ($p1 < .05$) in regression out count between middle words (M = 6.8) and last words (M = 8.3) in unfamiliar metaphors, such that last words had 1.5 more regressions out than middle words. There was a significant difference ($p1 < .01$) in regression out count between first

words (M = 4.8) and last words (M = 8.3) in unfamiliar metaphors, such that last words had 3.5 more regressions out than first words.

For the idiom condition, there was a significant difference ($p1 < .05$) in regression out count on first words between familiar idioms (M = 3.6) and unfamiliar idioms (M = 2.1), such that first words in familiar idioms had 1.5 more regressions out than unfamiliar idioms. There was a significant difference ($p1 < .05$) in regression out count on middle words between familiar idioms (M = 3.7) and unfamiliar idioms (M = 6.9), such that middle words in unfamiliar idioms had 3.2 more regressions out than familiar idioms. There was no significant difference ($p1 = ns$) in regression out count on last words between familiar idioms (M = 2.5) and unfamiliar idioms (M = 3.5). For words within familiar idioms, there was no significant difference ($p1 = ns$) in regression out count between first words (M = 3.6) and middle words (M = 3.7). There was a significant difference ($p1 < .05$) in regression out count between middle words (M = 3.7) and last words (M = 2.5) in familiar idioms, such that middle words had 1.2 more regressions out than last words. There was a marginal difference ($p1 = .08$) in regression out count between first words (M = 3.6) and last words (M = 2.5) in familiar idioms, such that first words had 1.1 more regressions out than last words. For words within unfamiliar idioms, there was a significant difference ($p1 < .01$) in regression out count between first words (M = 2.1) and middle words (M = 6.9), such that middle words had 4.8 more regressions out than first words. There was a significant difference ($p1 < .01$) in regression out count between middle words (M = 6.9) and last words (M = 3.5) in unfamiliar idioms, such that middle words had 3.4 more regressions out than last words. There was a significant difference ($p1 < .01$) in regression out count between first words (M = 2.1) and last words (M = 3.5) in unfamiliar idioms, such that last words had 1.4 more regressions out than first words.

In summary, metaphors again showed consistent familiarity effects in reading time and fixation count. However, there were few familiarity effects on regression data as a function of familiarity. Idioms also showed consistent familiarity effects in reading time and fixation count measures. Metaphors generally had longer reading times, larger fixation counts, and larger regression counts than idioms. Neither trope showed a familiarity effect on number of regression loops, but there were more loops in metaphors than idioms.

Discussion

The goal of this study was to explore the hypothesis that processing strategies across tropes, as reflected by eye movement behavior, will become more similar as familiarity of the phrases increases. This outcome would support an alternative model of figurative language processing, the Figurative Funnel, that uses familiarity as the basis for describing how figurative tropes are understood. The core of the Figurative Funnel is that all unfamiliar figurative tropes use a meaning construction process, and gradually move to a direct access process when the tropes become familiar. An important point is that the processes for unfamiliar tropes can vary from one another, creating a number of different strategies for interpretation that are trope-dependent. The truly familiar tropes however all are processed using the same direct access approach. Therefore, to support the Figurative Funnel, I must show this variety of processing in unfamiliar tropes, and unity of processing in familiar tropes.

Hypothesis 1: Familiarity Effects

The core of the Figurative Funnel is that tropes behave differently based on their familiarity. As such, familiarity effects should consistently be found for all the measures for both experiments. As expected, metaphors and idioms showed robust familiarity effects in most measures, at both the phrase and word levels. The effects found were shorter reading times, fewer fixations, and fewer regressions into and out of words in familiar phrases compared to words in unfamiliar phrases. This replicates the familiarity effects I found in previous metaphor research based on eye tracking (Campbell, 2014), and is consistent with many other studies that show familiarity effects using other measures (Blasko & Connine, 1993; Bowdle & Gentner, 2005; Campbell, 2014 Frisson & Pickering, 1999; Pickering & Frisson, 2001). Finding reliable familiarity effects for metaphors and idioms is consistent with predictions based on the

Figurative Funnel model.

While consistent familiarity effects were found in metaphors and idioms, they were rarely found in similes. This demonstrates that processing strategies for similes do not seem to change as a function of familiarity, at least as reflected by patterns of eye movements. One possible explanation for this is that similes may be less "figurative" than metaphors and idioms. Because they are overtly comparison statements, similes may not be affected as much by familiarity because the reader always performs a comparison to understand the phrases.

Already then there is evidence that similes are processed differently than metaphors and idioms, but it seems that similes may not differ in how they are processed based on familiarity alone. To support the Figurative Funnel, for a given trope there must be a change in processing style for unfamiliar versus familiar versions of the trope. If the tropes use different processes for familiar and unfamiliar phrases, I expected this to be reflected by different eye movement patterns for familiar and unfamiliar phrases. It is also possible however, for different processes to be reflected by similar patterns. For example, if there is a longer reading time on the vehicle word of a metaphor, with no regressions back to the target word, two different processes are still possible. The first is that they are using the categorization approach and do not need to return to the target to assert the vehicle's properties. The second is that they are doing the mental comparison process, and do not need to actually return their gaze to the target word in order to do it. In addition, I expected similar eye movement patterns across tropes for familiar phrases. To determine if the results match my expectations based on the Figurative Funnel model, below I review the results for each trope to determine if there is a shift in processing style within each trope as a function of familiarity. I then review the results for between-trope comparisons (i.e., metaphors versus similes and metaphors versus idioms) to see if familiar versions of each trope

are processed similarly.

Hypothesis 2: Metaphors

The two experiments consistently showed that unfamiliar metaphors were more difficult to process than familiar metaphors. At the phrase-level, unfamiliar metaphors had longer reading times and more fixations than familiar metaphors. The number of regressions into and out of the metaphor phrase, as well as the number of regression loops, did not differ based on familiarity, however. This may be because reader focus on the component words of metaphors more so than they do the context, regardless of how familiar the metaphor is. In order to determine if a different process is occurring between familiar and unfamiliar metaphors, it is important to consider the differences in patterns at the word-level because this is likely where different patterns will emerge as the reader internally activates or constructs the meaning.

Unfamiliar metaphors had longer reading times on the vehicle words than familiar metaphors across the two experiments. However, unfamiliar metaphors had longer reading times on the target words than familiar metaphors in only one of the two experiments. This finding is unexpected, as previous work (Campbell, 2014) consistently showed less time on the target words of familiar metaphors compared to unfamiliar. Across the two experiments, the target words of familiar metaphors had an equal number of fixations as unfamiliar metaphors. This same pattern held true for the vehicle words as well. It seems then that familiarity effects are produced most strongly by the middle and last words of metaphors. This demonstrates both the importance of the vehicle word, as well as the strength of the eye tracking approach as it allowed the source of the reading time differences as a function of familiarity to be located to specific words.

In Experiment 1, there were no familiarity effects for the target and vehicle words in the

metaphors. That is to say, an equal number of regressions were made into target and vehicle words of familiar and unfamiliar metaphors, and an equal number of regressions were made out of target and vehicle words of familiar and unfamiliar metaphors. In Experiment 2 though, there were some familiarity effects in that target and middle words of familiar metaphors had fewer regressions in than unfamiliar metaphors. There were no familiarity effects for regression out data in Experiment 2. As expected, more regressions were made out of vehicle words, and more regressions were made into target words, for both familiar and unfamiliar metaphors. But this finding is expected given the structure of the metaphor. The target is first, and so it is more likely to be regressed back into, while the vehicle is last and most likely to be regressed out of.

The similarity in the regression patterns, reading times, and fixation counts supports the conclusion that a similar process is being used between familiar and unfamiliar metaphors. For regressions, readers did not have demonstrably different patterns of eye movements between the two familiarity conditions. Familiarity effects were generally consistently found for reading time, which seems to support the conclusion that different processes are being used for familiar and unfamiliar metaphors. Keep in mind, however, that longer processing times could reflect different processes being used or a single process being used that takes longer for unfamiliar metaphors. For reasons outlined below, I believe that this demonstrates that a similar process is used for familiar and unfamiliar metaphors. What this process may be, and the implications of it is discussed later in the text.

Hypothesis 2: Similes

Unlike metaphors, similes at the phrase-level consistently showed no differences in processing difficulty between familiar and unfamiliar similes. That is to say, familiar and unfamiliar similes had similar reading times, fixation counts, regressions into the simile,

regressions out of the simile, and number of regression loops. The combination of no difference in patterns of eye movements (demonstrated by regressions and fixations) and no difference in difficulty (demonstrated by reading time) makes me conclude that similes use the same process for comprehension no matter how familiar the phrase is. Just as with metaphors though, it is important to note how similes may have differed at the word level, to see if there are any clues that imply different processes being used.

The description of any pattern differences in similes can be described very briefly: there were none based on familiarity. For all measures for both the target and vehicle words, there were no differences found as a result of the familiarity of the simile. The only differences found were that more regressions were made into target words than vehicles, and more regressions were made out of vehicles than targets. This did not change as a function of familiarity though. This difference in regressions into the target and out of the vehicle is expected given that similes share virtually the same structure as metaphors.

There seems to be no evidence suggesting that a different process is used to understand similes as a function of familiarity. Not only were no eye movement patterns found that may demonstrate a different process, but the amount of time it took to process familiar and unfamiliar similes was the same. This is evidence that low- and high-familiarity similes use the same process for comprehension, and that there are no differences in difficulty of comprehension as a result of familiarity.

Hypothesis 2: Idioms

At the phrase-level, idioms showed consistent familiarity effects in nearly every measure. Familiar idioms had shorter reading times and fewer fixations than unfamiliar idioms. The shorter reading time and fewer fixations are clear indicators that familiar idioms are easier to

process than unfamiliar idioms. The meaning of unfamiliar, nondecomposable idioms cannot be constructed using clues from the words in the phrase alone, therefore, they must use contextual clues during meaning construction, whereas familiar idioms can be accessed directly and easily. However, there were no differences in the number of regressions to the context, or from the last sentence to the trope, as a function of familiarity. This could be because the contexts were all strongly supportive, it is possible that the readers did not need to consistently return to the context for clues on how to interpret the idiom.

Word comparisons for idioms are different from metaphors and similes in that the idioms used in this study do not have the corresponding keywords that are critical for constructing the meaning of the phrase. In metaphors and similes, the goal is to bridge the gap between the target and vehicle words. In idioms, depending on their compositionality, the words used to compose the phrase can contribute a great deal of the meaning to very little meaning. In addition, metaphors and similes have a structured set of words, whereas the structure of idioms varies. Therefore, comparisons at the word level are not centered around specific crucial words. Instead, the focus is on finding a similar pattern of results that were found at the phrase level.

Idioms showed consistent familiarity effects. Familiar idioms had faster reading times, lower fixation counts, and fewer regressions into and out of words in the idioms compared to unfamiliar idioms. In terms of reading time, the first, middle, and last words of idioms were always read faster when they were part of a familiar idiom. This pattern was found for fixation counts on words as well, though first words in familiar idioms did not have significantly fewer fixations than found for unfamiliar idioms. An important pattern to note is that the first words in familiar idioms had both longer reading times and more fixations than the last words. This could reflect readers accessing (or beginning to access) the meaning of the idiom quickly after reading

the first word or two, and requiring very little reading time on the final word as they have already recognized the idiom and started accessing the meaning directly from their lexicon. This is supported as well by the fact that there were fewer fixations on the last words than first words, indicating that readers may not have needed to fixate the last words for every idiom as the words in the phrases were expected (because the idioms were known) and were already available. Contrast this with the fact that similes showed equal reading times and fixation counts on the first and last words, and we can see that familiar idioms are being processed differently than the other tropes. In addition, unfamiliar idioms had equal reading times and fixation counts on the first and last words of the idiom, showing that the reader needed time to fixate each word in the idiom in order to process it. This is additional evidence that not only do idioms show familiarity effects, but that they may use a different process as they shift from unfamiliar to familiar.

In terms of regression patterns, idioms showed generally consistent familiarity effects as well, with fewer regressions into and out of words in familiar idioms than unfamiliar idioms. One exception to this is that the last words in familiar idioms had the same number of regressions out as in unfamiliar idioms. Just as was seen in metaphors and similes, idioms showed more regressions into the first word than the last word, but only showed more regressions out of the last word than the first word in unfamiliar idioms, and not in familiar idioms. This pattern of regressions is expected though, as the structure of the phrase lends itself to more regressions out of the final word and into the first word.

Looking at the combination of movement data (regressions and fixations) and processing difficulty data (reading time), I believe that idioms show some evidence of a shift in processing style as they go from unfamiliar to familiar. Unfamiliar idioms not only are more difficult to process than familiar idioms, as demonstrated by increased reading time and fixation count, but

the differences at the word level indicate that they are processed differently than familiar idioms. Specifically, when looking at the total reading time and fixation count together, the first word and last words of unfamiliar idioms are equal to one another. Contrast this with familiar idioms, which have faster reading time and fewer fixations on last words than first words, and we see a different pattern. Readers must construct the meaning of unfamiliar idioms, which requires them to attend to all of the words of the idiom, and attend to the words more carefully than they would have to in familiar idioms. In addition, unfamiliar idioms have more regressions into the last word than familiar idioms, which is evidence that readers may be reading further into the context looking for clues in order to construct the meaning of unfamiliar idioms. Familiar idioms have few regressions into the last word, which is evidence that the reader has already accessed the meaning and has little reason to return to the phrase. However, at the phrase level, there were no differences found in the number of regressions to the context, or from the last sentence of the context to the phrase, made as a function of familiarity.

Hypothesis 3: Comparisons across Tropes

The within-trope comparisons presented above provide evidence of whether or not a given trope shifts its processing style as it goes from unfamiliar to familiar. As a short summary, the early evidence is that metaphors use the same process regardless of familiarity. The fixation time and regression data indicate that this process is more difficult for unfamiliar metaphors. The evidence also demonstrates that familiar and unfamiliar similes use the same process as one another. The case for this is very strong given that there were no familiarity effects at the phrase and word level, which indicates that the same process is used for familiar and unfamiliar similes, and that this process has no noticeable differences in difficulty based on familiarity. While metaphors and similes appear to use the same process for familiar and unfamiliar phrases, the

evidence for idioms is that they use different processes for familiar and unfamiliar phrases. This is reflected by consistent familiarity effects as well as differences in regression and fixation patterns based on familiarity.

These within-trope comparisons are an essential part of the Figurative Funnel, as they demonstrate whether or not there are shifts in processing from meaning construction to direct access within a trope. It is also necessary to compare the tropes to assess the second assumption of the Figurative Funnel. That assumption is that if tropes use different processes when they are unfamiliar, those processes will be reflected by different patterns of eye movements; and that if familiar tropes use the same direct access process, the patterns of eye movements will be very similar to one another. Therefore, the comparisons that follow are between the tropes, based on the trope comparisons set up in the two experiments.

Hypothesis 3: Metaphors vs. Similes

Starting with the main effect of trope, there were no differences between metaphors and similes across the measures. These results initially show the similarity between the two tropes, in that neither one is inherently easier to process than the other when collapsed across familiarity. For reading time at both the phrase and word levels, metaphors and similes were virtually identical to one another at both levels of familiarity. That is to say, familiar metaphors had the same reading time as familiar similes. In addition, the first words of familiar metaphors and similes were equal, as were the middle and last words. One exception was that first words in unfamiliar metaphors had longer reading times than first words in unfamiliar similes. This same pattern, with metaphors and similes being the same across all conditions except for the first words of unfamiliar tropes, was seen in the fixation count data as well.

In terms of regression data, metaphors and similes were identical to one another at both

the phrase and word levels. That is to say, first words of familiar tropes were equal, as were middle words, last words, and the phrase as a whole. Some marginal differences were found for unfamiliar tropes, such that middle and last words in unfamiliar metaphors tended to have more regressions out than unfamiliar similes.

I believe that the results support two possible conclusions. The first is that metaphors and similes use the same process as one another. This is demonstrated through the numerous equalities across the measures at both the phrase and word level, as well as in both familiarity conditions. While unfamiliar metaphors may be more difficult to process than familiar metaphors, there is no obvious indication that they are processed differently than unfamiliar similes.

An alternative explanation is that unfamiliar metaphors are in fact processed differently from similes, and that some of the results have begun to show this trend. Specifically, it is the combination of two differences that support this conclusion. First is that first words in unfamiliar metaphors are read more slowly than first words in unfamiliar similes. Second is that first words in unfamiliar metaphors have more fixations than first words in unfamiliar similes. Taken together, this shows a pattern in which unfamiliar metaphors require more effort assimilating information between the target and vehicle words compared to unfamiliar similes. Familiar metaphors and similes however, have essentially no differences from one another. It is more likely that the same process is being used.

However, it could be that this small difference in pattern (more time and fixations on first words in unfamiliar metaphors than similes) could be indicative of category creation and assertion, as described by the categorization model of metaphor processing. This model predicts that the reader must create a category for the vehicle word of unfamiliar metaphors and then

assert those properties onto the target word. If that is the case, the increase in reading time and fixations on first words in unfamiliar metaphors may be representing this assertion process. It is not clear from the data though that such a distinctly different process is occurring in unfamiliar metaphors compared to similes.

It is difficult to make a concrete conclusion about the between-trope comparisons of metaphors and similes and how they relate to the Figurative Funnel. While the two tropes are identical when familiar, as predicted by the Figurative Funnel, they are also nearly identical when unfamiliar, which contradicts a major assumption of the Figurative Funnel. An argument could be made that the few differences in the unfamiliar condition may be evidence for a different process between metaphors and similes, but I believe that argument is weak. Based on the results of Experiment 1, I conclude that metaphors and similes are processed similarly to one another, regardless of familiarity.

Hypothesis 3: Metaphors vs. Idioms

Starting with the main effect of trope, metaphors were more difficult to process than idioms. Metaphors have longer reading times, more regressions out of the phrase, and more regression loops. These differences show that regardless of familiarity, metaphors are at least more difficult to process than idioms, and may use a different strategy for comprehension. In terms of reading time, metaphors took longer to read than idioms at both the phrase level, and on first and last words. An important distinction though, is that unfamiliar idioms and metaphors both had equal reading time on the first and last words, but familiar idioms had less reading time on the last word than first word. These results match the fixation count data, in which metaphors had more fixations on first and last words than idioms in both familiarity conditions. Again, unfamiliar metaphors and idioms had equal fixation counts between first and last words, but

familiar idioms had fewer fixations on the last word than first word. Taken together, this is evidence that idioms are not only easier to process than metaphors generally, but that they do not require the same level of attention to the individual words as metaphors. Specifically, readers need less time to process the last word of familiar idioms, when they fixate on the last word at all, compared to both familiar and unfamiliar metaphors. In part, this could be because readers do not know they have read a metaphor until they reach the final word, whereas for idioms, readers may have already accessed (or begun accessing) the meaning of the idiom before reaching the final word of very familiar idioms.

For the regression data, there was no main effect of trope for number of regressions into the phrase. While there were not many differences at the word level, metaphors had more regressions into middle words than idioms. One interesting difference is that unfamiliar idioms showed more regressions into the last word than unfamiliar metaphors. I take this as some evidence that the readers continued on into the last sentence of the passage to search for contextual clues when reading unfamiliar idioms, but did not tend to do this for unfamiliar metaphors. Metaphors consistently showed more regressions out of both familiar and unfamiliar metaphors than did idioms. At the word-level, metaphors tended to have more regressions out of the phrase than idioms. This was seen in familiar middle and last words, and unfamiliar first and last words. There also were more regression loops in metaphors than idioms, at both levels of familiarity.

Based on these findings, it seems as though two different processes are being used to understand idioms and metaphors. While that alone is not surprising, and is in fact expected for unfamiliar tropes according to the Figurative Funnel, this seems to also be the case for familiar metaphors and idioms. Familiar idioms tend to be processed much more easily than familiar

metaphors, as evidenced by the shorter reading times, smaller fixation counts, and fewer regressions. In addition, metaphors and idioms show a distinct difference in how much the last word of the phrase is prioritized compared to metaphors. As noted above, readers generally do not know they are reading metaphors until they read the vehicle word, even in the case of familiar metaphors. Familiar idioms, on the other hand, may not need much attention on the last word if the reader is able to access (or begin to access) the idiom directly from the lexicon before they finish reading the phrase. Final conclusions regarding whether the results for metaphors and idioms support the Figurative Funnel will be presented below.

Metaphors vs. Metaphors

Before making any conclusions about the Figurative Funnel, it is important to note that metaphors were used in two different experiments. As a whole, there were very few differences between the metaphor data found in Experiment 1 and Experiment 2. In terms of reading times and fixation counts, metaphors showed significant or marginally significant familiarity effects across both studies, such that familiar metaphors were read faster and required fewer fixations than unfamiliar metaphors. One difference of note is that in Experiment 1, first words had more fixations than last words, but this difference did not exist in Experiment 2. I do not have any reason as to why first words received more fixations in Experiment 1 than 2, as they used exactly the same stimuli. Both studies showed the same results in terms of regressions into and out of the metaphors at the word-level, however there appears to be more regressions to the context in Experiment 2 ($M = 8.7$) than in Experiment 1 ($M = 6.1$). Beyond that, there were no differences of theoretical interest in metaphor data between the two experiments.

Conclusions about the Figurative Funnel

A key question to address is whether any evidence was found that supports my Figurative

Funnel model. In order to be supported, two major assumptions must be met: one within-trope, and one between-trope. Within a given trope, there must be a distinct shift in how the trope is processed as it gains familiarity. Specifically, processing must move from meaning construction when the trope is unfamiliar to direct access for familiar tropes. The second assumption is that if there is a difference between the tropes in how they are processed when they are unfamiliar, the eye movement patterns will differ, so that it is clear that different types of meaning construction are used that are trope-dependent. Familiar tropes should all use direct access of their meaning, and so differences in processing will not be evident due to the similarity of this process across tropes. I will begin by addressing how each trope applies to the first assumption. I will then make conclusions about the between-trope assumptions at the unfamiliar and familiar levels.

For metaphors, I concluded that it is likely the case that the same process is being used by both familiar and unfamiliar metaphors, and that this process is more difficult to accomplish for unfamiliar metaphors. This is due to the fact that the regression and fixation patterns showed no differences as a result of familiarity, and it was only the amount of reading time that showed any consistent familiarity effects. In Experiment 1 the target and vehicle words were equal in reading time and fixation count regardless of familiarity, and in Experiment 2 the first words had more fixations, but this was seen in both familiar and unfamiliar metaphors.

It is the fact that there were so few familiarity effects for regression measures, coupled with the fact that any word-level differences were matched between familiarity conditions, that leads me to conclude that there is a similar process being used to process familiar and unfamiliar metaphors. Therefore, metaphors do not meet the first assumption of the Figurative Funnel, in that they do not show a distinctly different process being used as the trope shifts from unfamiliar to familiar.

Even more so than metaphors, similes showed absolutely no familiarity effects in any measure. With no differences in regression and fixation patterns, I must conclude that there is no difference in the process that is being used based on familiarity. In addition, with no reading time differences as a function of familiarity, it seems to be the case that unfamiliar similes are no harder to process than familiar similes. It is easy to conclude then that similes do not meet the first assumption of the Figurative Funnel.

Unlike metaphors and similes, idioms showed a number of familiarity effects, such as the reading time and regression patterns on the words of the idiom. I believe that there is evidence that idioms have a shift in processing style based on familiarity. This first piece of evidence is that for unfamiliar idioms, there was an equal amount of reading time and fixations on the first and last words, supporting the belief that the reader must attend more carefully to all of the words of an idiom in order to understand it. For familiar idioms however, there were more fixations and longer reading times on the first word of the phrase than the last word. This may be because readers are quickly accessing the meaning of the idiom directly from the lexicon, and so will need to make only a short fixation on the last word, if they fixate it at all.

The second piece of evidence for different processes is that there are more regressions into the last words of unfamiliar idioms compared to familiar idioms. Because idioms are so contextually dependent, the reader may continue past an unfamiliar idiom in order to find more clues to its meaning, only to return to it to continue constructing a meaning or to verify the meaning that was constructed. Familiar idioms, on the other hand, are accessed so easily and directly, especially in the supportive contexts provided in this study, that there is very little need for readers to return to the idiom after reading it. Taken together, I believe that an argument could be made that different processes are used in idioms based on their familiarity, and

therefore idioms meet the first assumption of the Figurative Funnel.

The second assumption of the Figurative Funnel is that the tropes may use different forms of meaning construction when they are unfamiliar, creating different patterns of eye movements, and therefore comparisons between the tropes purely at the unfamiliar level are necessary. Very few differences were found at the word-level between unfamiliar metaphors and similes. While some difference approached significance, the only significant difference was that first words in unfamiliar metaphors had longer reading time than unfamiliar similes. As I have said before, reading time differences generally represent differences in processing difficulty. It is the regression and fixation differences, of which there were none, that show different types of processing being used between the tropes. With no differences being found in the fixation counts and regression patterns, I must conclude that unfamiliar metaphors and similes use a similar processing style. This is not a violation of the second assumption of the Figurative Funnel though. If the two tropes use the same process, it is only natural that the eye movements at the unfamiliar end are similar to one another. Therefore, the results show that tropes that use similar processing strategies show similar eye movement patterns.

When comparing unfamiliar metaphors to unfamiliar idioms, there are clear differences in how the two tropes are processed. Not only were unfamiliar metaphors more difficult to process as shown through increased reading time, but metaphors and idioms had different regression and fixation patterns. Unfamiliar metaphors had more fixations on first and last words than did idioms. This coupled with the reading time data shows that the individual words in unfamiliar metaphors are more essential to understanding the phrase than the individual words of unfamiliar idioms. There were many more regressions out of unfamiliar metaphors at the phrase and word levels than unfamiliar idioms. Most importantly, unfamiliar idioms had more

regressions into the last word of the phrase, which indicates that readers are making more effort to find context clues in unfamiliar idioms than metaphors. The evidence points to unfamiliar metaphors focusing processing more internally (within a phrase) to construct a meaning, such as using information from the target and vehicle words, whereas unfamiliar idioms are more reliant on contextual clues, and yet are easier to process.

The second assumption of the Figurative Funnel states that tropes will use similar processes when they are familiar. Specifically, they will all use direct meaning access. Therefore, comparisons at the familiar end should show similarities amongst the three tropes. As stated before, I believe that metaphors and similes use a very similar, if not the same process to be understood. This is clearly the case for familiar metaphors and similes, in that there were almost no differences at all between the two tropes. The two tropes share the same reading times, fixation counts, and regression patterns when they are familiar. The one difference is that more regression loops occurred in familiar metaphors than similes. This difference though, is not enough for me to conclude that the two tropes use different processes. It seems clear then that the process used by metaphors and similes is not only similar when they are unfamiliar, but essentially the same process when they are familiar.

For metaphors and idioms, there were many instances where familiar idioms were easier to process, as seen in decreased reading time, fixation counts, and regression counts. For familiar metaphors, readers put equal emphasis on the first and last words of the phrase, in both reading time and fixation count. For familiar idioms, on the other hand, there was a decrease in the amount of time and fixations needed on the final word. More regressions were made out of words in unfamiliar metaphors as well. These differences make me believe that familiar idioms use a different process than familiar metaphors. This is largely because metaphors must be read

entirely before they can begin to be comprehended, whereas the meaning of highly familiar idioms can be accessed (or begun to be accessed) from the lexicon before the phrase has been fully read. Therefore, the familiar versions of these two tropes do not use a similar process.

Overall then, it would seem that there is little evidence supporting the Figurative Funnel. I believe that there is strong evidence that meaning construction is used by all three tropes when they are unfamiliar. Unfamiliar metaphors and similes must have their meanings constructed using a feature matching comparison process in which the target and vehicle words are of equal importance to understanding the phrase. Unfamiliar idioms must have their meanings constructed based on contextual clues rather than any literal meaning of their individual words.

The results indicate that both assumptions of the Figurative Funnel are violated. The first assumption, in which within a trope there will be a shift in processing style as a function of familiarity, was not seen in metaphors and similes. The evidence from this study is that metaphors and similes both use a comparison process regardless of their familiarity, thus violating the first assumption. Idioms, on the other hand, did show evidence of a change in processing style based on familiarity. Unfamiliar idioms were more difficult to process and required regressions back into the phrase in order for meaning to be constructed. Familiar idioms were processed very easily, and oftentimes did not even need attention on the last word of the phrase before the meaning was accessed.

The second assumption is that the tropes would all show a similar process when they are familiar, and different processes when unfamiliar. There was evidence for this assumption between familiar metaphors and similes because they use the same process. However, that process was the same when they are unfamiliar, and thus violates the first assumption. Metaphors and idioms showed different processes when they were unfamiliar, which supports the Figurative

Funnel. However, they also showed very different styles when they were familiar. I do not believe that the patterns shown by familiar metaphors and similes are evidence that any direct access is occurring, or at least not nearly to the same extent as idioms. Therefore, the second assumption of the Figurative Funnel was violated. With both assumptions violated, the current study does not support the idea of the Figurative Funnel. Why this might be the case is discussed later in the paper.

Connections to other models/research

If the Figurative Funnel was not supported by the results of the present experiments, then it is prudent to see how the results might apply to other models and previous research. Current models of figurative language processing do not make predictions about eye movement patterns, and very few studies have been done using eye tracking with figurative language; therefore, conclusions about eye movement patterns reflect my interpretation of the results, not actual predictions from theories.

A major assumption of the Career of Metaphor model (Bowdle & Gentner, 2005) is that unfamiliar metaphors and similes are processed the same way. This study showed support for this idea in that unfamiliar metaphors were processed the same as unfamiliar idioms. However, familiar metaphors were also processed in the same way as familiar similes, an assumption not made by the Career of Metaphor model, which instead predicts familiar metaphors to be processed more easily through a categorization process. This study then does support the idea that unfamiliar metaphors are treated the same way as similes, but does not show the advantages in ease of processing and change in processing style that is expected in familiar metaphors.

According to the categorization model (Glucksberg, 2008), unfamiliar metaphors use the same process as familiar metaphors, but the process is just more difficult to accomplish. This

process is one in which the vehicle word is associated with a category, as an exemplar, and the characteristics of that category are asserted onto the target. Oftentimes in past research I have argued that this would result in shorter reading times on target words compared to vehicles, given the importance of the vehicle in this model. However, I now am beginning to believe that I have been underselling the importance of the target word. If there is indeed an assertion process from the vehicle to the target, it follows that there would be more similar reading times and fixation counts between the target and vehicle words if readers regress to the target. This was demonstrated in both metaphor studies, in which target and vehicle words had similar reading times and fixation counts. In fact, in one study there were more fixations on target words than vehicles for both familiar and unfamiliar metaphors. These findings might be representative of the category assertion process, where the reader returns to the target to assert the characteristics of the vehicle. Coupled with that idea is that this pattern was consistent across familiar and unfamiliar metaphors, and that differences were in magnitude of the measures not the patterns, supports the idea that the same process is used for metaphors regardless of familiarity, and that that process is just more difficult for unfamiliar metaphors. I am not making a definitive conclusion that the categorization model is fully correct, but this is a possible interpretation of the results of these studies.

The hybrid model of idiom processing (Titone & Libben, 1999) states that familiar idioms have their meanings accessed directly, whereas unfamiliar idioms require their meaning to be constructed. My results directly support this model. One important distinction to make is that the idioms used in my study were nondecomposable, meaning their meaning could not be constructed based on the meaning of the individual words. More research must be done with decomposable idioms to provide further support for this model.

One of the few studies that have used eye tracking to directly compare two figurative tropes is Ashby et al.'s (2017) study comparing metaphors and similes. They found differences in first pass reading time, such that unfamiliar metaphors were initially more difficult to process than familiar metaphors and similes, which were similar. I did not report first pass reading time data however, and cannot make a direct comparison between my study and theirs.

Limitations and Future Directions

One limitation with this study is the difficulty in mapping eye movements to cognitive processes. As seen in the conclusions made about metaphors, it is possible that the metaphor data supports the use of categorization for all metaphors, as well as comparison for all metaphors. As such, this work is preliminary, and the necessary start of a catalog of eye tracking measures of figurative language that will allow for more comparisons in the future.

Future research should include a greater variety of tropes. For example, metonymy have been studied using eye tracking methods (Frisson & Pickering, 1999), but metonymies were not included here because of their very different syntactic structure. An experiment that allows for the direct comparison of similes to idioms is an obvious and necessary next step. It should be noted though, that given the extreme similarity between metaphors and similes, the comparisons between idioms and metaphors made in this paper could be synonymous to the comparisons made between idioms and similes. Using a within-subjects design could also be beneficial. It is possible that readers learned to expect a metaphor, simile, or idiom in the third sentence of the passages despite the inclusion of filler passages. Presenting multiple tropes to individual participants would further hide the purpose of the research. Using a within-subject design also would control for individual reader differences, but would lead to a longer experiment with possible fatigue effects.

The present experiments only used strongly supportive contexts. Future research should manipulate the strength of the context in order to see if that affects processing style. This is especially relevant given how context dependent idioms are, and any context manipulation would be able to show support for that notion.

Conclusions

In summary, the Figurative Funnel model was not supported. That said, some very interesting conclusions can be made. First, similes use the same comparison process regardless of their familiarity. This makes sense considering that similes invite the reader to make a comparison because of the inclusion of the word "like". Readers are given the prompt to compare the target and vehicle word because they have been told they are alike. What is interesting though is that this process does not seem to be more difficult when the similes are unfamiliar compared to familiar. With no familiarity effects at the phrase level, the evidence indicates that similes use a comparison process that is so comfortable to the reader that they have no trouble using it even when the simile is unfamiliar. It is also possible that the strong contexts used here made unfamiliar versions of the similes "too easy" to understand.

The second interesting finding is that metaphors appear to use the same process regardless of their familiarity. This was demonstrated with the similar regression and fixation patterns across familiarity, and that any differences found were with reading time; a measure of difficulty. This finding is counter to the Career of Metaphor model that assumes metaphors shift in processing style as familiarity increases. What process is used for metaphors is unclear based on the findings of this study. I have made an argument that it may be the categorization process described by Glucksberg. However, as seen in the next point, this may not be the case.

The third interesting finding is that there was evidence that metaphors and similes use a

very similar process to one another. This is unexpected, as the assumption is that similes are universally understood through a comparison process. If metaphors use the same process as similes regardless of familiarity, then metaphors may always use a comparison between the target and vehicle words. This finding is of course not conclusive, but future research is needed to better understand the relationship between metaphors and similes.

The final interesting finding is that idioms were the one trope that showed evidence of a shift in processing style based on familiarity. As far as decomposable idioms are concerned, this study showed that unfamiliar idioms must have their meaning constructed, but they can have their meaning accessed directly when they familiar.

While the Figurative Funnel model was not supported by this study, the findings described above show some expected and some unexpected results that have major implications for figurative language research. The complexities of figurative language processing make producing definitive conclusions difficult, but also provide opportunities for advancing current theories.

Works Cited

Ashby, J., Roncero, C., de Almeida, R., & Agauas, S. (2017). The early processing of metaphors and similes: Evidence from eye movements. *The Quarterly Journal of Experimental Psychology,* DOI: 10.1080/17470218.2016.1278456

Baptista, N. I., Macedo, E. C., & Boggio, P. S. (2015). Looking more and at different things: Differential gender eye-tracking patterns on an irony comprehension task. *Psychology & Neuroscience, 8*(2), 157-167.

Blasko, D. & Connine, C. (1993). Effects of familiarity and aptness on metaphor processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(2), 295-308.

Bowdle, B., & Gentner, D. (1997). Informativity and asymmetry in comparisons. *Cognitive Psychology, 34*(3), 244-286.

Bowdle, B., & Gentner, D. (2005). The career of metaphor. *Psychological Review, 112*(1), 193-216. Burgess, C., & Chiarello, C. (1996). Neurocognitive mechanisms underlying metaphor comprehension and other figurative language. *Metaphor and Symbol, 11*(1), 67-84.

Buckingham, D. (2006). A multilingual didactic approach to idioms using a conceptual framework. *Language Design, 8*, 35-45

Campbell, S. (2014). Evaluating the career of metaphor and categorization models of metaphor processing: An eye-tracking perspective

Campbell, S. (2015). Comparing models of figurative language processing: The role of familiarity.

Campbell, S., & Raney, R. (2016). A 25-year replication of Katz et al.'s (1988) metaphor norms. *Behavioral Research Methods*.

Chettih, S., Durgin, F., & Grodner, D. (2012). Mixing metaphors in the cerebral hemispheres: What happens when careers collide? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(2), 295-311.

Colston, H. L. (2005). Social and cultural influences on figurative and indirect language. *Figurative language comprehension: Social and cultural influences* (pp. 99-130). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, Mahwah, NJ.

Connine, C., Blasko, D., Brandt, R., & Layer, J. (1992). Idiomatic processing: Syntactic frozenness and subjective familiarity. *Psychological Research, 53*(3), 225-232.

Cronk, B., & Schweigert, W. (1992). The comprehension of idioms: The effects of familiarity, literalness, and usage. *Applied Psycholinguistics, 13*(2), 131-146.

Davies, M. (May, 2016) *Corpus of Contemporary American English*. Retrieved from corpus.byu.edu.

Filik, R., Leuthold, H., Wallington, K., & Page, J. (2014). Testing theories of irony processing using eye-tracking and ERPs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(3), 811-828.

Filik, R., & Moxey, L. M. (2010). The on-line processing of written irony. *Cognition, 116*(3), 421-436.

Frisson, S., & Pickering, M. (1999). The processing of metonymy: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(6), 1366-1383.

Frisson, S., & Pickering, M. (2007). The processing of familiar and novel senses of a word: Why reading Dickens is easy but reading Needham can be hard. *Language and Cognitive Processes, 22*(4), 595-613.

Gentner, D., & Bowdle, B. F. (2001). Convention, form, and figurative language processing. *Metaphor and Symbol, 16*(3-4), 223-247.

Gentner, D., & Wolff, P. (1997). Alignment in the processing of metaphor. *Journal of Memory and Language, 37*, 331-355.

Gibbs, R. W., Jr. (1994). *Figurative thought and figurative language*. San Diego, CA, US: Academic Press, San Diego, CA.

Gibbs, R. W., Jr. (2002). A new look at literal meaning in understanding what is said and implicated. *Journal of Pragmatics, 34*(4), 457-486.

Gibbs, R. W., Jr., Leggitt, J. S., & Turner, E. A. (2002). What's special about figurative language in emotional communication? *The verbal communication of emotions: Interdisciplinary perspectives* (pp. 125-149). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, Mahwah, NJ.

Gibbs, R., & Nayak, N. (1989). Psycholinguistic studies of the syntactic behavior of idioms. *Cognitive Psychology, 21*(1), 100-138.

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

Glucksberg, S. (2003). The psycholinguistics of metaphor. *TRENDS in Cognitive Sciences, 7*(2), 92-96.

Glucksberg, S. (2008). How metaphors create categories - quickly. In R. Gibbs (Ed.), *The Cambridge handbook of metaphor and thought* (1st ed., pp. 67-83). New York, NY: Cambridge University Press.

Glucksberg, S., Gildea, P., & Bookin, H. (1982). On understanding nonliteral speech: Can people ignore metaphors? *Journal of Verbal Learning and Verbal Behavior, 21*(1), 85-98.

Glucksberg, S., & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychological Review, 97*(1), 3-18.

Glucksberg, S., McGlone, M., & Manfredi, D. (1997). Property attribution in metaphoric comprehension. *Journal of Memory and Language, 36*, 50-67.

Glucksberg, S., Newsome, M., & Goldvarg, Y. (2001). Inhibition of the literal: Filtering metaphor-irrelevant information during metaphor processing. *Metaphor and Symbol, 16*(3), 277-293.

Gokcesu, B. (2008). Metaphor processing and polysemy. *ProQuest Information and Learning, 69*(2).

Hasson, U., Estes, Z., & Glucksberg, S. (2001). Metaphors communicate more effectively than do similes. *Abstracts of the Psychonomic Society, 42nd annual meeting, 6,* 103.

Johnson, A. T. (1996). Comprehension of metaphors and similes: A reaction time study. *Metaphor & Symbolic Activity, 11*(2), 145-159.

Jones, L., & Estes, Z. (2006). Roosters, robins, and alarm clocks: Aptness and conventionality in metaphor comprehension. *Journal of Memory and Language, 55*(1), 18-32.

Katz, A., Paivio, A., Marschark, M., & Clark, J. (1988). Norms for 204 literary and 260 nonliterary metaphors on 10 psychological dimensions. *Metaphor and Symbolic Activity, 3*(4), 191-214.

Keysar, B. (1989). On the functional equivalence of literal and metaphorical interpretations in discourse. *Journal of Memory and Language, 28*(4), 375-385.

Libben, M. R., & Titone, D. A. (2008). The multidetermined nature of idiom processing. *Memory & Cognition (Pre-2011), 36*(6), 1103-21.

Miller, G. (1979). Images and models, similes and metaphors. In A. Ortony (Ed.), *Metaphor and*

Thought. New York: Cambridge University Press.

Minkoff, S., & Raney, G. (2000). Letter-detection errors in the word *the*: Word frequency versus syntactic structure. *Scientific Study of Reading*, *4*(1), 55-76.

Nippold, M., & Taylor, C. (2002). Judgments of idiom familiarity and transparency: A comparison of children and adolescents. *Journal of Speech, Language, and Hearing Research, 45*(2), 384-391.

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

Pickering, M. J., & Frisson, S. (2001). Processing ambiguous verbs: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(2), 556-573.

Olkoniemi, H., Ranta, H., & Kaakinen, J. K. (2016). Individual differences in the processing of written sarcasm and metaphor: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(3), 433-450.

Roncero, C., de Almeida, R., Martin, D., & de Caro, M. (2016). Aptness predicts metaphor preference in the lab and on the Internet. *Metaphor and Symbol, 31*(1), 31-46.

Schweigert, W. (1991). The muddy waters of idiom comprehension. *Journal of Psycholinguistic Research, 20*(4), 305-314.

Schweigert, W. A., & Moates, D. R. (1988). Familiar idiom comprehension. *Journal of Psycholinguistic Research, 17*(4), 281-296.

Shibata, M., Toyomura, A., Motoyama, H., Itoh, H., Kawabata, Y., & Abe, J. (2012). Does simile comprehension differ from metaphor comprehension? A functional MRI study. *Brain and Language, 121*(3), 254-260.

Swinney, D., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal*

of Verbal Learning and Verbal Behavior, 18(5), 523-534.

Therriault, D., & Raney, G. (2007). Processing and representing temporal information in narrative text. *Discourse Processes*, *43*(2), 173-200.

Titone, D., & Connine, C. (1994). The comprehension of idiomatic expressions: Effects of predictability and literality. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 20,* 1126-1138.

Titone, D. A., & Connine, C. M. (1999). On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics, 31*(12), 1655-1674.

Utsumi, A. (2011). Computational exploration of metaphor comprehension processes using a semantic space model. *Cognitive Science, 35*(2), 251-296.

Table 1

Experiment 1 Means By Subject (M1) and By Item (M2) For Reading Time (in ms) For the Phrase, Fixation Count For the Phrase, Regression to Context Count, Regression From Last Sentence to Trope Count, and Regression Loop Count as a Function of Trope (Metaphor, Simile) and Familiarity (Familiar, Unfamiliar). Standard Error in Parentheses.

Figurative Language Condition	Reading Time (ms)	Fixation Count	Regression To Context Count	Regression From Last Sentence to Trope Count	Regression From Last Word to First Word Count	Regression Loop Count
Familiar Metaphor	M1 = 978 (49.6) M2 = 988 (28.3)	M1 = 3.3 (.20) M2 = 3.3 (.15)	M1 = 5.5 (.80) M2 = 5.5 (.64)	M1 = 1.7 (.49) M2 = 1.7 (.34)	M1 = 3.3 (.35) M2 = 3.4 (.35)	M1 = 2.3 (.30) M2 = 2.4 (.29)
Familiar Simile	M1 = 951 (49.6) M2 = 951 (28.3)	M1 = 3.3 (.20) M2 = 3.3 (1.5)	M1 = 5.6 (.80) M2 = 5.6 (.78)	M1 = 2.1 (.49) M2 = 2.1 (.39)	M1 = 3.0 (.35) M2 = 2.9 (.35)	M1 = 2.1 (.30) M2 = 2.0 (.30)
Unfamiliar Metaphor	M1 = 1068 (47.9) M2 = 1062 (34.2)	M1 = 3.5 (.18) M2 = 3.5 (1.5)	M1 = 6.8 (.98) M2 = 6.8 (.64)	M1 = 2.0 (.43) M2 = 2.0 (.34)	M1 = 3.6 (.34) M2 = 3.5 (.34)	M1 = 2.4 (.34) M2 = 2.5 (.33)
Unfamiliar Simile	M1 = 956 (47.9) M2 = 954 (34.2)	M1 = 3.2 (.18) M2 = 3.2 (.15)	M1 = 6.4 (.98) M2 = 6.4 (.78)	M1 = 2.1 (.43) M2 = 2.1 (.39)	M1 = 3.4 (.34) M2 = 3.4 (.35)	M1 = 2.3 (.34) M2 = 2.1 (.34)
Familiar Average	M1 = 965 (35.0) M2 = 970 (20.0)	M1 = 3.3 (.15) M2 = 3.3 (.10)	M1 = 5.6 (.56) M2 = 5.6 (.63)	M1 = 1.9 (.34) M2 = 1.9 (.32)	M1 = 3.1 (.25) M2 = 3.2 (.26)	M1 = 2.2 (.21) M2 = 2.2 (.23)
Unfamiliar Average	M1 = 1012 (33.9) M2 = 1008 (24.2)	M1 = 3.4 (.13) M2 = 3.4 (.10)	M1 = 6.6 (.69) M2 = 6.6 (.63)	M1 = 2.0 (.31) M2 = 2.0 (.32)	M1 = 3.5 (.24) M2 = 3.4 (.25)	M1 = 2.4 (.24) M2 = 2.3 (.23)
Metaphor Average	M1 = 1023 (47.2) M2 = 1025	M1 = 3.4 (.19) M2 = 3.4	M1 = 6.1 (.77) M2 = 6.1	M1 = 1.8 (.43) M2 = 1.8	M1 = 3.4 (.27) M2 = 3.4	M1 = 2.4 (.25) M2 = 2.4

	(22.1)	(.11)	(.45)	(.24)	(.26)	(.23)
Simile Average	M1 = 954 (47.2) M2 = 952 (22.1)	M1 = 3.3 (.19) M2 = 3.3 (.11)	M1 = 6.0 (.77) M2 = 6.0 (.55)	M1 = 2.1 (.43) M2 = 2.1 (.28)	M1 = 3.2 (.27) M2 = 3.2 (.24)	M1 = 2.2 (.25) M2 = 2.1 (.25)

Table 2

Experiment 1 Means By Subject (M1) and By Item (M2) for Reading Time (ms) Per Word Group as a Function of Familiarity (Familiar, Unfamiliar, Trope (Metaphor, Simile), and Word (First, Middle, Last). Standard Error in Parentheses.

	First Word	Middle Word	Last Word
Familiar Metaphor	M1 = 379 (20.1) M2 = 371 (21.2)	M1 = 257 (11.3) M2 = 264 (6.8)	M1 = 342 (27.1) M2 = 353 (18.7)
Familiar Simile	M1 = 333 (20.1) M2 = 328 (21.2)	M1 = 276 (11.3) M2 = 278 (6.8)	M1 = 343 (27.1) M2 = 345 (18.7)
Unfamiliar Metaphor	M1 = 380 (18.3) M2 = 378 (16.7)	M1 = 301 (12.3) M2 = 299 (10.4)	M1 = 387 (26.8) M2 = 385 (21.3)
Unfamiliar Simile	M1 = 327 (18.3) M2 = 323 (16.7)	M1 = 276 (12.3) M2 = 276 (10.4)	M1 = 354 (26.8) M2 = 356 (21.3)
Word Average	M1 = 355 (12.9) M2 = 350 (10.3)	M1 = 277 (7.4) M2 = 279 (4.3)	M1 = 357 (18.5) M2 = 359 (10.9)

Table 3

Experiment 1 By Subject Means (M1) and By Item Means (M2) for Fixation Count Per Word Group as a Function of Familiarity (Familiar, Unfamiliar), Trope (Metaphor, Simile), and Word (First, Middle, Last). Standard Error in Parentheses.

	First Word	Middle Word	Last Word
Familiar Metaphor	M1 = 1.5 (.08) M2 = 1.5 (.12)	M1 = .57 (.05) M2 = .58 (.03)	M1 = 1.2 (.12) M2 = 1.2 (.09)
Familiar Simile	M1 = 1.3 (.08) M2 = 1.3 (.12)	M1 = .66 (.05) M2 = .66 (.03)	M1 = 1.3 (.12) M2 = 1.3 (.09)
Unfamiliar Metaphor	M1 = 1.6 (.09) M2 = 1.6 (.10)	M1 = .66 (.05) M2 = .66 (.04)	M1 = 1.3 (.10) M2 = 1.3 (.12)
Unfamiliar Simile	M1 = 1.3 (.09) M2 = 1.3 (.10)	M1 = .66 (.05) M2 = .66 (.04)	M1 = 1.3 (.10) M2 = 1.3 (.12)
Word Average	M1 = 1.4 (.06) M2 = 1.4 (.06)	M1 = .64 (.03) M2 = .64 (.02)	M1 = 1.3 (.07) M2 = 1.3 (.06)

Table 4

Experiment 1 By Subject Means (M1) and By Item Means (M2) for Regressions In Count Per Word Group as a Function of Familiarity (Familiar, Unfamiliar), Trope (Metaphor, Simile), and Word (First, Middle, Last). Standard Error in Parentheses.

	First Word	Middle Word	Last Word
Familiar Metaphor	M1 = 5.6 (.86) M2 = 5.5 (.69)	M1 = 5.5 (.93) M2 = 5.5 (.72)	M1 = .90 (.37) M2 = .90 (.25)
Familiar Simile	M1 = 4.6 (.86) M2 = 5.9 (.58)	M1 = 7.8 (.93) M2 = 6.4 (.62)	M1 = 1.5 (.37) M2 = 1.4 (.26)
Unfamiliar Metaphor	M1 = 5.9 (.86) M2 = 4.6 (.69)	M1 = 6.4 (.76) M2 = 7.8 (.72)	M1 = 1.4 (.24) M2 = 1.5 (.25)
Unfamiliar Simile	M1 = 3.7 (.86) M2 = 3.7 (.58)	M1 = 8.5 (.76) M2 = 8.5 (.62)	M1 = .90 (.24) M2 = .85 (.26)
Word Average	M1 = 4.9 (.55) M2 = 4.9 (.29)	M1 = 7.0 (.56) M2 = 7.0 (.35)	M1 = 1.1 (1.8) M2 = 1.1 (.12)

Table 5

Experiment 1 By Subject Means (M1) and By Item Means (M2) for Regressions Out Count Per Word Group as a Function of Familiarity (Familiar, Unfamiliar), Trope (Metaphor, Simile), and Word (First, Middle, Last). Standard Error in Parentheses.

	First Word	Middle Word	Last Word
Familiar Metaphor	M1 = 4.0 (.63) M2 = 4.0 (.51)	M1 = 5.1 (.76) M2 = 5.1 (.63)	M1 = 6.2 (.92) M2 = 6.2 (.51)
Familiar Simile	M1 = 4.6 (.63) M2 = 3.9 (.61)	M1 = 6.0 (.76) M2 = 5.7 (.61)	M1 = 5.1 (.92) M2 = 5.9 (.72)
Unfamiliar Metaphor	M1 = 3.9 (.70) M2 = 4.6 (.51)	M1 = 5.7 (.83) M2 = 6.0 (.63)	M1 = 5.9 (.76) M2 = 5.1 (.51)
Unfamiliar Simile	M1 = 5.2 (.70) M2 = 5.2 (.61)	M1 = 5.9 (.83) M2 = 5.9 (.61)	M1 = 4.5 (.76) M2 = 4.5 (.72)
Word Average	M1 = 4.4 (.44) M2 = 4.4 (.26)	M1 = 5.6 (.53) M2 = 5.7 (.31)	M1 = 5.4 (.54) M2 = 5.4 (.34)

Table 6

Experiment 2 Means By Subject (M1) and By Item (M2) of Reading Time (in ms) For the Phrase, Fixation Count For the Phrase, Regression To Context Count, Regression From Last Sentence to Trope Count, and Regression Loop Count as a Function of Trope (Metaphor, Simile) and Familiarity (Familiar, Unfamiliar). Standard Errors in Parentheses.

	Reading Time (ms)	Fixation Count	Regression To Context Count	Regression From Last Sentence to Trope Count	Regression From Last Word to First Word Count	Regression Loop Count
Familiar Metaphor	M1 = 1087 (36.9) M2 = 1101 (34.5)	M1 = 3.7 (.15) M2 = 3.7 (.17)	M1 = 8.6 (.82) M2 = 8.6 (1.2)	M1 = 2.4 (.39) M2 = 2.4 (.37)	M1 = 3.8 (.32) M2 = 3.7 (.32)	M1 = 2.8 (.29) M2 = 3.0 (.30)
Familiar Idiom	M1 = 860 (36.9) M2 = 943 (34.5)	M1 = 2.9 (.15) M2 = 3.2 (.17)	M1 = 3.7 (.82) M2 = 3.7 (.71)	M1 = 2.8 (.39) M2 = 2.8 (.47)	M1 = 3.1 (.32) M2 = 3.1 (.31)	M1 = 2.0 (.29) M2 = 2.0 (.30)
Unfamiliar Metaphor	M1 = 1158 (46.9) M2 = 1146 (39.0)	M1 = 3.9 (.18) M2 = 3.9 (.19)	M1 = 8.8 (.90) M2 = 8.8 (1.2)	M1 = 1.7 (.50) M2 = 1.7 (.37)	M1 = 4.8 (.38) M2 = 4.7 (.39)	M1 = 3.3 (.32) M2 = 3.1 (.32)
Unfamiliar Idiom	M1 = 1000 (46.7) M2 = 1020 (39.0)	M1 = 3.3 (.18) M2 = 3.7 (.19)	M1 = 3.5 (.90) M2 = 3.5 (.71)	M1 = 2.8 (.50) M2 = 2.8 (.47)	M1 = 3.2 (.38) M2 = 3.2 (.37)	M1 = 2.0 (.32) M2 = 2.1 (.33)
Familiar Average	M1 = 973 (26.1) M2 = 1022 (24.4)	M1 = 3.3 (.11) M2 = 3.5 (.12)	M1 = 6.1 (.58) M2 = 6.1 (.69)	M1 = 2.6 (.28) M2 = 2.6 (.28)	M1 = 3.4 (.22) M2 = 3.4 (.24)	M1 = 2.4 (.21) M2 = 2.5 (.23)
Unfamiliar Average	M1 = 1079 (33.0) M2 = 1083 (27.6)	M1 = 3.6 (.13) M2 = 3.8 (.13)	M1 = 6.1 (.64) M2 = 6.1 (.69)	M1 = 2.2 (.36) M2 = 2.2 (.28)	M1 = 4.0 (.27) M2 = 4.0 (.23)	M1 = 2.6 (.22) M2 = 2.6 (.24)
Metaphor Average	M1 = 1123 (39.5) M2 = 1124	M1 = 3.8 (.15) M2 = 3.8	M1 = 8.7 (.80) M2 = 8.7	M1 = 2.0 (.39) M2 = 2.0	M1 = 4.3 (.26) M2 = 4.2	M1 = 3.0 (.24) M2 = 3.1

	(26.5)	(.13)	(.83)	(.26)	(.25)	(.22)
Idiom Average	M1 = 930 (39.5) M2 = 982 (26.5)	M1 = 3.1 (.15) M2 = 3.4 (.13)	M1 = 3.6 (.80) M2 = 3.6 (.50)	M1 = 2.8 (.39) M2 = 2.8 (.33)	M1 = 3.2 (.26) M2 = 3.1 (.26)	M1 = 2.0 (.24) M2 = 2.0 (.23)

Table 7

Experiment 2 By Subject Means (M1) and By Item Means (M2) for Reading Time (ms) Per Word Group as a Function of Familiarity (Familiar, Unfamiliar), Trope (Metaphor, Simile), and Word (First, Middle, Last). Standard Error in Parentheses.

	First Word	Middle Word	Last Word
Familiar Metaphor	M1 = 397 (17.2) M2 = 400 (23.9)	M1 = 300 (14.7) M2 = 314 (9.8)	M1 = 390 (19.8) M2 = 388 (18.7)
Familiar Idiom	M1 = 313 (17.2) M2 = 321 (23.9)	M1 = 273 (14.7) M2 = 351 (9.8)	M1 = 273 (19.8) M2 = 271 (18.7)
Unfamiliar Metaphor	M1 = 423 (19.4) M2 = 417 (19.7)	M1 = 320 (14.3) M2 = 326 (10.3)	M1 = 416 (21.8) M2 = 403 (25.6)
Unfamiliar Idiom	M1 = 347 (19.4) M2 = 344 (19.7)	M1 = 313 (14.3) M2 = 352 (10.3)	M1 = 341 (21.8) M2 = 324 (25.6)
Word Average	M1 = 370 (11.9) M2 = 371 (10.5)	M1 = 301 (9.3) M2 = 336 (5.7)	M1 = 355 (13.6) M2 = 346 (10.9)

Table 8

Experiment 2 By Subject Means (M1) and By Item Means (M2) for Fixation Count Per Word Group as a Function of Familiarity (Familiar, Unfamiliar), Trope (Metaphor, Simile), and Word (First, Middle, Last). Standard Error in Parentheses.

	First Word	Middle Word	Last Word
Familiar Metaphor	M1 = 1.7 (.09) M2 = 1.7 (.10)	M1 = .64 (.06) M2 = .64 (.03)	M1 = 1.4 (.09) M2 = 1.4 (.10)
Familiar Idiom	M1 = 1.2 (.09) M2 = 1.2 (.10)	M1 = .74 (.06) M2 = .96 (.03)	M1 = .96 (.09) M2 = 1.0 (.10)
Unfamiliar Metaphor	M1 = 1.7 (.08) M2 = 1.7 (.12)	M1 = .78 (.05) M2 = .78 (.04)	M1 = 1.5 (.10) M2 = 1.5 (.12)
Unfamiliar Idiom	M1 = 1.2 (.08) M2 = 1.3 (.12)	M1 = .89 (.05) M2 = 1.0 (.04)	M1 = 1.2 (.10) M2 = 1.3 (.12)
Word Average	M1 = 1.4 (.05) M2 = 1.5 (.06)	M1 = .76 (.04) M2 = .86 (.02)	M1 = 1.3 (.06) M2 = 1.3 (.06)

Table 9

Experiment 2 By Subject Means (M1) and By Item Means (M2) for Regressions In Count Per Word Group as a Function of Familiarity (Familiar, Unfamiliar), Trope (Metaphor, Simile), and Word (First, Middle, Last). Standard Error in Parentheses.

	First Word	Middle Word	Last Word
Familiar Metaphor	M1 = 6.3 (.73) M2 = 6.3 (.73)	M1 = 6.9 (.71) M2 = 6.9 (.94)	M1 = 1.7 (.30) M2 = 1.7 (.38)
Familiar Idiom	M1 = 3.8 (.73) M2 = 7.4 (.98)	M1 = 4.5 (.71) M2 = 8.7 (1.4)	M1 = 1.8 (.30) M2 = 1.4 (.44)
Unfamiliar Metaphor	M1 = 7.4 (.82) M2 = 4.5 (.84)	M1 = 8.7 (.89) M2 = 12.4 (1.1)	M1 = 1.4 (.40) M2 = 2.0 (.44)
Unfamiliar Idiom	M1 = 4.9 (.82) M2 = 6.5 (1.1)	M1 = 7.5 (.89) M2 = 17.5 (1.6)	M1 = 3.0 (.40) M2 = 4.0 (.51)
Word Average	M1 = 5.6 (.49) M2 = 6.2 (.48)	M1 = 6.9 (.51) M2 = 11.3 (.56)	M1 = 2.0 (.20) M2 = 2.3 (.25)

Table 10

Experiment 2 By Subject Means (M1) and By Item Means (M2) for Regressions Out Count Per Word Group as a Function of Familiarity (Familiar, Unfamiliar), Trope (Metaphor, Simile), and Word (First, Middle, Last). Standard Error in Parentheses.

	First Word	Middle Word	Last Word
Familiar Metaphor	M1 = 5.6 (.53) M2 = 5.6 (.75)	M1 = 5.7 (.66) M2 = 5.7 (.64)	M1 = 8.0 (.79) M2 = 8.0 (.75)
Familiar Idiom	M1 = 3.6 (.53) M2 = 4.8 (.65)	M1 = 3.7 (.66) M2 = 6.8 (.85)	M1 = 2.5 (.79) M2 = 8.3 (.87)
Unfamiliar Metaphor	M1 = 4.8 (.54) M2 = 3.7 (.87)	M1 = 6.8 (.87) M2 = 8.9 (.73)	M1 = 8.3 (.70) M2 = 2.8 (.86)
Unfamiliar Idiom	M1 = 2.1 (.54) M2 = 2.8 (.75)	M1 = 6.9 (.87) M2 = 11.8 (.98)	M1 = 3.5 (.70) M2 = 4.7 (1.0)
Word Average	M1 = 4.0 (.34) M2 = 4.2 (.39)	M1 = 5.7 (.48) M2 = 8.3 (.36)	M1 = 5.6 (.45) M2 = 5.9 (.47)

Table 11

Statistical Analyses for Experiment 1 By Subject (F1) and By Item (F2) for Reading Time (in ms) for the Phrase, Fixation Count for the Phrase, Regression To Context Count, Regression From Last Sentence to Trope Count, and Regression Loop Count as a Function of Trope (Metaphor, Simile) and Familiarity (Familiar, Unfamiliar).

	Reading Time	Fixation Count	Regression to Context	Last to Phrase	Last Word to First Word	Regression Loops
	df = 1, 38	*df* = 1, 38	*df* = 1, 38	*df* = 1, 38	*df* = 1, 38	*df* = 1, 38
Familiarity Main Effect	*F1* = 7.7, *p* < .01, MSE= 45315, β = .77 *F2* = 1.5, *p* = .23, MSE = 29223, β = .22	*F1* = .77, *p* = .39, MSE = .10, β = .14 *F2* = .26, *p* = .26, MSE = .10, β = .08	*F1* = 2.6, *p* = .12, MSE = 21.0, β = .35 *F2* = 1.3, *p* = .26, MSE = 21.0, β = .20	*F1* = .59, *p* = .45, MSE = .61, β = .12 *F2* = .15, *p* = .70, MSE = .61, β = .07	*F1* = 1.2, *p* = .28, MSE = 2.5, β = .19 *F2* = 1.0, *p* = .33, *MSE* = 3.1, β = .16	*F1* = .38, *p* = .54, MSE = .61, β = .09 *F2* = .44, *p* = .51, MSE = .58, β = .10
Trope Main Effect	*F1* = 1.1, *p* = .31, MSE = 96050, β = .17 *F2* = 5.4, *p* < .05, MSE = 106215, β = .62	*F1* = .16, *p* = .70, MSE = .22, β = .07 *F2* = .47, *p* = 50, MSE = .22, β = .10	*F1* = .01, *p* = .91, MSE = .31, β = .05 *F2* = .07, *p* = .79, MSE = .31, β = .06	*F1* = .14, *p* = .71, MSE = 1.0, β = .07 *F2* = .82, *p* = .37, MSE = 1.0, β = .14	*F1* = .45, *p* = .51, MSE = 1.3, β = .10 *F2* = .56, *p* = .56, MSE = 1.7, β = .09	*F1* = .24, *p* = .63, MSE = .61, β = .08 *F2* = .22, MSE = .65, β = .08
Familiarity X Trope	*F1* = 6.2, *p* < .05, MSE = 36466, β = .68 *F2* = 1.3, *p* = .27, MSE = 25526, β = .20	*F1* = 4.4, *p* < .05, MSE = .57, β = .53 *F2* = 1.5, *p* = .23, MSE = .57, β = .22	*F1* = .12, *p* = .73, MSE = .12, β = .06 *F2* = .23, *p* = .63, MSE = 1.0, β = .08	*F1* = .59, *p* = .45, MSE = .61, β = .12 *F2* = .50, *p* = .46, MSE = .61, β = .11	*F1* = .10, *p* = .76, MSE = .20, β = .06 *F2* = .1, *p* = .78, MSE = .18, β = .05	*F1* = .07, *p* = .79, MSE = .11, β = .06 *F2* = .06, *p* = .78, MSE = .10, β = .05

Table 12

Statistical Analyses for Experiment 1 By Subject (F1) and By Item (F2) for Reading Time (in ms) Per Word Group, Fixation Count Per Word Group, Regression In Count Per Word Group, and Regression Out Count Per Word Group as a Function of Word (First, Middle, Last), Trope (Metaphor, Simile) and Familiarity (Familiar, Unfamiliar).

	Reading Time	Fixation Count	Regression In	Regression Out
	df = 2, 76	*df* = 2, 76	*df* = 2, 76	*df* = 2, 76
Word Main Effect	*F1* = 21.3, *p* < .01, MSE = 163228, β = 1.0 *F2* = 23.8, *p* < .01, MSE = 153787, β = 1.0	*F1* = 85.1, *p* < .01, MSE = 13.6, β = 1.00 *F2* = 72.8, *p* < .01, MSE = 13.6, β = 1.0	*F1* = 67.0, *p* < .01, MSE = 712, β = 1.0 *F2* = 150.9, *p* < .01, MSE = 712, β = 1.0	*F1* = 1.8, *p* = .17, MSE = 36, β = .37 *F2* = 4.3, *p* < .05, MSE = 36.0, β = .73
Word X Trope	*F1* = 1.5, *p* = .23 MSE = 11586, β = .31 *F2* = 1.6, *p* = .21, MSE = 10138, β = .32	*F1* = 2.4, *p* = .10, MSE = .38, β = .47 *F2* = 2.0, *p* = .14, MSE = .45, β = .36	*F1* = 6.8, *p* < .01, MSE = 72.7, β = .91 *F2* = .99, *p* = .38, MSE = 6.3, β = .22	*F1* = 1.4, *p* = .26, MSE = 26.9, β = .29 *F2* = .46, *p* = .64, MSE = 3.4, β = .12
Word X Familiarity	*F1* = 4.0, *p* < .05, MSE = 5073, β = .69 *F2* = .54, *p* = .56, MSE = 2040, β = .14	*F1* = .50, *p* = .61, MSE = .01, β = .13 *F2* = .09, *p* = .91, MSE = .01, β = .06	*F1* = 1.8, *p* = .17, MSE = 6.3, β = .37 *F2* = 15.5, *p* < .01, MSE = 72.7, β = .99	*F1* = .92, *p* = .40, MSE = 3.4, β = .20 *F2* = 3.2, *p* < .05, MSE = 26.9, β = .59
	df = 1, 38	*df* = 1, 38	*df* = 1, 38	*df* = 1, 38
Familiarity X Trope	*F1* = 6.2, *p* < .05 MSE = 12170, β = .68 *F2* = 1.3, *p* = .26, MSE = 8592, β = .20	*F1* = 4.4, *p* < .05, MSE = .19, β = .54 *F2* = 1.5, *p* = .23, MSE = .19, β = .22	*F1* = 3.5, *p* = .07, MSE = 10.8, β = .44 *F2* = 1.9, *p* = .18, MSE = 10.8, β = .27	*F1* = .02, *p* = .89 MSE = .07, β = .05 *F2* = .01, *p* = .92, MSE = .07, β = .05
	df = 2, 76	*df* = 2, 76	*df* = 2, 76	*df* = 2, 76
Word X Familiarity X Trope	*F1* = 1.3, *p* = .27, MSE = 1717, β = .28 *F2* = .21, *p* = .81, MSE = 799,	*F1* = .24, *p* = .78, MSE = .01, β = .09 *F2* = .04, *p* = .96, MSE = .01,	*F1* = .33, *p* = .72, MSE = 1.2, β = .10 *F2* = .18, *p* = .83, MSE = 1.2,	*F1* = .75, *p* =.48 MSE = 2.8, β = .17 *F2* = .37, *p* = .69, MSE = 2.8,

	$\beta = .08$	$\beta = .06$	$\beta = .08$	$\beta = .11$

Table 13

Statistical Analyses for Experiment 2 By Subject (F1) and By Item (F2) for Reading Time (in ms) for the Phrase, Fixation Count for the Phrase, Regression To Context Count, Regression From Last Sentence to Trope Count, and Regression Loop Count as a Function of Trope (Metaphor, Simile) and Familiarity (Familiar, Unfamiliar).

	Reading Time	Fixation Count	Regression to Context	Last to Phrase	Last Word to First Word	Regression Loops
	df = 1, 38	*df* = 1, 38	*df* = 1, 38	*df* = 1, 38	*df* = 1, 38	*df* = 1, 38
Familiarity Main Effect	*F1* = 27.0, *p* < .01, MSE = 224402, β = .99 *F2* = 2.9, *p* = .10, MSE = 74796, β = .38	*F1* = 13.8, *p* < .01, MSE = 2.3, β = .95 *F2* = 3.7, *p* = .06, MSE = 2.3, β = .46	*F1* = 0.0, *p* = 1.0, MSE = .00, β = .05 *F2* = 0.0, *p* = .84, MSE = .00, β = .05	*F1* = 1.2, *p* = .27, MSE = 2.5, β = .19 *F2* = .77, *p* = .39, MSE = 2.5, β = .14	*F1* = 2.8, *p* = .10, MSE = 6.1, β = .38 *F2* = 3.0, *p* = .10, MSE = 5.8, β = .40	*F1* = 1.0, *p* = .32, MSE = 1.5, β = .17 *F2* = 1.1, *p* = .40, MSE = 1.9, β = .20
Trope Main Effect	*F1* = 11.9, *p* < .01, MSE = 742473, β = .92 *F2* = 14.4, *p* < .01, MSE = 404701, β = .96	*F1* = 11.4, *p* < .01, MSE = 10.8, β = .91 *F2* = 4.7, *p* < .05, MSE = 3.1, β = .56	*F1* = 20.2, *p* < .01, MSE = 520.2, β =.99 *F2* = 28.3, *p* < .01, MSE = 520.2, β = .99	*F1* = 2.1, *p* = .16, MSE = 12.8, β = .29 *F2* = 3.3, *p* = .08, MSE = 12.8, β = .42	*F1* = 8.7, *p* < .01, MSE = 24.2, β = .82 *F2* = 8.5, *p* < .01, MSE = 23.1, β = .83	*F1* = 9.5, *p* < .01, MSE = 21.0, β = .85 *F2* = 9.1, *p* = < .01, MSE = 24.3, β = .83
Familiarity X Trope	*F1* = 2.88, *p* = .10, MSE = 23909, β = .38 *F2* = .20, *p* < .66, MSE = 5216, β = .07	*F1* = 2.1, *p* = .16, MSE = .35, β = .29 *F2* = .55, *p* = .46, MSE = .35, β = .11	*F1* = .20, *p* = .65, MSE = .80, β = .07 *F2* = .04, *p* = .84, MSE = .80, β = .06	*F1* = 1.2, *p* = .27, MSE = .25, β = .19 *F2* = .63, *p* = .43, MSE = 2.5, β = .12	*F1* = 1.9, *p* = .18, MSE = 4.1, β = .27 *F2* = 1.6, *p* = .20, MSE = 4.5, β = .22	*F1* = .69, *p* = .41, MSE = 1.0, β = .13 *F2* = .61, *p* = .49, MSE = 1.1, β = .11

Table 14

Statistical Analyses for Experiment 2 By Subject (F1) and By Item (F2) for Reading Time (in ms) Per Word Group, Fixation Count Per Word Group, Regression In Count Per Word Group, and Regression Out Count Per Word Group as a Function of Word (First, Middle, Last), Trope (Metaphor, Simile) and Familiarity (Familiar, Unfamiliar).

	Reading Time	Fixation Count	Regression In	Regression Out
	df = 2, 76	*df* = 2, 76	*df* = 2, 76	*df* = 2, 76
Word Main Effect	*F1* = 17.2, *p* < .01, MSE = 104185, β = 1.0 *F2* = 4.5, *p* < .05, MSE = 25699, β = .75	*F1* = 55.3, *p* < .01, MSE = 9.5, β = 1.0 *F2* = 50.1, *p* < .01, MSE = 7.9, β = 1.0	*F1* = 55.4, *p* < .01, MSE = 516, β = 1.0 *F2* = 160.9, *p* < .01, MSE = 1418.8, β = 1.0	*F1* = 6.8, *p* < .01, MSE = 72.1, β = .91 *F2* = 23.1, *p* < .01, MSE = 286.9, β = 1.0
Word X Trope	*F1* = 5.8, *p* < .01, MSE = 34910, β = .86 *F2* = 16.7, *p* < .01, MSE = 95636, β = 1.0	*F1* = 11.0, *p* < .01, MSE = 1.9, β = .99 *F2* = 17.7, *p* < .01, MSE = 2.8, β = 1.0	*F1* = 6.6, *p* < .01, MSE = 61.1, β = .90 *F2* = 1.89, *p* = .16, MSE = 30.5, β = .38	*F1* = 8.4, *p* < .01, MSE = 89.6, β = .96 *F2* = 3.0, *p* = .06, MSE = 35.1, β = .56
Word X Familiarity	*F1* = 1.1, *p* = .33, MSE = 1953, β = .24 *F2* = .54, *p* = .59, MSE = 3652, β = .14	*F1* = 4.4, *p* < .05, MSE = .12, β = .74 *F2* = .34, *p* = .71, MSE = .05, β = .10	*F1* = 4.6, *p* < .05, MSE = 18.8, β = .76 *F2* = 36.6, *p* < .01, MSE = 322.8, β = 1.0	*F1* = 10.4, *p* < .01, MSE = 53.0, β = .99 *F2* = 26.4, *p* < .01, MSE = 328.1, β = 1.0
	df = 1, 38	*df* = 1, 38	*df* = 1, 38	*df* = 1, 38
Familiarity X Trope	*F1* = 2.9, *p* = .10, MSE = 7912, β = .38 *F2* = .20, *p* = .66, MSE = 1739, β = .07	*F1* = 2.1, *p* = .16, MSE = .12, β = .29 *F2* = .54, *p* = .47, MSE = .11, β = .11	*F1* = 3.2, *p* = .08, MSE = 13.1, β = .41 *F2* = 3.0, *p* = .09, MSE = 61.9, β = .39	*F1* = 2.0, *p* = .17, MSE = 7.0, β = .28 *F2* = 1.4, *p* = .24, MSE = 15.4, β = .21
	df = 2, 76	*df* = 2, 76	*df* = 2, 76	*df* = 2, 76
Word X Familiarity X Trope	*F1* = .88, *p* = .42, MSE = 1525, β = .20 *F2* = .46, *p* = .63, MSE =	*F1* = 2.2, *p* = .11, MSE = .06, β = .44 *F2* = .54, *p* = .58, MSE = .07,	*F1* = .69, *p* = .49, MSE = 2.8, β = .16 *F2* = .34, *p* = .65, MSE = 5.5,	*F1* = 2.1, *p* = .13 MSE = 10.5, β = .41 *F2* = .42, *p* = .66, MSE = 4.9,

	3163, $\beta = .12$	$\beta = .14$	$\beta = .10$	$\beta = .12$

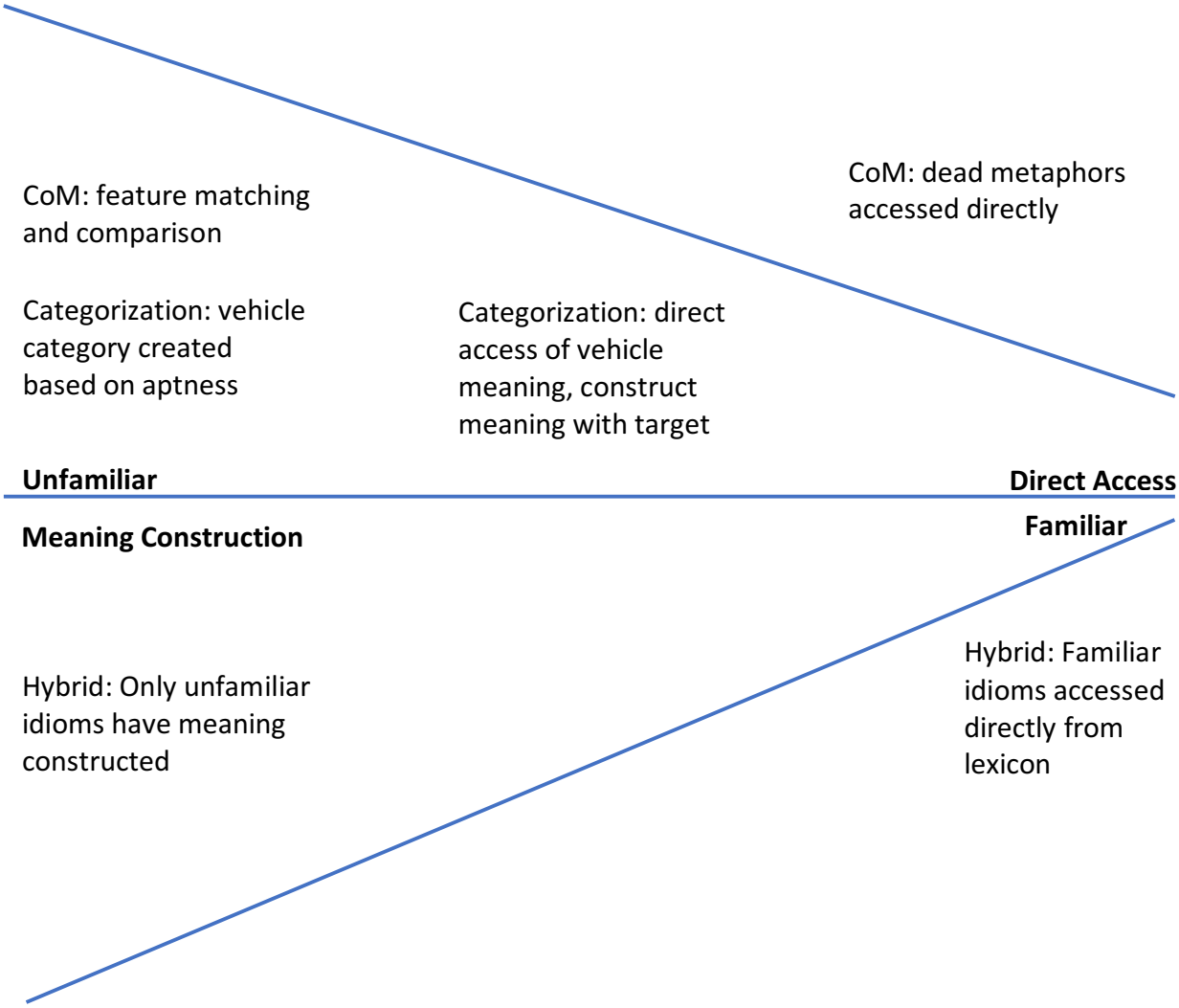

Figure 1. Figurative Funnel. The goal of this figure is to demonstrate the idea that at the unfamiliar end, the models predict a number of different processes used to construct meaning. However, as tropes become more familiar, the models make similar predictions based on the process of directly accessing the trope's meaning.

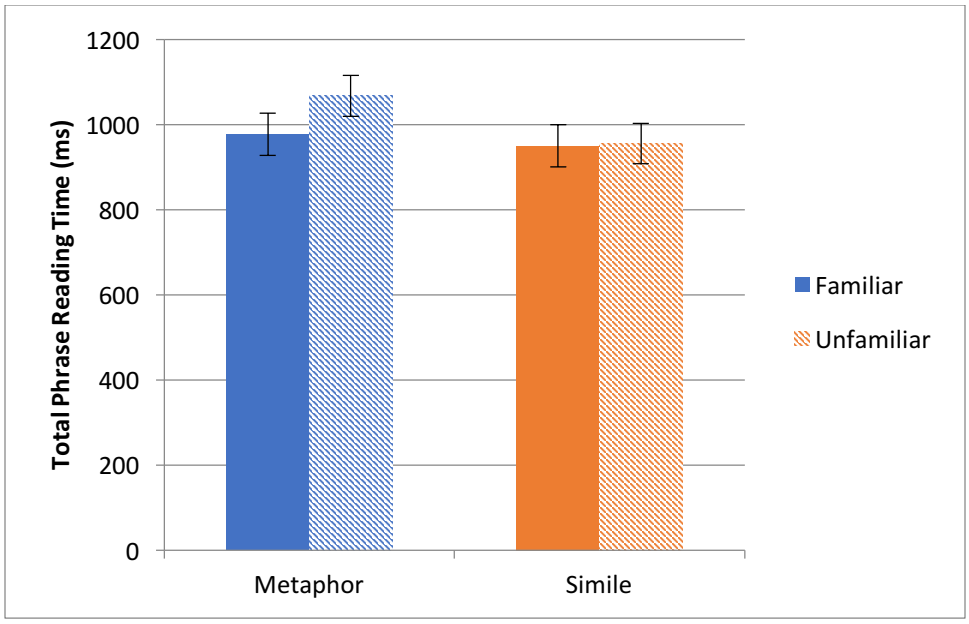

Figure 2. Experiment 1 total reading time of the phrase as a function of trope (metaphor, simile) and familiarity (familiar, unfamiliar).

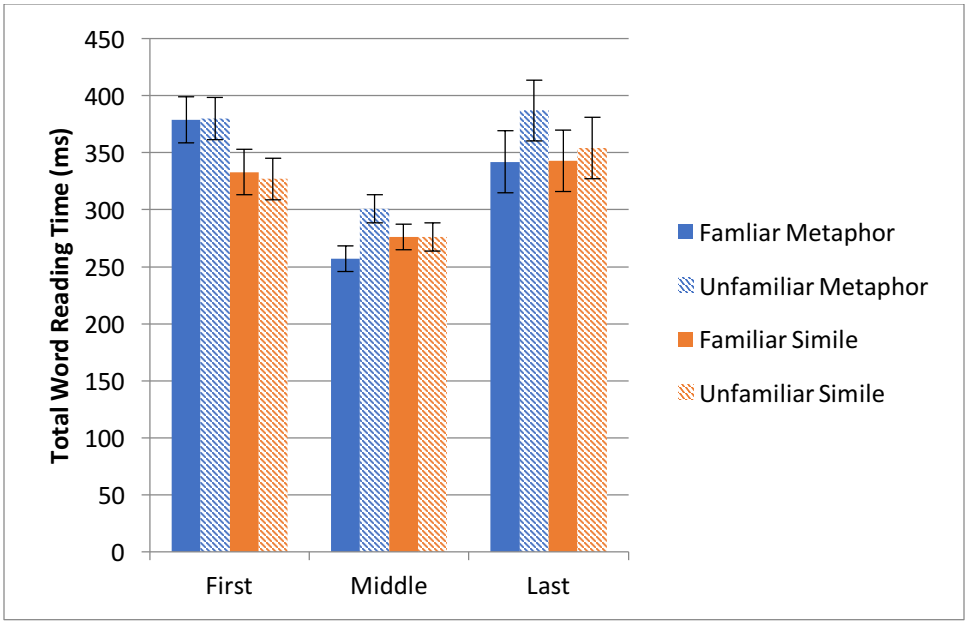

Figure 3. Experiment 1 total reading time of the words as a function of word (first, middle, last), trope (metaphor, simile) and familiarity (familiar, unfamiliar).

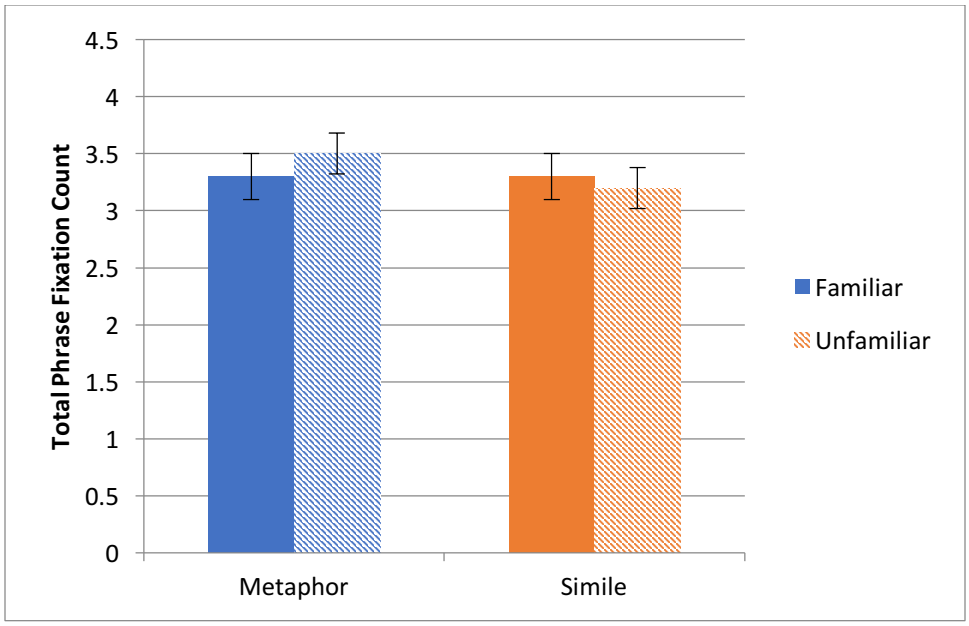

Figure 4. Experiment 1 total fixation count of the phrase as a function of trope (metaphor, simile) and familiarity (familiar, unfamiliar).

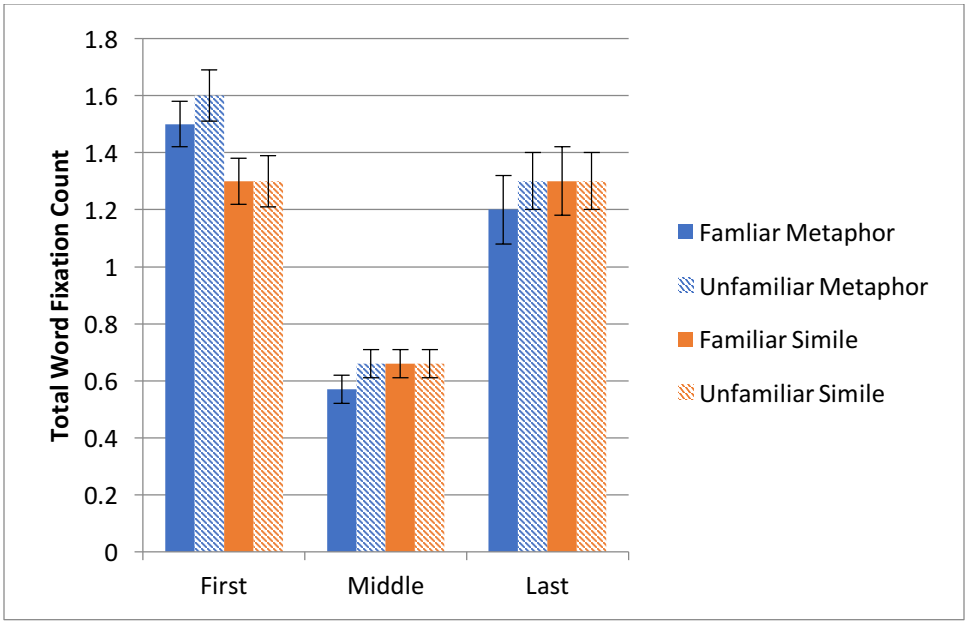

Figure 5. Experiment 1 total fixation count of the words as a function of word (first, middle, last), trope (metaphor, simile) and familiarity (familiar, unfamiliar).

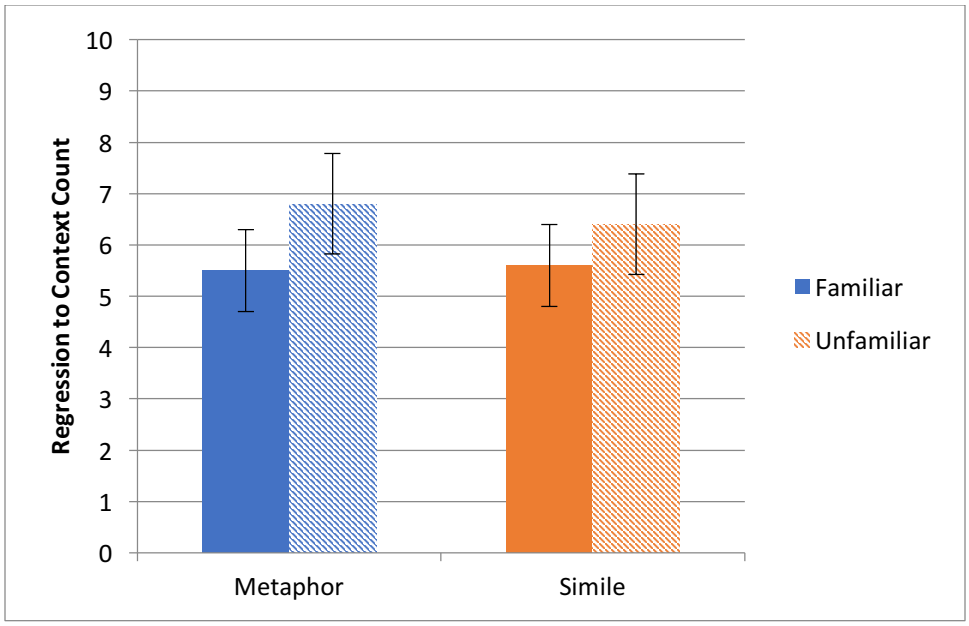

Figure 6. Experiment 1 total regressions to the context count as a function of trope (metaphor, simile) and familiarity (familiar, unfamiliar).

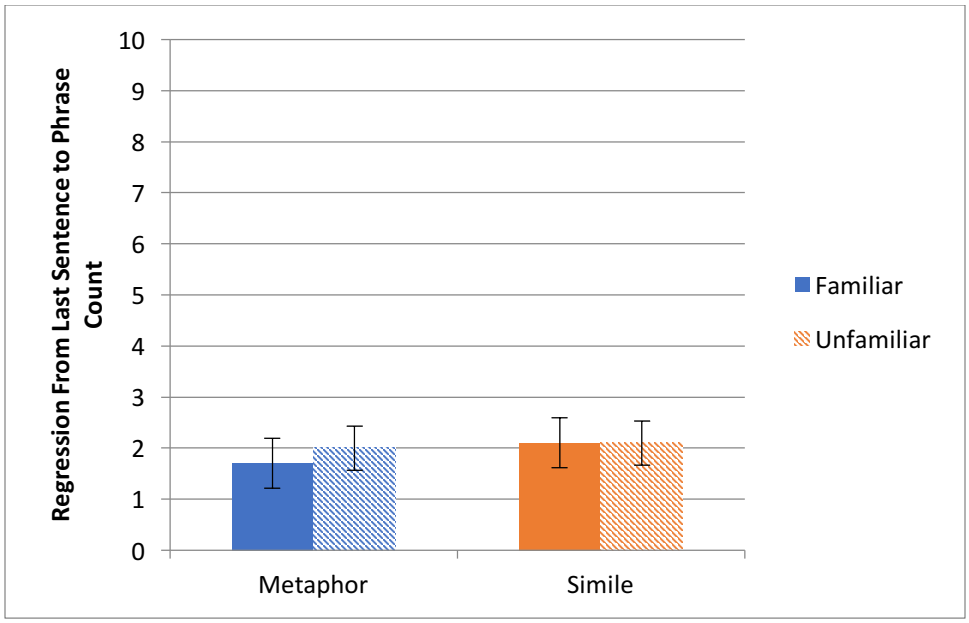

Figure 7. Experiment 1 total regressions from the last sentence to the phrase as a function of trope (metaphor, simile) and familiarity (familiar, unfamiliar).

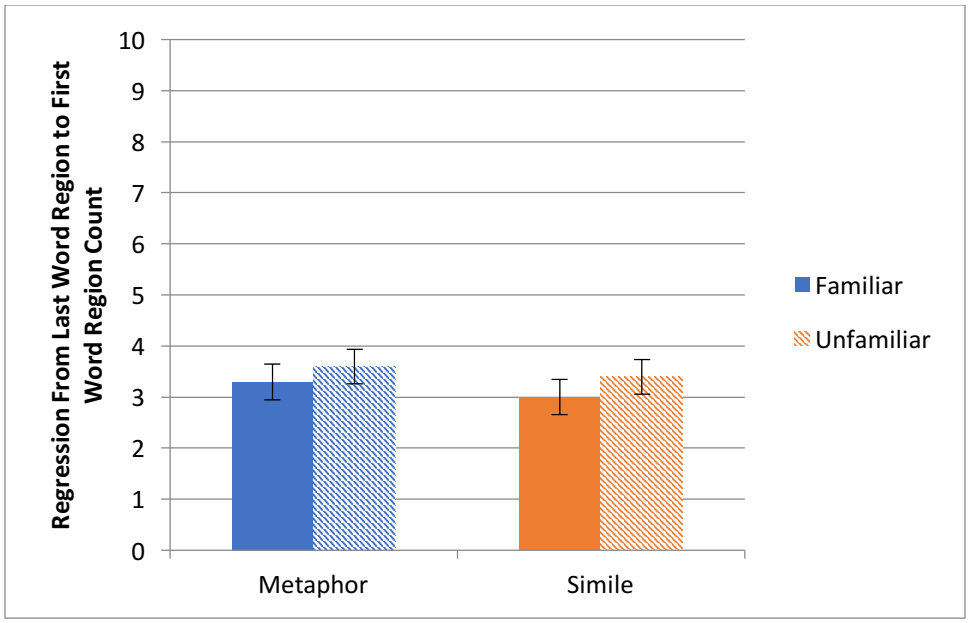

Figure 8. Experiment 1 total regressions from the last word region to the first word region as a function of trope (metaphor, simile) and familiarity (familiar, unfamiliar).

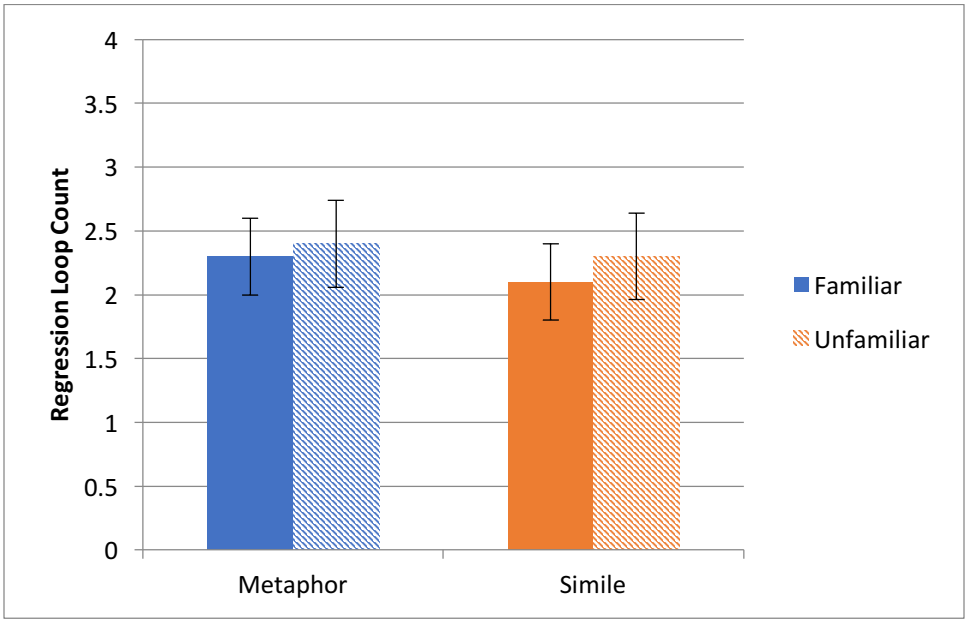

Figure 9. Experiment 1 total regression loop count as a function of trope (metaphor, simile) and familiarity (familiar, unfamiliar).

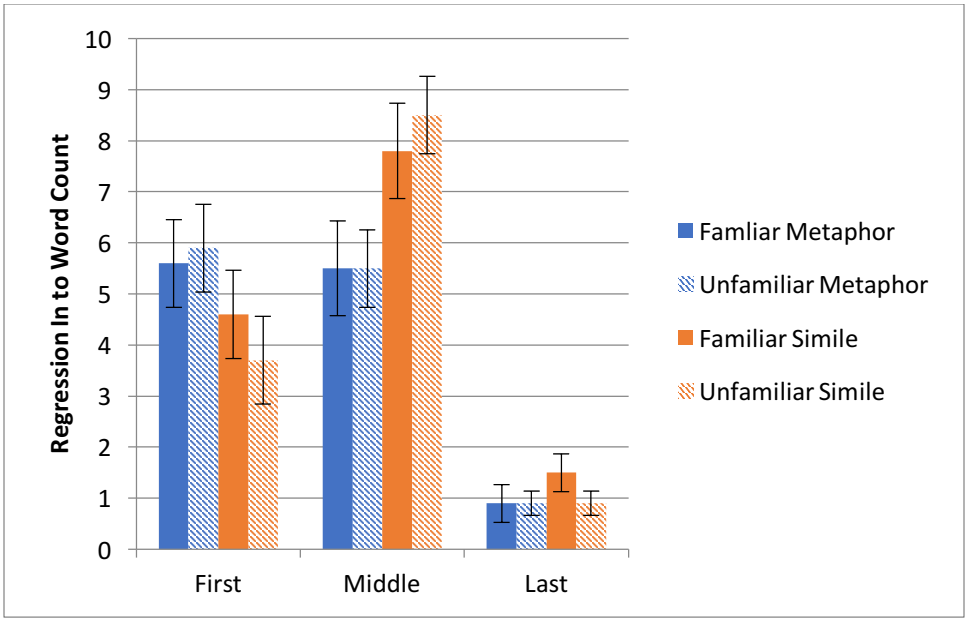

Figure 10. Experiment 1 total regressions in count of the words as a function of word (first, middle, last), trope (metaphor, simile) and familiarity (familiar, unfamiliar).

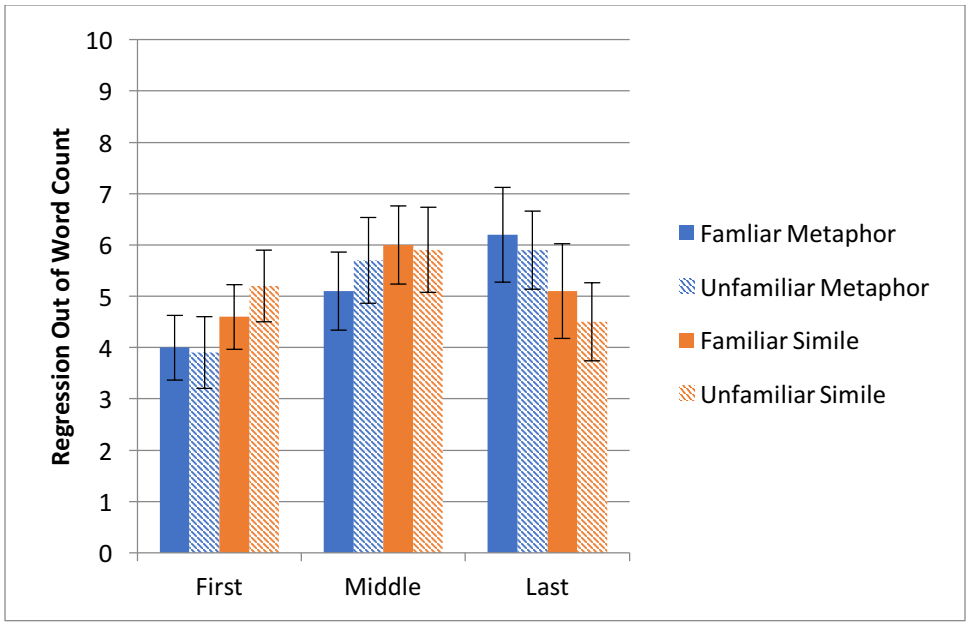

Figure 11. Experiment 1 total regressions out count of the words as a function of word (first, middle, last), trope (metaphor, simile) and familiarity (familiar, unfamiliar).

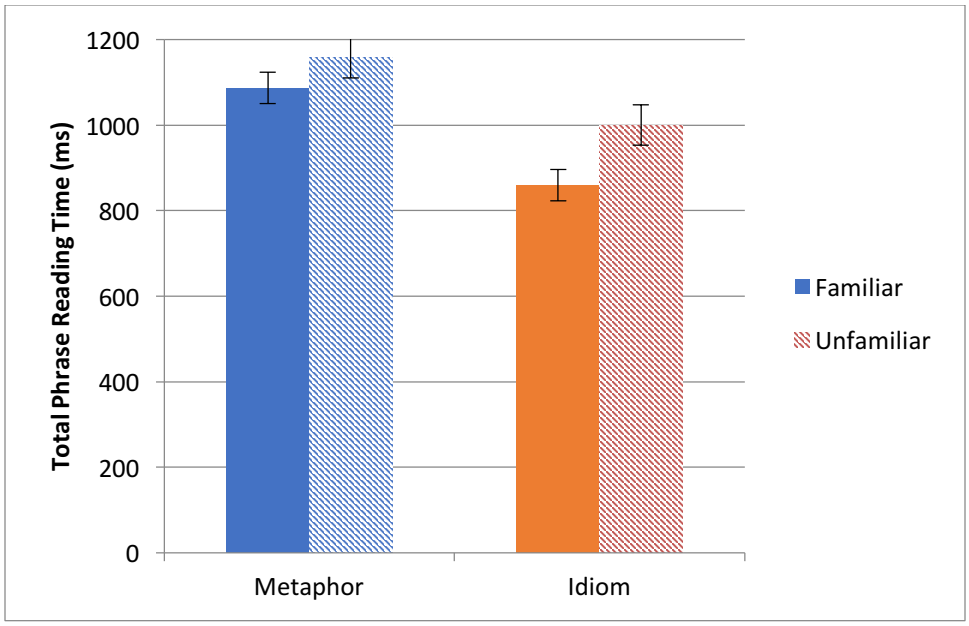

Figure 12. Experiment 2 total reading time of the phrase as a function of trope (metaphor, idiom) and familiarity (familiar, unfamiliar).

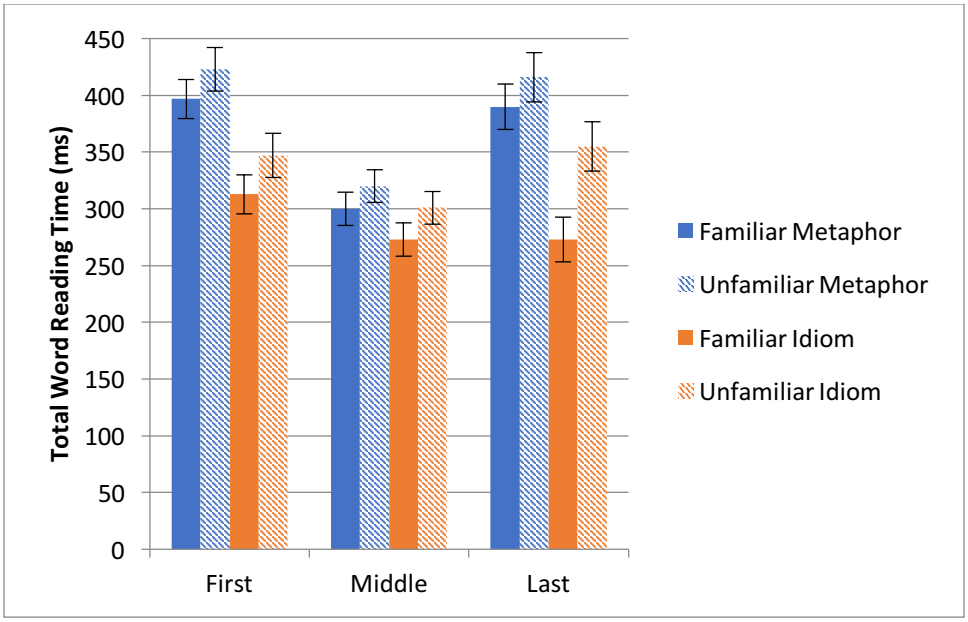

Figure 13. Experiment 2 total reading time of the words as a function of word (first, middle, last), trope (metaphor, idiom) and familiarity (familiar, unfamiliar).

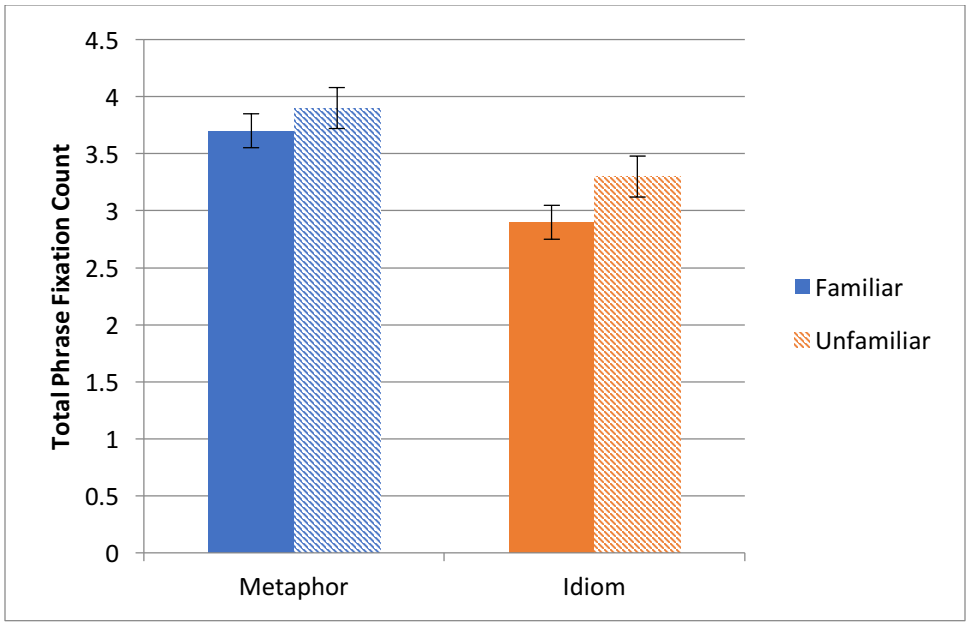

Figure 14. Experiment 2 total fixation count of the phrase as a function of trope (metaphor, idiom) and familiarity (familiar, unfamiliar).

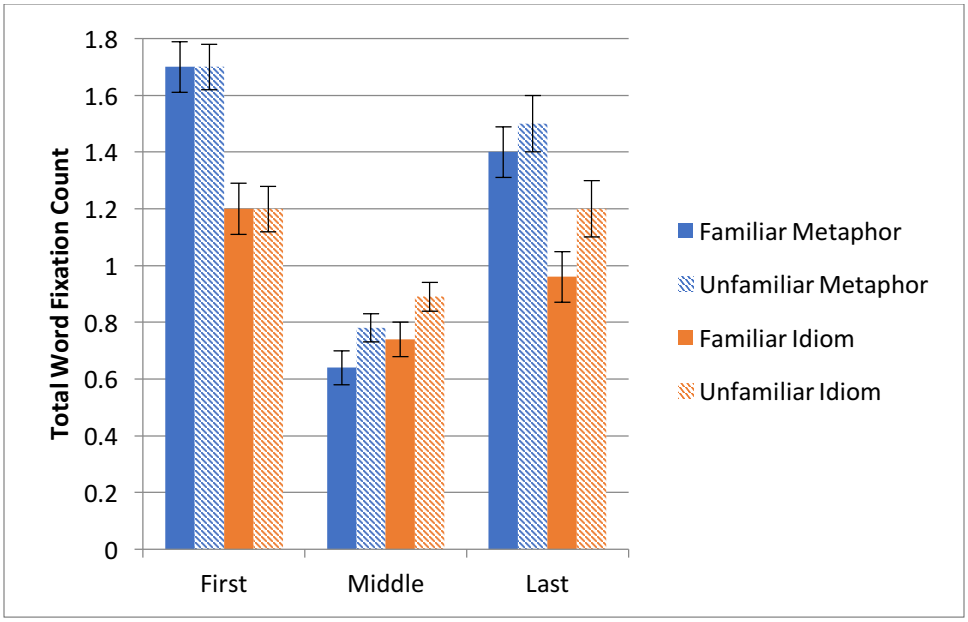

Figure 15. Experiment 2 total fixation count of the words as a function of word (first, middle, last), trope (metaphor, idiom) and familiarity (familiar, unfamiliar).

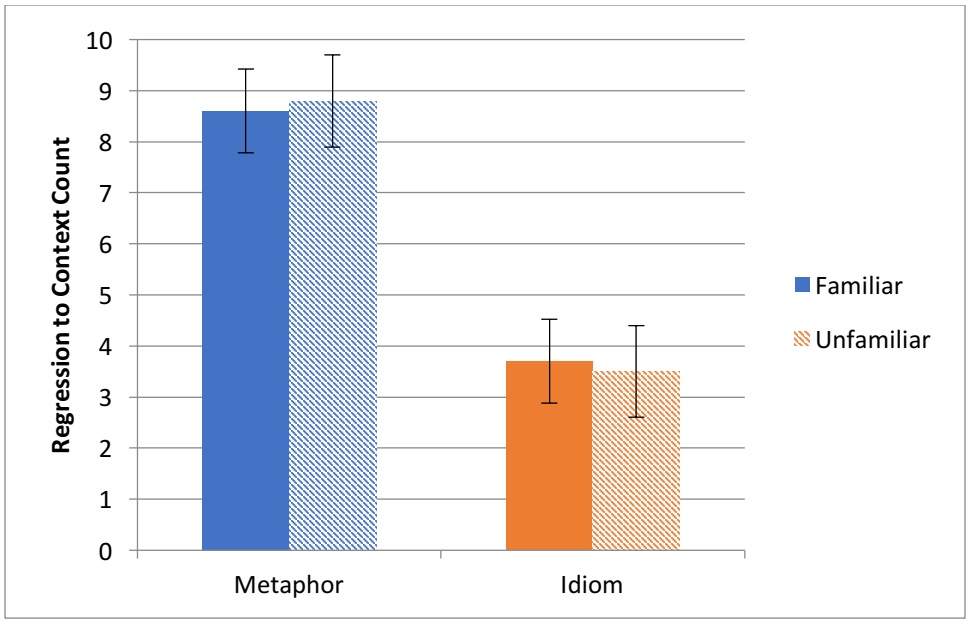

Figure 16. Experiment 2 total regressions to the context count as a function of trope (metaphor, idiom) and familiarity (familiar, unfamiliar).

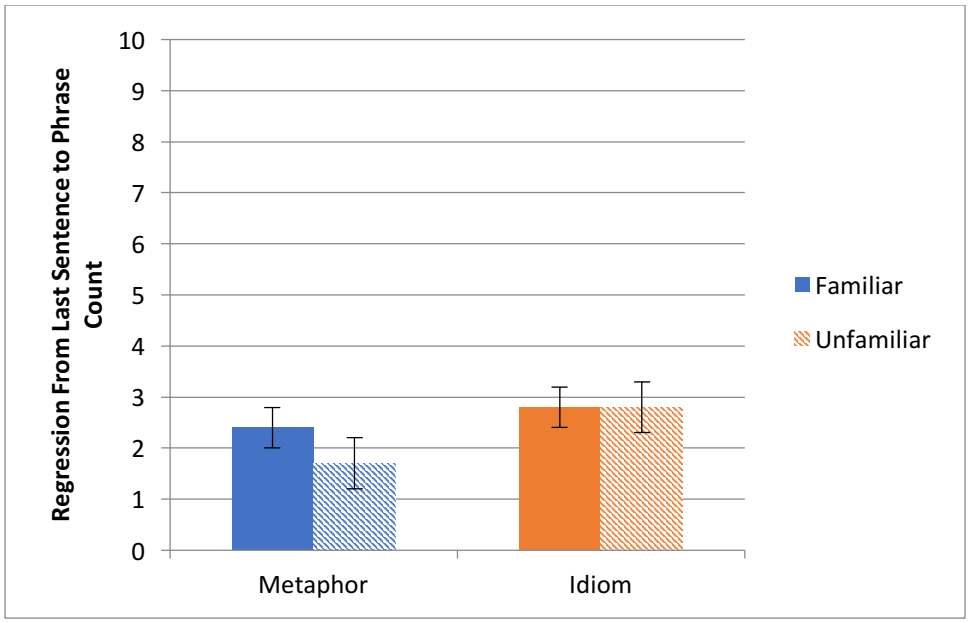

Figure 17. Experiment 2 total regressions from the last sentence to the phrase count as a function of trope (metaphor, idiom) and familiarity (familiar, unfamiliar).

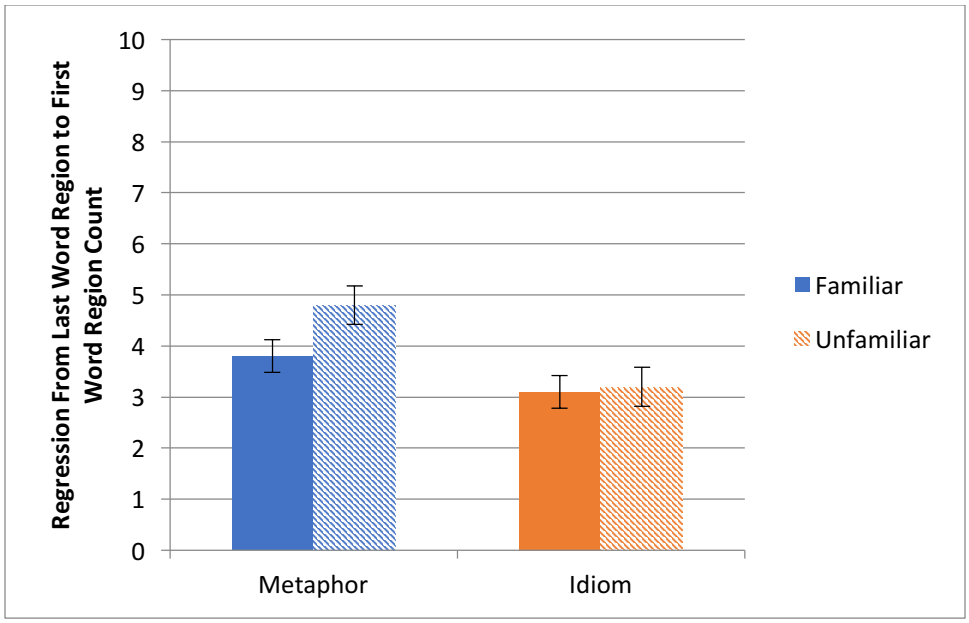

Figure 18. Experiment 2 total regressions from the last word region to the first word region count as a function of trope (metaphor, idiom) and familiarity (familiar, unfamiliar).

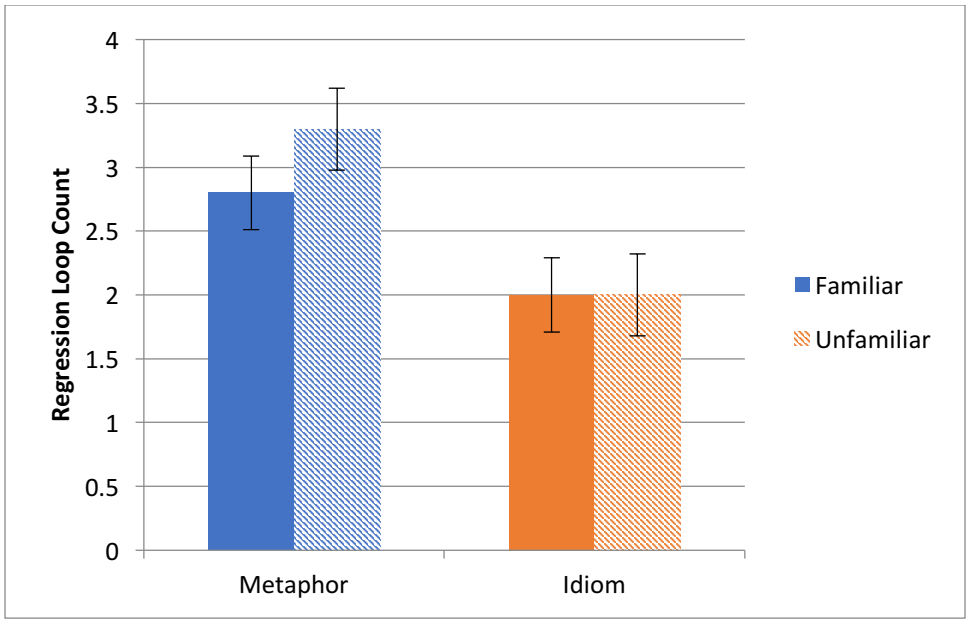

Figure 19. Experiment 2 total regression loop count as a function of trope (metaphor, idiom) and familiarity (familiar, unfamiliar).

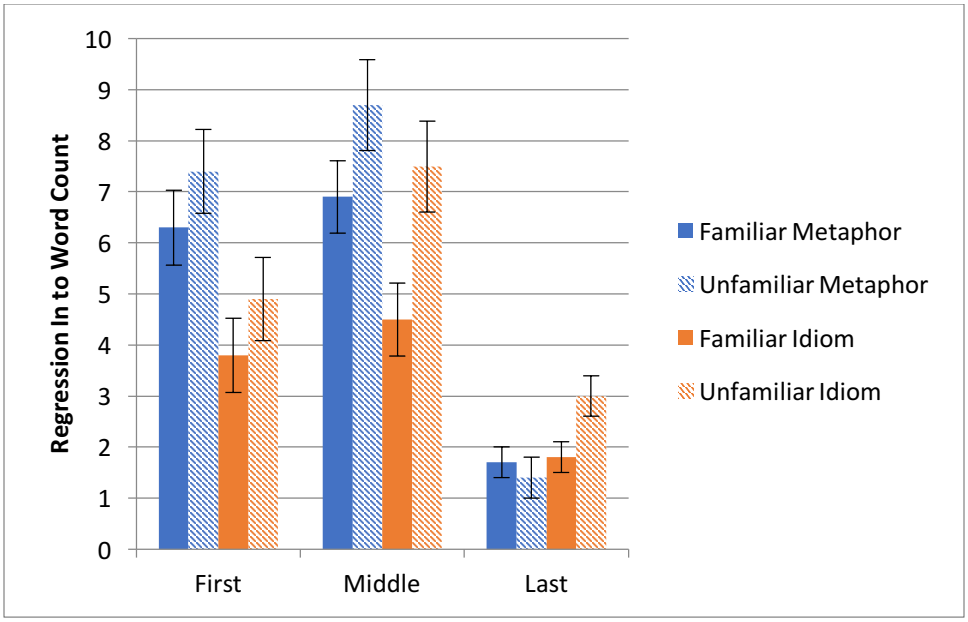

Figure 20. Experiment 2 total regressions in count of the words as a function of word (first, middle, last), trope (metaphor, idiom) and familiarity (familiar, unfamiliar).

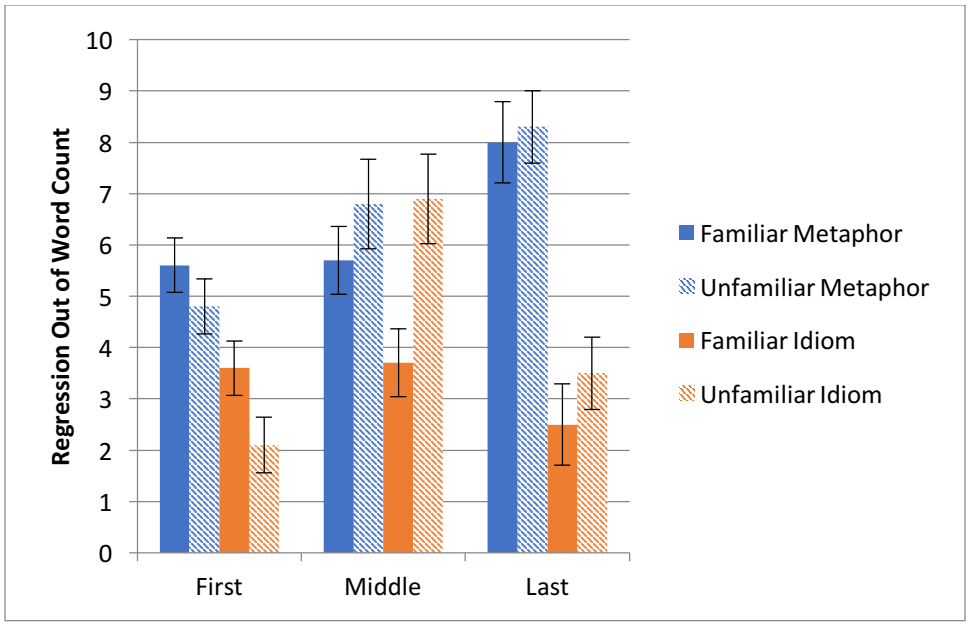

Figure 21. Experiment 2 total regressions out count of the words as a function of word (first, middle, last), trope (metaphor, idiom) and familiarity (familiar, unfamiliar).

Appendix A

There are a number of eye movement measures that can be assessed, and so it is important to determine which will be most interesting and will have the largest theoretical implications. Below I will describe the measures that I believe are most important when studying familiarity effects in figurative language comprehension. Definitions of the measures will be illustrated using the hypothetical fixation patterns below. The numbers below the words indicate fixation locations.

Sample fixation pattern 1 (metaphor = mind is a sponge):

Amanda's 3-year-old niece learns things easily. Young
1 2 3 4 6 5 7 8 9 10 11

children are able to absorb new information quickly.
12 13 14 15 16 17 18 19 20
26

Amanda says the *mind is a sponge*. Her niece is
21 22 23 24 27 28 29
25 30

very adorable.
31 32

Sample fixation pattern 2 (idiom = *bull in a china shop*)

Roger felt like a *bull in a china shop*.
1 2 3 4 5 6 7 8
9 10

Total Reading Time: The amount of time spent on an area of interest, including refixations, rereading, and regressions back onto the area. Total reading time of the metaphor phrase *mind is a sponge* in sample 1 would include fixations 23, 24, 25, 27, and 30. Reading time will be measured both as reading time of the entire figurative phrase and then of the individual words

within the phrase. In order to create a consistency in the number of individual words, measures will be collected on the First Word of the phrase, the Last Word of the phrase, and all words in between will be grouped into the Middle Words. Reading time reflects the level of effort required to understand and integrate a word in a text. Therefore, differences in reading times at both the phrase and word level will be indicative of how difficult it is for the reader to either access or create the correct meaning of the figurative trope.

Total Fixation Count: The number of times a reader stops (fixates) on an area of interest, including refixations, rereading, and regressions back onto the area. Total fixation count for the metaphor phrase *mind is a sponge* in sample 1 would include fixations 23, 24, 25, 27, and 30. Fixation count will also be measured as both fixation count for the phrase and fixation count for the individual words using the same grouping of words (First, Middle, Last) as for total reading time. Fixation count and reading time are highly related with one another, with typically more fixations resulting in longer reading times. When readers refixate a word, it is oftentimes a sign that they are having difficulty understanding the meaning of the word, or are working on integrating it into the text based on what they have read. Therefore, fixation count will be highly dependent on how easily readers are able to create or access the meaning of the trope.

Regression to the Context: When the reader is within the figurative phrase and their next fixation is to a point prior to the phrase, it is considered a regression to the context. The regression needed to originate from the trope, and end in a word that was at least two words prior to the first word in the phrase. This was done in order to make sure that the regression was made to the context intentionally, and not the result of a regression intended to the first word. A regression to the context in sample 1 would be the saccadic movement from fixation 25 to fixation 26. This phrase-level measure will help determine how often readers need to return to

the context for clues, or to consolidate a constructed meaning with the context. As an example, if unfamiliar idioms require contextual clues to be understood, then we would expect to find more regressions to the context in unfamiliar idioms than other less context-dependent tropes.

Regression From the Last Sentence to the Phrase: This occurs when the reader has gone past the figurative phrase into the final contextual sentence, and then returns from that sentence back to the figurative phrase. This regression needed to originate from two words after the last word of the trope, and end in any word in the phrase. This was done in order to make sure the regression was coming from the last sentence intentionally, and returning to the phrase. A regression from the last sentence to the phrase would be the saccadic movement from fixation 29 to fixation 30. This phrase-level measure is used to determine if readers move forward into the context looking for more clues for understanding the trope, only to return to the trope to finalize processing. Similar to the regression to context example, if unfamiliar idioms require contextual clues then the reader may reader into the last sentence of the context in search of an explanation of the phrase, only to return after. This would occur more often in unfamiliar idioms than in less context-dependent tropes.

Regression From the Last Word Region to the First Word Region: This is a specific pattern in which the reader leaves the last word region of the trope and regresses back to the first word of trope. Last word region is defined as the last word in the phrase and the word immediately prior to it. First word region is defined as the first word in the phrase and the word immediately following it. Because the middle words of the tropes are oftentimes shorter words (e.g. is, or a), they are still able to read the adjacent words. Therefore, we look at regressions to and from regions instead of the individual words. This pattern represents the reader consolidating a trope's meaning while keeping their attention within the trope, rather than looking to the context.

As an example, this may be related to the category assertion process described in the categorization model, in which the reader asserts the characteristics of the vehicle word onto the target word.

Regression Loop: After a reader has made a regression from the last word region to the first word region, if their next fixation is on the last word region it is considered a regression loop. For example, fixate Vehicle, fixate Target, fixate Vehicle. A regression loop in sample 2 is the series of fixations going from 8 to 10. This measure further represents the reader focusing their efforts within the trope rather than outside of it (before or after the trope), and may be reflective of some meaning construction processes. As an example, this may be related to the comparison process described in the Career of Metaphor used for unfamiliar metaphors, in which the reader is regressing between the target and vehicle words as they construct the meaning.

Regression In to a Word: Regressions can be counted to a specific word, rather than a region as described above. These are marked as regressions into an individual word. A regression in to a word in sample 2 is the saccadic movement from fixation 8 to fixation 9, which would be considered a regression in to the word *bull*.

Regression Out of a Word: Similar to regressions into a word, these are measures in which the reader is leaving a specific word and regressing to a previous point in the text. A regression out of a word in sample 2 is the saccadic movement from fixation 8 to fixation 9, which would be considered a regression out of the word *china*.

There are many possible measures when doing eye tracking research, such as first and second pass, but not all of them of them can be included for a number of reasons. First, doing so would add to an already lengthy set of analyses, and would have a negative impact on interpretation of the results. Second, these measures have either been found to be non-significant in the past, or

are not of theoretical importance. For example, in previous work (Campbell, 2014), I found no consistent familiarity effects during first and second pass, and instead focused on total time.

Appendix B

Familiar Metaphors/Similes (Similes were created by adding *like* to a metaphor)
Amanda's 3-year-old niece learns things easily. Young children are able to absorb new information quickly. Amanda says the *mind is (like) a sponge*. Her niece is very adorable.

Cancer is a very serious illness to have. It is a disease that will grow and spread uncontrollably throughout somebody's body. In many ways, a *tumor is (like) a plague*. Cancer awareness remains a very important issue to this day.

Unfamiliar Metaphors/Similes
The crew had set sail for three weeks now. They began to cast out their nets hoping to catch their prey. The captain said a *fisherman is (like) a spider*. It would be another two weeks before they would make port.

Steven saw some dark clouds looming ahead. He knew that something bad was brewing. Steven thought that a *storm is (like) a coffeepot*. In a few minutes it began to rain heavily.

Familiar Idioms
It was the weekend after graduation and Becky was hosting a party. She invited all of her closest friends to celebrate. The party was a great success and Becky *had a ball*. She was glad she had met so many incredible people while at school.

Gavin was a small town boy who recently competed in a televised singing competition where he won second place. He was constantly bragging about his accomplishments. His friends know he likes to *toot his own horn* all the time. Gavin was a talented singer.

Unfamiliar Idioms
Andrew had been called into his boss's office for his end of the year evaluation. Looking over the files, Andrew started getting a bit worried. His boss was upset and told him that this year Andrew did not *cut the mustard*. Andrew was given a second chance to start performing better.

The local teachers union was considering a strike. Days before they planned on voting on a strike, news came of the school budget being cut again this year. Around here, this *was par for the course*. The teachers decided to strike.

Appendix C

Language History Questionnaire

Subject # _____ Sex _____ Age _____ What country were you born in? _____

Years living in U.S. _____ Years in U.S. Schools _____

(1) What is the FIRST language you spoke? If your parents spoke two languages to you, list BOTH languages.

(2) List from MOST fluent to LEAST fluent all of the languages that you know (write on the back of this page if you need more space). Note that the language you learned first is not necessarily the language you now know best. Specify the age at which you began to learn the language (if it is your native language you should specify age as "birth") and where you learned it (e.g., school, home, church).

	Language	Age learned	Location learned
Most fluent	_____	_____	_____
	_____	_____	_____
	_____	_____	_____
Least fluent	_____	_____	_____

(3) Answer the following questions. Complete only those questions that apply to you.
At what age did you begin speaking **English**? _____
At what age did you begin reading **English**? _____
At what age did you begin speaking your most fluent language **OTHER THAN** English? _____
At what age did you begin reading your most fluent language **OTHER THAN** English? _____

(4) Complete the following ratings. If you think you are more proficient in either English or your OTHER language, your ratings should reflect this difference. Answer only those questions that apply to you.

	NOT fluent									VERY fluent
For ENGLISH:										
How fluent are you in speaking?	1	2	3	4	5	6	7	8	9	10

How fluent are you in understanding?	1	2	3	4	5	6	7	8	9	10
How fluent are you in reading?	1	2	3	4	5	6	7	8	9	10

For your most fluent language OTHER THAN English:

How fluent are you in speaking?	1	2	3	4	5	6	7	8	9	10
How fluent are you in understanding?	1	2	3	4	5	6	7	8	9	10
How fluent are you in reading?	1	2	3	4	5	6	7	8	9	10

Vocabulary Test (Version 6/09/2004) **Subject_____**

Directions: Choose the BEST definition for each word.

1. **ASCEND**
   A. to go up or mount
   B. consent
   C. improve with time
   D. to leave behind
   E. to replace a leader

2. **WARY**
   A. tired out
   B. rude; uncouth
   C. perturbed
   D. brand-new
   E. cautious; careful

3. **NURTURE**
   A. helped by man
   B. to feed or nourish
   C. to educate
   D. to protect by nature
   E. to cook

4. **INFINITESIMAL**
   A. very long
   B. very slow
   C. well defined
   D. uncompromising
   E. very small

5. **BELLIGERENT**
   A. informative
   B. blunt
   C. tiring
   D. war-like
   E. pro-active

6. **INDIFFERENT**
   A. similar
   B. unconcerned
   C. diffident
   D. solicitous
   E. opposite

7. **PERJURE**

A. to save from indignity
B. to improve or rectify
C. to demand support
D. to lie under oath
E. day by day

8. VERBOSE
A. slow
B. impressive
C. complicated
D. wordy
E. meaningless

9. OPAQUE
A. transparent
B. slippery
C. impenetrable by light
D. gem-like
E. financially well-off

10. SYNTHESIS
A. musical rendition of a written work
B. a theory of immoral behavior
C. the combination of parts to form a whole
D. watching or guarding
E. properties of artificial chemicals

11. SPONTANEITY
A. unwanted laughter
B. uncontrollable danger
C. unplanned action
D. unneeded socialism
E. stand-up attitude

12. VALIDATE
A. to prove
B. to get paid back
C. to expire
D. to run away
E. to complete successfully

13. SUBORDINATE
A. to hypothesize in abstract
B. to practice with instruction
C. to levy upon others
D. to go on vacation

E. to rank in importance

14. MEAGER
A. not full, inadequate
B. to beg
C. without self-respect
D. in good shape, healthy
E. wise, full of advice

15. EQUIVOCAL
A. premier, establishing new precedent
B. popular, known by everyone
C. exciting, causing a commotion
D. peculiar, one of a kind
E. uncertain, having two meanings

16. REBUKE
A. to dispute
B. poor reputation
C. to scold harshly
D. to stop at midpoint
E. to overfill

17. ECLECTIC
A. providential
B. of religious origins
C. purified
D. out of fashion
E. from various sources

18. TERSE
A. concise
B. private
C. angry
D. outdated
E. harsh-sounding

19. ILLUSORY
A. bright
B. deceptive
C. unhealthy
D. making a reference to
E. sometimes friendly, sometimes undependable

20. DIVULGE

A. to discourage
B. to pay for
C. to turn away
D. to reveal
E. to infiltrate

21. REPROVE
A. to reverse an argument
B. to be clean of
C. to express disapproval
D. to grovel for forgiveness
E. to encourage hope

22. IMPLAUSIBLE
A. could happen at any moment
B. not believable
C. unyielding
D. considered tactless
E. to serve or worship

23. INCONTROVERTIBLE
A. useless
B. prone to trouble making
C. indisputable
D. successful
E. unprotected

24. QUERY
A. excavation
B. prey
C. inquiry
D. strange occurrence
E. strange, odd

25. DISPERSE
A. to seize one's assets
B. to live in exile
C. to break up and scatter
D. to weaken connections
E. to make vacant

26. VACILLATE
A. to prepare for action; lubricate
B. to show indecision; to waver
C. to hold firmly, to be stubborn

D. to wait until the last second, delay
E. to scatter; to create chaos

27. SUPERFLUOUS
A. gay, happy
B. reserved, waiting
C. trivial; unimportant
D. unnecessary; excessive
E. undecided; variable

28. AUTONOMOUS
A. unknown identity
B. having many names
C. uncontrollable
D. independent existence
E. self-confidence

29. PRECEDENT
A. an expectation
B. most important event
C. a leader
D. a prior occurrence
E. a forgotten time

30. BOLSTER
A. to disagree, strongly
B. to defend, proudly
C. to reinforce, strengthen
D. to agonize, repeatedly
E. brutalize, mercilessly

Curriculum Vitae
Spencer J. Campbell

CONTACT INFORMATION

2912 ½ 5 ½ Avenue
Rock Island, IL 61201

Telephone: 630-730-0603
Email: spencercampbell@augustana.edu

EDUCATION

B.A. May 2010. Augustana College.
Major: Psychology Minor: Mathematics

M.A. May 2014. University of Illinois at Chicago.
Major: Cognitive Psychology

PhD Expected December 2017. University of Illinois at Chicago.
Major: Cognitive Psychology Minor: Cognitive Science

HONORS AND AWARDS

Upshaw Teaching Excellence Award, UIC
Augustana College Summer Research Fellowship
Psi Chi Regional Research Award
Wallace-Wallin Outstanding Psychology Student Scholarship
Augustana College Outstanding Psychology Senior Award
Ralph Waldo Hansen Excellence in Psychology Scholarship

TEACHING EXPERIENCE

Teaching Fellow - Augustana College

FYI 101 - First Year Inquiry, Rhetoric and the Liberal Arts: Fall 2017 (2 sections)
PSYC 100 - Introduction to Psychology: Winter 2017 (2 sections), Spring 2018 (2 sections)
PSYC 246 - Research Methods in Psychology: Spring 2018

Instructor - University of Illinois at Chicago

PSCH 242 - Introduction to Research in Psychology: Spring 2016, Summer 2016
PSCH 353 - Laboratory in Cognition and Memory: Spring 2017

Teaching Assistant - University of Illinois at Chicago

PSCH 353 - Laboratory in Cognition and Memory: Spring 2011, Fall 2013, Spring 2013, Fall 2014, Spring 2014, Fall 2015, Spring 2015, Fall 2016 – Instructor: Dr. Gary Raney
PSCH 242 - Introduction to Research in Psychology: Fall 2011, Spring 2012, Spring 2013 Instructor: Dr. Evelyn Behar
PSCH 303 - Writing in Psychology: Summer: 2012, Fall 2012 - Instructor: Dr. Eric Gobel

PUBLICATIONS

Campbell, S., & Raney, G. (2015). A 25-year replication of Katz et al.'s metaphor norms. *Behavior Research Methods, 48*(1), 330-340.

Raney, G., Campbell, S., & Bovee, J. (2014). Using eye movements to evaluate the cognitive processes involved in text comprehension. *Journal of Visualized Experiments,* (83), e50780, doi:10.3791/50780 (2014).

Corts, D., & Campbell, S. (2009). Novel metaphor extensions in political satire. In Barnden, Lee, Littlemore, Moon, Philip, & Wallington (Eds.), *Corpus-based approaches to figurative language.* University of Birmingham.

CONFERENCE PRESENTATIONS

Campbell, S., & Raney, G. (2016). Familiarity and context have independent effects on metaphor reading times. Presented at Annual Meeting of the Psychonomic Society 2016, Boston, MA.

Campbell, S., & Raney, G. (2016). The role of language background on providing norms for metaphors. Presented at Midwestern Psychological Association 2016, Chicago, IL.

Raney, G., Roy-Charland, A., Bovee, J., Miller, K., & Campbell, S. (2015). Determining what makes a text difficult to comprehend is difficult. Presented at Annual Meeting of the Psychonomic Society 2015, Chicago, IL.

Raney, G., Miller, K., Bovee, J., Roy-Charland, A., & Campbell, S. (2015). *Comprehension and quiz difficulty norms for 32 English passages.* Presented at Midwestern Psychological Association 2015, Chicago, IL.

Campbell, S., & Raney, G. (2014). *Reevaluating past metaphor norms*. Presented at Annual Meeting of the Psychonomic Society 2014, Long Beach, CA.

Campbell, S., & Raney, G. (2014). *A metaphor then is a metaphor now*. Presented at Society of Text and Discourse Processing 2014, Chicago, IL.

Raney, G., Bovee, J., Miller, K., Campbell, S., Fayz, L., & Brill-Schuetz, K. (2014). *Phonological similarity effects in native and non-native speakers.* Presented at Midwestern Psychological Association 2014, Chicago, IL.

Campbell, S., & Raney, G. (2013). *The role of familiarity and context strength in metaphor processing.* Presented at Annual Meeting of the Psychonomic Society 2013, Toronto, Canada.

Campbell, S., & Raney, G. (2013). In Raney and Ashby symposium on Using eye movements to study online processing of figurative language. Symposium presented at APS Annual Convention, Washington D.C.

Campbell, S., & Raney, G. (2012). *Individual differences in metaphor processing: An eye-tracking approach.* Presented at Midwestern Psychological Association 2012, Chicago, IL.

Campbell, S., & Raney, G. (2011). *Life is a pencil: Using eye tracking to explore metaphor processing.* Presented at Annual Meeting of the Psychonomic Society 2011, Seattle, WA.

INVITED PRESENTATIONS

Campbell, S. (2014). Categorization and its role in metaphor processing. Presented at Augustana College Psychology Department.

Campbell, S. (2014). An overview of past and present models of metaphor processing. Presented at Augustana College Psychology Department.

Campbell, S. (2013). How to make the most of your undergraduate experience: The evolution of a line of research. Presented at Augustana College Psychology Department.

Campbell, S. (2013). Familiarity and context strength impact metaphor processing: An eye-tracking study. Presented at Northwestern University Psychology Department, Gentner laboratory.

DEPARTMENT PRESENTATIONS

Campbell, S. (2014). A metaphor then is a metaphor now. Presented at University of Illinois at Chicago Cognitive Psychology Brown Bag.

Campbell, S. (2012). A masters thesis is a pencil. Presented at University of Illinois at Chicago Cognitive Psychology Brown Bag.

Campbell, S. (2010). The eye is the boss: Examining the career of metaphor. Presented at University of Illinois at Chicago Cognitive Psychology Brown Bag.

REVIEW EXPERIENCE

Ad hoc reviewer for Behavior Research Methods.

DEPARTMENTAL SERVICE:

Psychology Department IRB assistant, University of Illinois at Chicago, 2010-2015.

WORK EXPERIENCE

Teaching Fellow. 2017-Present. Instructor of psychology and liberal arts focused courses at Augustana College. Duties include lecture, assignment, paper, and test preparation and grading.

Graduate Research Assistant. 2010-2017. Conducting graduate level research on figurative language processing and the role of vocabulary on text comprehension. Duties included designing research paradigms and methodologies, running experiments, data collection and analysis, mentoring undergraduate research assistants, and presentation of research.

Graduate Teaching Assistant. 2010-2017. Teaching assistant for a number of undergraduate psychology courses. Courses included: Research Methods, Cognitive Psychology Lab. Duties included: teaching discussion classes, grading assignments and examinations, holding office hours, assisting in the classroom lectures.

ACADEMIC AND PROFESSIONAL AFFILIATIONS

Society for the Teaching of Psychology
Psychonomic Society
Association for Psychological Science
Psi Chi: The National Honor Society in Psychology