CUTTING THE MUSTARD:

AN EXPERIMENTAL INVESTIGATION OF IDIOMATIC EXPRESSIONS IN THE

LEXICON

by

Edward M. Holsinger

A Dissertation Presented to the
FACULTY OF THE USC GRADUATE SCHOOL
UNIVERSITY OF SOUTHERN CALIFORNIA
In Partial Fulfillment of the
Requirements for the Degree
DOCTOR OF PHILOSOPHY
(LINGUISTICS)

December 2011

Acknowledgements

I am extremely indebted to the members of my committee for their valuable comments, discussion, guidance and support throughout the process of developing this dissertation. In particular, I would like to thank Dr. Elsi Kaiser, for her unwavering support and mentorship during my time at USC and for the innumerable hours that she has spent reading, commenting and discussing this work as well as providing me with the knowledge and skill necessary to undertake it. Without her mentorship this document would not exist. I also wish to thank my wife, Maria, for her patience and support throughout this process, and for proofreading draft after draft for inevitable errors.

Table of Contents

List of Tables

List of Figures

Abstract

Much of the work in theoretical and experimental linguistics has focused upon compositional language, viewing non-compositional expressions such as idioms (e.g. *kick the bucket*) as exceptional cases in the language system. Early research into these expressions treated them as unanalyzed *words-with-spaces* (Bobrow & Bell, 1973; Swinney & Cutler, 1979), while more recent work places more emphasis upon their structural properties (Cacciari & Tabossi, 1988; Cutting & Bock, 1997; Sprenger, Levelt & Kempen, 2006). However there is still disagreement regarding the representation of idiomatic expressions in the lexicon, and the role of literal processing in the processing of idioms. This dissertation examines three primary questions (i) how are idioms represented structurally, (ii) how are idioms stored in the lexicon, and related to other elements in lexical space, and (iii) how do individuals process expressions ambiguous between an idiomatic and literal sense. We report the results of three distinct experimental investigations designed to address these questions. The results of our experiments suggest that idiomatic expressions are represented as structural frames in the lexicon and are sensitive to syntactic context during on-line processing (Exp 3a, 3b) and that idiomatic expressions, such as *kick the bucket,* are lexically related to their literal components *kick* and *bucket* such that access to the idiomatic representation is mediated by activation of its literal pieces (Exp, 3a, 3b, 4a, 4b). Additionally, we present evidence that some degree of literal processing of idiomatic expressions is obligatory (Exp 1, 2, 4a, 4b) even in semantic contexts that strongly disfavor the literal interpretation of these

strings. We present our results as they apply to models of idiom representation and processing.

Chapter 1: General Introduction

Most theoretical work in linguistics revolves around a class of constructions that can be classified as compositional. The notion of compositionally as applied to expressions of language is generally attributed to Frege (1884). Most generally, compositionality can be expressed as in (1).

(1)   The meaning of a complex expression is equivalent to:

     (a)   the meaning of the constituents of that expression and

     (b)   the way in which those constituents are composed.

Thus the meaning of a complex expression, such as *John eats beef* is equivalent to the individual constituents of that expression (minimally including *John, eat* and *beef*), and the way in which those constituents are composed, in this case some grammatical argument relation with the verb *eat*.

The notion of compositionality is perhaps one of the most powerful and well-defended ideas in theoretical linguistics, and with good reason. Compositionality is one of the few ways in which linguists can explain how individuals are capable of producing and understanding a theoretically infinite space of possible sentences with only a finite set of symbols and rules of composition. Additionally, compositionality helps us explain the overall systematicity apparent in language. This is the observation that if a speaker

understands what *John eats beef* means then they also necessarily understand what *Beef eats John* means (see Fodor & Pylyshyn, 1988, Fodor 1987, 2001).

The scope and importance of compositionality, therefore, makes it all the more interesting to examine those cases in which it fails. There are myriad cases of the failure of compositionality in the literature, including certain constructions (see Higginbotham, 1986; von Fintel 1998), propositional attitudes (see Pelletier, 1994), and quotation (Botterell & Stainton, 2005, see Pagin & Westerstahl, 2010 for a review of issues in compositionality). In many of these cases the purported failure is one of relatively subtle semantic computation, and while interesting need not concern us here. Instead I will be examining so called idiomatic expressions, which are pervasive across human languages, and under the standard view plainly ignore the conditions set forth in (1).

It is because of this general failure to behave compositionally, that we choose to examine idiomatic expressions. Specifically, I will be examining the structure and operation of the mental lexicon, which can be viewed as the cognitive system that contains the atomic units necessary for language to function. These units form the base constituents over which compositionality operates. Thus idioms provide a unique test case. On the one hand the apparently arbitrary nature of their form-meaning pairing argues strongly that these expressions must be learned and should be stored lexically like any other morpheme. On the other hand, the fact that many of these expressions are ambiguous

between a literal and idiomatic interpretation raises interesting questions regarding the means by which the language system identifies and processes these expressions.

Additionally, unlike other elements in the lexicon, there is extensive intuitive and experimental evidence that idiomatic expressions are not completely unrelated to their compositional companions. This poses additional problems both for the conception of the mental lexicon, and for the notion of compositionality as applied to sentence processing.

1.2 Structure of this Document

This document represents work done as a partial investigation into the structure and dynamics of the mental lexicon. I will specifically be looking at idiomatic expressions such as *kick the bucket*, *find her feet*, or *chew out*. These expressions are valuable to us, as they behave like structures in some ways, and like words in others. This duality makes idioms a valuable tool for investigating the lexicon and the relationship between lexical representations and grammatical computation.

In Chapter 1 we will first review the literature investigating the mental lexicon, starting from the sorts of formal approaches and concerns that lead to the need for such a structure. We will then turn to experimental investigations of the lexicon, and establish a working view of the lexicon that we can then use going forward. We will then turn our attention to idioms and discuss why they are interesting to us. We will discuss the blurry line between idiom and metaphor on one hand and idiom and fixed expressions on the

other, and then turn to specific issues with integrating idioms into models of the lexicon. Finally we will review current experimental work in idiom identification, representation and the relationship between idiomatic and literal representations and motivate the experiments presented in the following chapters.

In Chapter 2 we will present the results of a self-paced reading experiment designed to investigate what sorts of cues individuals use to determine whether a string of words is to be taken literally or idiomatically, and to examine the time-course of access to idiomatic and literal expressions during normal processing on the one hand, and during recovery from misplaced expectations on the other.

Chapters 3 and 4 will present eye-tracking studies investigating the time course of activation of idioms such as *kick the bucket*, and comparing that to semantically related expressions such as *kick the pail*. In Chapter 3 we investigate whether syntactic considerations are used on-line to guide individuals' interpretations, and when participants consider idiomatic and/or literal meaning in unbiased out-of-the-blue contexts such as *John kicked the bucket last Thursday*. In Chapter 4 we extend this approach to the investigation of biased contexts, in which preceding sentential information biases the interpretation of the ambiguous string toward either the idiomatic or literal interpretation.

In Chapter 5 we attempt to pull the results of these experiments together and, with guidance from the literature, attempt to provide a clearer picture of how idioms are represented and integrated into the mental lexicon.

1.3 The Mental Lexicon

Much of the business of linguistics over the past few decades has been involved in investigating general mechanisms and principles underlying human language. Today, there is still a great deal of disagreement regarding how much of the human language faculty is innate, how much is domain specific to language, and how much of what we have discovered about systems such as syntax is generalizable cross-linguistically. Ignoring these generalizations, however, it is important to remember that the bulk of the actual data over which these potentially universal systems operate, is both arbitrary and learned. If we were to metaphorically use the language of chemistry to discuss linguistics, the prime candidate for the atom would be the morpheme: the smallest meaning-bearing unit. These meaning-bearing units form the bulk of our linguistic data, and with only minor exceptions (such as onomatopoeic words) act as arbitrary mappings between linguistic forms and concepts. This arbitrariness between linguistic form and meaning is easily demonstrated by considering word meaning across languages as in (2) below.

(2) Words expressing the concept DOG

    (2a) dog (English)

    (2c) hund (German)

(2d) gou3 (Mandarin)

(2e) koira (Finnish)

(2f) kalib (Arabic)

Indeed this lack of relationship between words for the same concept in different languages has long been exploited in historical linguistics as a means of tracing relationships between language families. Hence the fact that the word for cat is similar in English and German suggests a relationship between these languages.

These symbols are thus learned during the process of acquiring one's native language, and are stored or associated with information regarding their pronunciation, meaning and grammatical properties. Syntactic, semantic and phonological processes operate over these units, and much of the day-to-day business of language production and comprehension is involved in their access and use to form larger structures. As such these units — words and morphemes — occupy a central role in language, and the cognitive structure which is responsible for their storage and retrieval has come to be known as the mental lexicon.

The mental lexicon is the cognitive structure that serves as a mental dictionary or database for linguistic information. The contents of this lexicon are widely taken to be the individual words (e.g. *dog, cat*) and perhaps other morphemes (e.g. *-ed, -ing*) along with information regarding their meaning, pronunciation, and grammatical roles. The

information contained in the lexicon is, in a very real sense, the information that makes the symbolic system of human language possible. The lexicon contains information both language-specific (e.g. pronunciation of *DOG* in English or Mandarin), as well as language general (e.g. the concept of *DOG*). It contains information theorized to be innate (e.g. grammatical categories) as well as information that must be learned (e.g. pronunciation, meaning of individual symbols). Having a better understanding of what information is contained in the mental lexicon, and how that information is represented, accessed and utilized will go a long way in informing both theoretical models and computational implementations of language.

Additionally, the lexicon sits at an important nexus in the human language system. This system must interface with a great variety of other cognitive systems ranging from grammatical and logical systems to motor control structures to perceptual systems. Thus, the study of the lexicon is not only the study of the cognitive data necessary for human language, but also serves as a window into a wide range of cognitive systems, helping us answer important questions regarding their structure, operation and domain-specificity with regard to language. Without worrying too much right now about the specifics, we can view the lexicon as the place where *exceptional* information is stored. That is the lexicon would contain the arbitrary information necessary for language (*e.g. dog is a Noun, wonder must take a sentential argument, cat starts with an unvoiced velar stop*).

1.3.1 Approaches

The lexicon was originally proposed in Chomksy (1965) as a framework for containing idiosyncratic information pertaining to individual words and morphemes. In this conception the lexicon consists as a set of entries, each of which is itself a set consisting of the phonological and syntactic features of a given lexical item. Specifying this information in the lexicon, thus freed the grammatical system from having to deal with these properties directly. Instead, syntax could focus upon general grammatical regularities and delegate much of the idiosyncrasies of individual words to direct specification in the lexicon.

By the 70's the conception of the lexicon had changed dramatically both in formal and experimental circles. In the formal sphere Chomsky's *Remarks on Nominalization* (1970), extended the breadth and power of the lexicon to potentially include morphological features and operations, heralding a wealth of investigation and exploration of the lexicon as a formal device across all subfields of linguistics (see Jackendoff, 1990, 1992; Selkirk, 1982; Halle & Marantz, 1993; Kiparsky, 1982, Pollard & Sag, 1994; Pustejovsky, 1995). Around the same time, a deep investigation of the psychological and computational architecture of the mental lexicon began.

Early models of the lexicon were generally either concerned with conceptual organization (Collin & Quillian, 1969, Smith, Shoben, Rips, 1974) or access and retrieval (Forster, 1976; Morten 1970), but these two avenues of inquiry fed into and enriched each other.

This confluence of ideas from theoretical, computational and experimental approaches rapidly lead to models with very complex structure and behavior and gave way to the exploration of the interfaces between the lexicon and other linguistic systems in a push toward a coherent model of sentence production from lexicon to utterance (Garret, 1980; Bock & Levelt, 1994; Levelt & Meyer, 1999; Kempen & Huijbers, 1983, Collins & Loftus, 1975; Seidenberg & McClelland, 1989; McClelland & Rumelhart, 1981; Dell, 1986)

1.3.2 A Working Model

Much of the experimental evidence for the internal organization of the lexicon has come from investigations of the incidence and properties of speech errors (Meringer & Mayer, 1895; Fromkin, 1973; Dell, 1986) or research into priming and interference (Stroop, 1935; Lupker, 1982; Schriefers, Meyer & Levelt, 1990; Roelofs, 1997; Levelt, 2001, Levelt, Roelofs & Meyer, 1999, see also Levelt 1999 for a review). Models developed under the former paradigm focus upon the observation that speech errors are not randomly distributed, but that mixed form-meaning errors tend to be overrepresented in corpora (see Fromkin, 1973, Dell 1986). The classic example in being the speech error *evoid*, which is a blend of *avoid* and *evade* which have semantic, grammatical and phonological similarities. Models developed under the latter paradigm focus upon the observation that lexical access can be enhanced or impeded by accessing lexical content related to some target word (e.g. hearing *doctor* makes *nurse* easier to access in some cases).

Explaining the incidence and underlying causes of speech errors resulted in Dell's (1986) spreading-activation theory of sentence production, which has persisted as a leading model of lexical architecture to the present day. Dell's model is essentially an interactive symbolic network model in which individual nodes associated to each other via bi-directional links (see also McClelland & Rumelhart, 1981; Rumelhart & Norman, 1982). Activity in the system is governed by the principle of *spreading activation* and the principles of *spreading*, *summation* and *decay*. Spreading is the process by which an active node will automatically spread some amount of its activation to adjacent (i.e. linked) nodes in the network, with the proportion of activation spread to each neighboring node being dependent upon the strength of the connection. Summation is essentially a formal statement that if a single node receives activation from multiple sources, then the resulting activation is additive, and decay provides a mechanism by which activation of all nodes in the network gradually decays over time.

As our primary concern in this dissertation will be to examine the lexical representation of idiomatic expressions we will be assuming a lexical model in the spirit of Dell (1986), in which the overall architecture is that of a symbolic network, with individual nodes associated with other nodes via connections which can vary in strength and where all connections assumed to be bi-directional. We acknowledge, of course, that other proposals do exist (see for example Roelofs, 1997; Elman 2009), however for the purposes of this investigation we will be concerning ourselves primarily with the facts

regarding the representation of idiomatic expression rather than the particulars of one or another model of lexical access and representation.

We also generally assume that the system is fully symbolic and hence we take the less controversial view of conceptual representations as being atomic units rather than feature clusters (see Bierwisch & Schreuder, 1992; Roelofs, 1997). In this dissertation we will not be concerned how conceptual representations are structured, but instead focus upon how these representations are related to linguistic symbols. Thus the details of this distinction needn't cause us great concern, and thus we adopt the less complex and arguably less controversial view of conceptual representation.

It is also worth noting that in this document we will be taking a relatively mainstream view of the process of composition. That is, compositional phrases are constructed by grammatical operations applied over individual morphemes, which form the base of the linguistic process. Recent work into usage-based, exemplar and construction grammars have challenged the view that that there is a hard line between grammatical processes and lexical content (see for example, Goldberg, 1995, 2006; Tomasello, 2003; Bod, 1998). Again, however, this document will primarily be concerned with the representation of idioms, and their relationship to compositional grammatical structures rather than argue for or against a particular view of the lexicon – grammar interface.

Thus for the purposes of this dissertation, we will generally assume a lexicon in which words can be formally represented as nodes in a network. These nodes are associated

with other nodes via bi-directional links, which connect the word-level nodes to relevant grammatical, conceptual and phonological information. In what follows we will discuss idiomatic representation in these terms, primarily as a matter of convenience: contemporary models of idiomatic representation are expressed, or are expressible easily in this functional paradigm. We intentionally leave the exact implementation of the lexicon somewhat vague, but note that our results depict details regarding the processing of idioms and their relationship to other elements of the lexical and grammatical system, and hence should be accounted for regardless of the specifics of one's lexical architecture.

1.4 Idioms

While much of the previously discussed research has focused primarily on the mechanisms used to compose and compute linguistic structures, normal language use also consists of a large number of conventionalized and figurative expressions. We can operationally define a conventionalized expression as one in which the form of the expression is highly restricted. This minimally includes idiomatic expressions (e.g. *kick the bucket, cross that bridge when we come to it*), common collocations (e.g. *cats and dogs, war and peace, I'm fine*) and cliché expressions (e.g. *It was a dark and stormy night*). Figurative expressions, on the other hand, may or may not exhibit productive usage, but are operationally distinct in that the meaning of these expressions is not easily computable compositionally. Expressions of this sort include idioms as well as larger metaphoric usages including things like irony and simile.

Use of figurative and conventionalized expressions, including idioms is characteristic of normal language use (Lakoff & Johnson, 1980; Pollio, Smith & Pollio, 1990), and perhaps essential for fluent native-like proficiency (Pawley & Syder, 1983). Additionally, conventionalized expressions account for a significant proportion of the lexical items in a given language (Jackendoff, 1995; Weinreich, 1969). Despite the prevalence of these sorts of expressions, many of the grammatical and lexical systems that have been developed treat these expressions as exceptions to the more general processes of lexical retrieval and grammatical computation.

In this dissertation we will be explicitly examining idiomatic expressions, which are simultaneously a subset of the class of figurative expressions (which also includes forms such as metaphor and irony) and of conventionalized expressions (which also includes forms such as cliché and affective greetings). Before describing the properties and experimental investigations of idioms, it is worthwhile to establish a general understanding of how idioms differ from figurative language on the one hand, and from conventionalized expressions on the other, and examine in more detail why we believe that idiomatic expressions provide such a valuable window in the structure and operation of the mental lexicon. It is worth noting that while we will be discussing idioms as a class of expressions for convenience, there is no general consensus in the literature that such a uniform class exists, and attempts to define or operationalize idiomatic expressions as a unified class have been generally problematic (Wasow, Sag & Nunberg, 1983; Fraser,

1970; Gibbs, 1985; Gibbs, 1984) despite findings that individuals do tend to have intuitions regarding these expressions (Gibbs & Gonzales, 1985, though see also Keysar & Bly, 1999; Tabossi, Fanari & Wolf, 2008). Despite this, work in theoretical, experimental and computational domains converges upon an operationalization of idiomatic expressions that focuses upon their conventionalized form and lack of compositional interpretation. In what follows we will review the distinction between idioms and other figurative language, and between idioms and conventionalized expressions and attempt to provide an operational definition of these expressions that we can use to move forward.

1.4.1 Idioms vs. Metaphor

The distinction between idiomatic expressions and other figurative language on the one hand, and conventionalized expressions on the other, is vague at best. When distinguishing idioms from other figurative expressions the key distinctions are that idioms are conventionalized and less productive. Thus an idiom, such as *kick the bucket* in English, is a conventionalized ways of expressing *die*. It is syntactically unproductive, evidenced by the fact that it is heavily restricted in the number of syntactic operations it can undergo while maintaining its meaning. It is also semantically unproductive, thus *slap the bucket, punt the bucket* or *kick the trough* are literal phrases rather than alternative expressions of the idiom. Other figurative expressions are typically more productive and hence less conventionalized. For example, figurative expressions of the

form *X is a Y* can be applied for essentially any X and Y productively given some appropriate context (the 405 freeway is a parking lot, LA drivers are a pack of wolves).

However some care needs to be taken here, as many figurative expressions are conventional to varying degrees (e.g. *men are pigs, you are my sunshine,* etc.) which blurs the line somewhat. The relevant distinction however is in conventionalization of their interpretation. Idioms mean whatever it is that they mean, whereas even highly conventional figures of speech such as *men are pigs* are more flexible in the meaning that they convey. Thus an expression such as *men are pigs* leaves the exact details of the comparison between men and pigs open to context and interpretation, only specifying that men and pigs overlap in some contextually relevant set of conceptual properties. Likewise for an expression such as *you are my sunshine*, or *LA drivers are a pack of wolves* the exact subset of properties which are highlighted as overlapping in these comparisons are flexible. For idioms, however this is not the case. *Kick the bucket* means *die* or perhaps *die suddenly*.

Distinguishing idioms from other conventionalized expressions seems, on the surface to be a rather straightforward task. Conventionalized expressions such as clichés are compositional, idioms are not. On greater thought, however, this is not entirely accurate. Clichés such as *it was a dark and stormy night*, or *damned if you do, damned if you don't* are in some sense compositionally meaningful, but are also conventionally used and understood in highly circumscribed ways.

There is some experimental evidence to distinguish between idiom and metaphor in terms of processing cost. Ortony, Schallert, Reynolds & Antos (1978) conducted an experiment in which participants were given sentences ambiguous between a literal and metaphorical sense (e.g. *Regardless of the danger, the troops marched on,*) or between a literal and idiomatic sense (e.g. *let the cat out of the bag*). The sentences were preceded by context which biased one of the two interpretations. Thus to use our previous examples participants might see a story about children misbehaving for their somewhat inhumane babysitter, in which case *the troops marched on* would need to be interpreted metaphorically. For literally biased trials participants would see a story about soldiers in battle. Similarly for the idiomatic conditions, preceding story contexts would bias an interpretation in which a literal cat was let out of a literal bag, or an interpretation in which the expression is meant idiomatically. They found that it took participants longer to understand metaphorical senses than literal senses (as measured by reaction times) when the preceding context was short (single sentence), but not when the context was longer. For idioms they observed the opposite effect: idiomatically biased interpretations were understood faster than literally biased interpretations. In general, these results suggest that the processes underlying the interpretation of metaphor differ in important ways from those at underlying the interpretation of idioms (see also Caillies & Declercq, 2010)

The finding that idioms are processed rapidly has been replicated many times in the literature and can be viewed as one of the core processing characteristics of idioms (Gibbs & Gonzales, 1985; Gibbs, 1980; Gibbs & Nayak, 1989; Cacciari, Padovani & Corradini, 2007; Swinney & Cutler, 1979). Indeed this finding is one of the motivations for claiming idiomatic expressions are represented as 'big words', stemming from the idea that lexical access is more rapid than lexical access plus grammatical computation. Recent research into multi-word expressions, however suggests that this processing advantage may largely be due to conventionality (Tabossi, Fanari & Wolf, 2009) or simple phrase frequency (Arnon & Snider, 2010), hence this cannot be used reliably to differentiate between idioms and other conventional and/or frequent phrasal expressions.

Operationally we will differentiate from idioms and metaphorical expressions by appealing to conventionality and intuition. The idioms we use in this paper are either phrasal verbs (e.g. *look up, chew out*) or verb phrase idioms (*kick the bucket, find her feet*). For our purposes these expressions are sufficiently compositionally opaque and unproductive to distinguish them from metaphor. To distinguish these expressions from other conventional expressions we again appeal to compositional opacity. Admittedly this operational definition works much better with our verb-phrase idioms than with some of the phrasal verbs (*X up* is relatively productive and potentially a metaphorical extension in English), thus in the cases of phrasal verbs we have been extra careful in the selection and norming of our stimuli.

1.4.2 How Idioms Enter Vocabulary

The standard assumption regarding the historical origins of idiomatic expressions is that they represent *dead metaphors* which have, for some reason or other, lost their metaphorical connections and become conventionalized (Weinreich, 1969; Chomsky 1965; Aitchison, 1987). This assumption derives from series of observations that face several challenges and complications. First is the observation that idioms are *grammatical* in the sense that their surface form seems to exhibit the sorts of features that one would expect from a grammatical structure. Hence the majority of idioms appear to be ambiguous (if perhaps pragmatically odd) between an idiomatic and literal interpretation. The logic here is that *if* idioms *were* once examples of normal metaphorical language, then it had better be the case that they were grammatical. Of course there are a large number of examples that do not adhere to this observation. In English we have idiomatic phrases such as *by and large*, *no can do*, *trip the light fantastic*, and *believe you me* none of which are grammatical in the typical sense (see Wasow et al, 1983 for further examples). While this doesn't falsify the *dead metaphor* account, it at least suggests that idiomatic expressions may not always enter the lexicon via this route.

The second observation is that idiomatic expressions are essentially conventionalized metaphors which have lost their compositional meaning. This too is not entirely correct on a number of grounds. Gibbs (1993) in the aptly titled *Why idioms are not dead metaphors* points out that intuition and occasionally even etymological views as to the

metaphorical origins of idiomatic expressions often turn out to be suspect or incorrect. Further, it is largely unclear whether many idioms have fully lost any attachment to metaphorical and compositional interpretation. Wasow & Sag (1983) point out that many idioms can be modified with expected results (e.g. *leave no legal stone unturned*) (see also Cacciari & Glucksberg, 1991). Additionally there is evidence that the overall meaning of an idiomatic expression is influenced by its individual components (Nunberg, 1978; Fillmore, Kay & O'Connor, 1988; Gibbs & Nayak, 1989, Glucksberg, 1993) (e.g. ?*The wounded soldier kicked the bucket for several days*), and that the meaning of individual idiomatic expressions are not entirely unrelated but seem to share a great deal of underlying metaphorical conceptual content (Lakoff & Johnson, 1980; Gibbs, 1993).

A further complication for the dead metaphor view comes from evidence that individuals have intuitions regarding the degree to which idioms are compositionally analyzable (Gibbs & Nayak, 1989; Gibbs, Nayak, Cutting, 1989; see also Tabossi et al., 2008). In terms of language acquisition, there is evidence that implicates decomposability and metaphorical content as a deciding factor in learning new idioms both for children (Caillies & Le Sourn-Bissaoui, 2008; Hsieh & Hsu, 2010) and second language learners (Boers, 2000). Thus the view that idiomatic expressions are non-compositional, dead metaphors cannot be entirely accurate, as there is sufficient evidence that in many cases idioms are neither fully non-compositional, nor metaphorically dead in the relevant sense.

1.4.3 Idioms as a Window into the Lexicon

Idioms exhibit both word-like and structural properties. Like words they can be accessed rapidly as compared to structural computation, and the mapping between their conceptual meaning and their linguistic form is arbitrary. Like structures, however, they exhibit grammatical sensitivities and induce structural effects such as structural priming. Additionally, with few exceptions, idiomatic expressions are also ambiguous between a literal-compositional and idiomatic interpretation. This collection of properties makes idiomatic expressions a valuable tool for exploring the architecture and dynamics of the mental lexicon and the interface between the lexicon and grammatical systems.

With regard to the lexicon, the arbitrariness of idiomatic meaning suggests a word-like direct mapping between form and meaning, however their structural properties and ambiguity present problems for current theories of lexical access. Research into ambiguity resolution is extensive, however a common thread is that the ambiguity to be resolved is between forms at the same level of linguistic abstraction. Thus in a typical lexical ambiguity scenario the choice is between homophonous words (3), for structural ambiguity between possible syntactic forms (4) or syntactic category (5).

(3) John saw the man at the (river/financial) bank.

(4) John saw the man with the telescope.

(5) I know the desert trains (soldiers to be tough / could resupply the camp)

The third case is perhaps most relevant for the processing of idioms, as the ambiguity draws upon both the lexical and syntactic domains. In examining these expressions, Frazier & Rayner (1987) proposed a delay strategy in which uncertainty regarding the syntactic category of an ambiguous term would cause the parser to delay syntactic integration of these items until such time as it settles upon an appropriate interpretation (though see MacDonald, 1993). The details needn't concern us, however, the crucial point is that these ambiguities result in a situation where the appropriate syntactic structure cannot be computed without making decisions regarding the grammatical property of an ambiguous lexical item.

For idioms, however, the ambiguity is between a syntactic structure with a computable compositional meaning and a non-compositional structure with an arbitrary word-like meaning. Critically this differs from the previous example in several important ways. First, while the Frazier & Rayner's proposal of delaying structural integration until the ambiguity is decided may make sense for their cases, the prevailing view in the idiomatic literature is one in which structure building proceeds with priority over idiomatic interpretation (Cacciari & Tabossi, 1988; Cutting & Bock, 1997; Sprenger et al, 2006). Additionally, the crux of this ambiguity is not whether a given lexical item in the phrase is to be interpreted literally or not, but instead whether the entire phrasal unit *as a whole* should be interpreted as a lexical unit or should be processed structurally as a well-formed syntactic phrase. Examining this process of this ambiguity resolution allows us to

address questions regarding the interplay between syntactic computation and lexical access, and directly examine the interface between syntax and the lexicon.

Additionally, the dualistic properties of idioms allow us to ask interesting questions regarding the architecture of the lexicon. As mentioned earlier, recent research on multi-word expressions suggests that in addition to individual morphemes, the lexicon may also store larger syntactic chunks perhaps to facilitate access to frequent syntactic constructions. While this notion is somewhat controversial, the notion that idiomatic expressions are stored lexically is not (Katz & Postal, 1963; Weinreich 1969; Swinney & Culter, 1979). Due to their arbitrary nature, idioms must be learned and stored lexically, however their structural properties allow us to investigate how structural chunks, such as those proposed for multi-word expressions, are represented, stored and accessed.

Idiomatic expressions also grant us a unique view into the relationship between structures and their component pieces that is not confounded by computational systems. If there is a relationship between the verb *kick* and the idiom *kick the bucket* then that relationship presumably exists solely within the domain of the lexical system, in much the same way that *beach* and *bagel* exhibit a relationship by virtue of sharing the same onset phoneme. Exploration of this sort of relationship with compositional structures is difficult, as any effects obtained will necessarily be confounded by possible computational activity. Thus even if it is the case that frequent multi-word expressions such as *all over the place* are stored lexically as structural units, any effects suggesting a relationship between *all* and

all over the place could either be a result of their lexical representation or the fact that one is a compositional component of the other.

1.5 Models

Many models have been proposed that deal with idiomatic expressions in theoretical, experimental and computational frameworks. In general these models can be grouped into four categories which we will discuss in more detail below. Generally, these models differ in how they represent idiomatic expressions and how they propose these representations are accessed. *Words-with-spaces* models generally interpret idioms as big words, existing either at the same level of representation as normal words or in a special list of exceptions. Decompositional models propose that idioms, or at least many idioms, are partially compositional and are understood and interpreted by making reference to conceptual metaphorical information. The other two types of models both assume some level of structural representation for idiomatic expressions. The Configuration Hypothesis grew out of work on idiom comprehension and proposes a priority of literal computation with idiomatic interpretations becoming available only later. Finally the Hybrid Representation Hypothesis comes from work involving idiom production, and suggests a phrasal representation in the lexicon, with access mediated via the literal components of the idiom.

1.5.1 Words-with-spaces

The earliest models of idiomatic representation treated idioms essentially as words-with-spaces (Katz & Postal, 1963; Weinreich 1969), an approach which still holds some sway in computational circles (cf. Sag, Baldwin, Bond & Copestake, 2002 for discussion). In a simple lexical view, words are linguistic representations with arbitrary direct mappings to particular syntactic, semantic and conceptual information. Early research into idioms argued that idioms are word-like, in the sense that they occupy the same level of representation, being directly associated with semantic and conceptual information without a need for compositional interpretation. For example, Bobrow & Bell (1973) argued that idioms are stored in a separate system accessed via a special, non-compositional processing mode. Evidence for this claim comes from experiments showing 'literalness priming' effects: Participants are more likely to interpret an ambiguous string if as idiomatic if they have recently been exposed to several idiomatic strings and as literal when preceded by literal strings (see Figure 1).
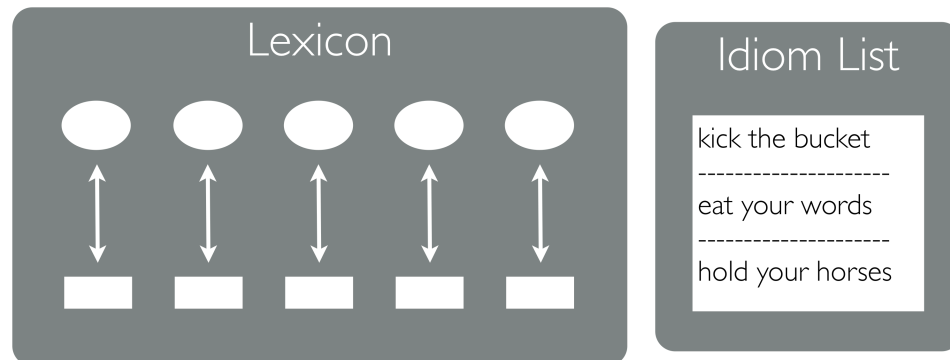
Figure 1: Organization of lexical space into a lexicon, containing individual words for grammatical computation and a separate list component for the storage of idioms.

Further evidence for the word-like nature of idioms was provided by Swinney & Cutler (1979)'s finding that idiomatic expressions are recognized as valid expressions faster than literal phrases (see also Gibbs & Gonzales, 1985; Gibbs 1980; Gibbs & Nayak, 1989; McGlone, Glucksberg & Cacciari, 1994). According to Swinney & Cutler (1979), idioms are stored in the lexicon like words. During processing accessing the idiom and computing the literal meaning of the expression proceeds in parallel, with the apparent speed advantage of idiomatic expressions emerging because idioms can be accessed directly in the mental lexicon without need for additional computational steps. They termed this model the *Lexical Representation Hypothesis* (see Figure 2). In addition to offering an intuitively appealing explanation for the rapid recognition of idioms, the Lexical Representation Hypothesis also allows us to delegate the resolution of the pervasive literal/non-literal ambiguity to the same sort of systems which handle other kinds of lexical ambiguity (see also Schweigert, 1986; Estill & Kemper, 1982).
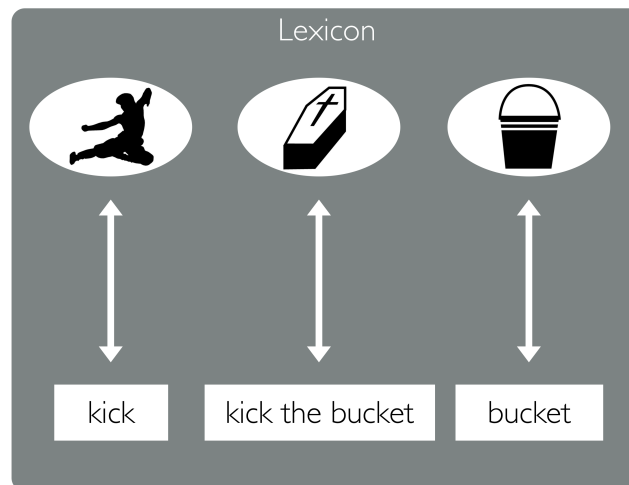
Figure 2: A portion of the lexicon according to the Lexical Representation Hypothesis. Note that idioms are essentially stored as large words, directly associated with their conceptual representations.

While the Lexical Representation Hypothesis predicts no relationship between the idiomatic and literal versions of an ambiguous string, later research shed doubt upon this prediction. There is considerable evidence that idiomatic expressions are, in fact related to their literal components (Wasow et al, 1983; Nunberg et al, 1994; Nayak & Gibbs, 1990; Katz, 1973; Nunberg 1978; Cacciari & Glucksberg, 1991 and others). In addition there is a body of experimental work which suggests that, contrary to the words-with-spaces view idiomatic expressions behave like structures. Gibbs & Gonzales (1985), for example provide evidence for varying degrees of syntactic rigidity among idioms (see also Fraser, 1970, Nunberg et al, 1994). It has also been demonstrated that individuals are aware of syntactic category information during the processing of idiomatic expressions (e.g. that *bucket* is a noun) and in recent work, Konopka & Bock (2009) found evidence for syntactic priming with phrasal verbs regardless of their level of idiomaticity (See also Peterson, Burgess, Dell & Eberhard, 2001). Taken as whole this literature strongly argues against a words-with-spaces view of idiomatic representation.

1.5.2 Decomposition & Direct Access

Several efforts have attempted to model idiomatic processing and representation in a way compatible with the evidence for a structural representation of idiomatic expressions. A general thread among these models is the proposal that idiomatic expressions are not as non-compositional as a words-with-spaces view would predict. Despite this common ground, however, theories differ greatly in what they mean by compositional, and in their

views regarding the relationship between the idiomatic and literal representations, particularly in which has priority.

One such theory is Idiom Decomposition Hypothesis (Gibbs & Nayak, 1989) and the closely related Direct Access Hypothesis (Gibbs, 1980). Gibbs (1980, 1985) proposed a view of idiom processing in which access to the idiomatic interpretation of a string can be performed directly, bypassing literal meaning computation altogether. In a series of three experiments, he demonstrates that individuals tend to process idiomatic usage of ambiguous expressions faster than literal usage when the idiomatic usage is more conventional than the literal usage. He also found that that this notion of conventionality plays a role in memory retrieval of these expressions. His conclusion from these results is that conventionality is the important factor in determining which interpretation is accessed, and claims that individuals tend to access the idiomatic interpretation *first* in many cases.

Unfortunately, Gibbs does not provide an operationalization of the notion of conventionality other than the general proposal that idiomatic expressions are conventionally used in their idiomatic sense. Presumably conventionality could be better expressed as a high frequency of idiomatic usage of a given expression as compared to its literal usage. However getting at this frequency is a complex task. For individual lexical items, frequency can be evaluated as the number of occurrences of a given lemma as compared to other appropriate lemmas (e.g. the frequency of the verb *kick* in various

27

morphological configurations as compared to the frequency verbs in general). For idioms this is a non-trivial task. Thus one might look at *kick the bucket* used idiomatically as compared to overall usage of *kick the bucket* used literally or idiomatically to get a rough estimate of the frequency of the idiomatic expression. However, the overall low frequency of expressions like *kick the bucket* in written corpora, confounded with the issue of the appropriate comparison class (i.e. since idioms are less syntactically productive than literal expressions, which of literal usages should count) makes this difficult to examine[1].

In later work this direct access hypothesis was tied to the proposal, following Lakoff & Johnson (1980), Nunberg (1978) and others that at least some idiomatic expressions *are* compositional in some relevant sense. This we will call the Idiom Decomposition Hypothesis. Gibbs & Gonzales (1985) provide evidence that individuals have reliable intuitions regarding the syntactic flexibility of idiomatic expressions, and in later work Gibbs & Nayak (1989) demonstrated that this degree of flexibility is related to individual's intuitions regarding the degree to which the expression can be semantically decomposed. His central argument is that while idiomatic expressions do not form a

---

[1] We do find some support for the frequency notion of conventionality for the stimuli used in our Experiment 2, reported in chapter two. An examination of potentially idiomatic expressions (e.g. *kick the bucket*) vs. the closely related non-idiomatic expressions we used as stimuli (e.g. *kick the pail*) in the Google Books Corpus (Davies & Mark, 2011, see also Michel et al. 2011) revealed a consistent pattern in which the frequency of the potentially idiomatic string was much higher than that of the related expression (mean ratio 32:1 occurrences). This examination did not, however, compare idiomatic usage of *kick the bucket* to non-idiomatic usage. Further examination of the frequency of these expressions, and particularly the frequency of their idiomatic usage would be valuable.

unified class, as evidenced by varying degrees of syntactic flexibility and semantic transparency, at least some of these expressions can in principle be computed with the same sorts of processes that are understood to underlie the processing of literal language.

Given this view, Gibbs and colleagues argue that the access and processing of idiomatic expressions, while perhaps similar to literal processing in some cases, does not require that any actual literal processing happens. Instead idiomatic interpretation can be arrived at by virtue of access to a system of metaphorical conceptions. Hence when interpreting an idiom such as *spill the beans*, the parser relies upon a system of conceptual metaphors such as *THE MIND IS A CONTAINER* and *THOUGHTS ARE OBJECTS* to navigate the figurative content of the expression (Gibbs, Bogdanovich, Sykes & Barr, 1997). This view of idioms proposes that these expressions are represented lexically as mappings between phrasal strings and this system of conceptual metaphor, with the potential for individual components of the string to be mapped separately to different aspects of this metaphorical system (Nayak & Gibbs, 1990). A consequence of this view, then, is that while something akin to literal processing may occur in the comprehension of these phrases, it is not the case that such processing will involve or result in access to the literal interpretation of the expression. For our purposes we will conflate the idiom decomposition hypothesis and the direct access hypothesis into one view which essentially states that idioms are partially compositional, and like other compositional expressions, can be interpreted directly.

A related approach is given by Glucksberg (1993), in which idiomatic expressions are related to their literal components via a mechanism they describe as Phrase Induced Polysemy. Under this approach, the individual lexical items which comprise an idiom take on new meanings by virtue of their usage in the expression. Thus for an idiom such as *pop the question* interpretation can proceed by assuming that *pop* has acquired the meaning *suddenly ask* and *the question* the meaning *marriage proposal* by virtue of their collocation in the idiom. This view differs from the decompositional hypothesis in that it does not rely upon a system of underlying conceptual metaphors to derive the interpretation of an idiom, but rather grounds the apparent decompositional properties of idioms in the collocational and usage properties of the expression.

1.5.3 Configuration Hypothesis

Another proposal is put forward by Cacciari & Tabossi (1988), who propose that during comprehension the parser proceeds with the normal process of literal interpretation until such time as it recognizes a given string as idiomatic. Once this occurs, the idiomatic meaning is retrieved and enters into the comprehension process. To support this view they used a cross-modal lexical decision task to probe whether idiomatic and literal interpretations were activated during the processing of Italian idioms such as *in seventh heaven*, and *go to the devil*. They found that when the idiomatic nature of the phrase was predictable, as measured by an idiom completion pretest, participants showed evidence of activation of the idiomatic interpretation but not the literal interpretation when probed on the offset of the expression. When the idiom was not predictable, however, participants

showed activation of the literal meaning at the offset of the expression, but did not exhibit signs of idiomatic activation until 300ms later.

These results argue against the idea, inherent in the Lexical Representation Hypothesis, that the literal and idiomatic interpretations of a given ambiguous string are unrelated and processed in parallel. Additionally, the results seem to suggest that the parser proceeds with processing structures literally until such time as it has recognized a string as idiomatic. Thus their results suggest some priority of literal processing, in opposition with the direct access model. Instead they propose the Configuration Hypothesis, in which idioms have a distributed lexical representation. They argue that literal processing proceeds automatically, and continues until the parser accumulates enough evidence that the string it is considering is idiomatic, at which point the parser engages in the retrieval of the idiomatic meaning and presumably suppresses the literal interpretation. They refer to this tipping point in idiom processing as the *idiom key* (see Tabossi & Zardon, 1993 for attempts to operationalize this notion).

This model is particularly relevant, as it provides an explanation of the decomposability and structural properties of idioms, while still providing an explanation about why they are processed more quickly than matched literal expressions. At least for highly predictable idioms, it suggests that the apparent speed advantage is a result of the parser only partially processing the string literally. How much literal processing is required depends heavily upon the properties and usage of the idiom in question, and the context

in which it occurs. Crucially how far down the literal parse the processor goes is dependent upon the degree to which the idiomatic string can be plausibly taken literally, and how well the parser can predict the intended meaning during online comprehension (Titione & Connie, 1994, 1999). In general, the configuration hypothesis proposes that in the absence of contextual biases the processing of an idiomatic string proceeds literally first, and its figurative interpretation becomes active only sometime later.

1.5.4 Hybrid Models

Recent work into idiom production also suggests a distributed representation and a primacy of literal processing. On the basis of speech error data, Cutting & Bock (1997) suggest that the production of idioms is sensitive not only to the idiomatic meaning of the phrase in question, but also to its syntactic structure and literal meaning. They presented participants with pairs of idioms which either overlapped in structure (e.g. *shoot the breeze, raise the roof*), both meaning and structure (e.g. *shoot the breeze, chew the fat*) and controls in which neither the meaning nor structure matched. Participants were then tasked with recalling and producing one of the idioms in the pair from memory as quickly as possible. They found that idiomatic expressions with matched structural form resulted in higher error rates, and that errors overwhelmingly observed grammatical category constraints. In a further experiment they found that overlap in meaning between paired idiom and literal phrases produces as many errors as with idiom pairs. The finding that, during idiom *production*, the literal meaning and associated syntactic structure are nevertheless activated is quite striking, given that the speaker presumably knows that she

is producing an idiomatic expression — hence one might expect that sensitivity to the structural properties and literal interpretation of an idiom would be unnecessary.
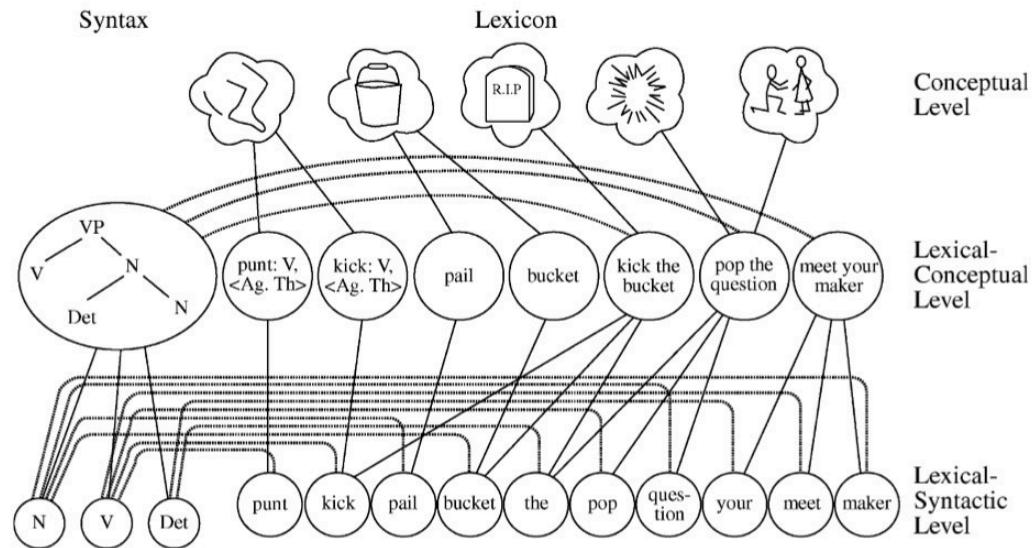
Figure 3: A portion of the lexicon relevant to the idioms *kick the bucket*, *pop the question* and *meet your maker*. Reproduced from Cutting & Bock (1997)

Thus Cutting & Bocks' data is strong evidence in favor of the *hybrid representation hypothesis* (see Figure 3. In their model, idiomatic expressions are represented as phrasal frames in a lexical-conceptual layer of the lexicon. Like words, idioms are associated directly with conceptual content. Like structures, access is mediated via the literal components of the expression and the lexical-conceptual representation is associated directly with a structural representation. This model predicts that structural and literal information will be recruited both during idiom production and comprehension. The model also predicts tight integration between the idiomatic representation of an ambiguous string and the literal meaning of its component parts. Activation of an

expression such as *kick the bucket*, for example, should result in activation of literal *bucket* which should in turn result in activation of semantically and phonologically related lemmas (e.g. *pail* and *buck*).
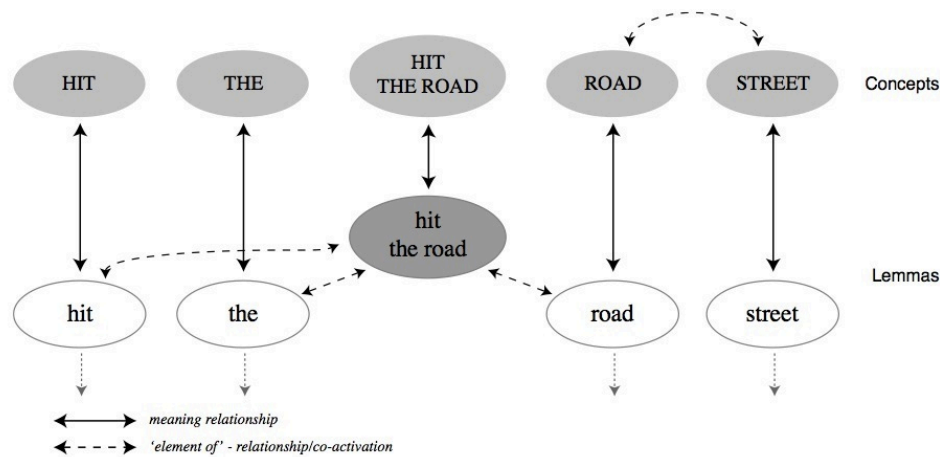
Figure 4: Model of the lexical representation of hit the road and relevant associated nodes in the Hybrid Representation Hypothesis. Figure reproduced from Sprenger et al (2006).

Sprenger, Levelt & Kempen (2006) examined these predictions explicitly and provided a refined model of idiom production (see Figure 4). Their experiments showed that identity priming of a word in an idiom (e.g. showing people *bucket* for *kick the bucket*) facilitated cued recall of the idiom. Additionally, this priming effect was found to be *greater* for idioms than for related literal strings as predicted by the architecture of the hybrid representation hypothesis. They also found that sentence completion of an incomplete idiomatic string was facilitated by priming words related both phonologically and semantically to the target word, further suggesting that the content of the literal lemmas that comprise the idiomatic string are activated during production. They propose a slightly revised model in which idiomatic representations are instantiated as *super-*

lemmas, which serve as a representation of the syntactic properties of the idiom, similar to the previous proposal. The introduction of these *super-lemmas* allows idiomatic representations to enter into normal competition with other lemmas during the production process (thus *kick the bucket* could compete with *die* at the same level of lexical representation during production). As an additional difference, Sprenger et al (2006) propose that *super-lemmas* act as grammatical functions over their set of component lemmas, providing constraints on the possible structural configurations allowed by the idiomatic expression without the need of appealing to a separate association with one or more phrasal frames (see Kuiper, Van Egmond, Kempen & Sprenger, 2007 for further discussion on the advantages of super-lemmas).

Chapter 2: Self-Paced Reading

The process of sentence comprehension is like the process of solving a complex jigsaw puzzle. The comprehender must assemble the input into a syntactic structure and then interpret the meaning of that structure. The process of assembling an appropriate syntactic representation is not a trivial task, especially given the rampant ambiguity present in language. The psycholinguistic literature on ambiguity resolution is extensive, and the historical and empirical details needn't concern us here. However, the experiment presented in this chapter is essentially an ambiguity resolution experiment with the crucial difference that we will be investigating idiomatic/literal ambiguity, thus it is worthwhile to provide a brief overview of psycholinguistic approaches to ambiguity resolution, and specifically to demonstrate how the sort of ambiguity in this case is different than what is typically investigated.

2.1 Ambiguity Resolution

Ambiguity in language occurs because in many domains there is no one-to-one mapping between form and meaning. One sort of ambiguity possible in language is a *lexical ambiguity*. For example, a sentence such as (6) is ambiguous between the senses paraphrased in (6a) and (6b).

(6)   John arrived at the bank.

　　　(6a)   John arrived at a financial institution.

　　　(6b)   John arrived at the margin of a river.

This ambiguity hinges upon the fact that there are two unrelated lexical entries for the word *bank* in English. The processing of these sorts of sentences has been the subject of extensive investigation (Foss, 1970; Swinney, 1979; Onifer & Swinney, 1981; Rayner & Duffy, 1986; Tanenhaus, Leiman & Seidenberg, 1979 and others).

Another sort of ambiguity is *syntactic*, in which the ambiguity hinges upon the underlying syntactic structure of the utterance. Consider examples (7) and (8) for example.

(7)   John saw the man with the telescope.

(8)   Time flies

In (7) the sentence is ambiguous between a reading in which John uses the telescope to see the man, and one in which John sees a man who is using a telescope. The ambiguity here is due to the sentence in (7) having more than one potential syntactic structure, namely whether the prepositional phrase *with the telescope* modifies the verb or the object. The sentence in (8) is particularly interesting, as it represents a sort of hybrid case between the lexical ambiguity discussed earlier, and the syntactic ambiguity in (7). Here the sentence is ambiguous between a reading in which *time* is interpreted as a noun and *flies* as a verb, resulting in the familiar idiom, and a reading in which *time* is interpreted as a verb and *flies* as a noun, resulting in the somewhat odd notion that one is *timing flies*

(e.g. a biology researcher timing her flies). The former sort of syntactic ambiguity, much like lexical ambiguity, has an extensive literature (Frazier, 1979; Altmann & Steedman, 1988; Ferreira & Clifton, 1986; Trueswell & Tanenhaus, 1991 and others), as does the latter (Frazier & Rayner, 1987; MacDonald, 1993).

As discussed in the introduction, the sort of ambiguity we will be investigating is in some sense different from any of these cases however. Consider the sentence in (9).

(9)   John kicked the bucket yesterday.

Like other ambiguities discussed, the ambiguity in (9) hinges upon there being multiple possible representations for the phrase *kick the bucket*. One reading may be paraphrased as *John died yesterday*, and the other as *John impacted the bucket with his foot yesterday*. However, unlike the above ambiguities, these sorts of idiom/literal ambiguities are not easily classified into either a syntactic or lexical category. The literal interpretation of the sentence requires that *kick the bucket* be interpreted as a verb phrase, and that the words *kick* and *bucket* be interpreted literally. For the idiomatic interpretation, however, it is unclear how to proceed. Certainly the parser must interpret the string *kick the bucket* to mean *die*, however the rest of the process is generally unclear.

One possibility is that the ambiguity hinges upon interpreting *kick the bucket* as a syntactic structure vs. a lexical unit. However this then raises the question of *when* the

parser decides upon one of these interpretations, and how it arrives at that decision. Additionally, as the discussion on idiomatic representation in Chapter 1 mentioned, the relationship between the literal and idiomatic versions of *kick the bucket* and between the idiom *kick the bucket* and its components *kick* and *bucket* is not clearly understood. For example, under a words-with-spaces approach, there is no proposed relationship between a phrase like *kick the bucket* and the verb *kick*. Alternatively, under a decompositional view the meaning of the term *kick* (either directly or via phrase induced polysemy) may contribute in some way to the idiomatic meaning. Likewise, the configuration hypothesis and hybrid representation hypothesis suggest that literal *kick* contributes to the access of the idiomatic expression either by virtue of the parsing process (in the case of the former) or via a principled relationship in the lexicon between the idiom and the literal lemma (in the latter). Thus while sentences like (9) are clearly ambiguous, characterizing that ambiguity proves to be a difficult endeavor. The main aim of this chapter is to further explore that relationship.

2.2 Contextual Effects on Idiom Resolution

One pervasive finding in the ambiguity resolution literature is that the human language system is capable of making use of many sources of information to aid in the process (Garnsey, Pearlmutter, Myrs & Lotocky, 1997; MacDonald, Pearlmutter & Seidenberg, 1994; MacDonald, 1994; Spivey, Tanenhaus, Eberhard & Sedivy, 2002; Spivey & Tanenhaus, 1998; Tanenhaus, Carlson & Trueswell, 1989; Trueswell & Tanenhaus, 1994; Tanenhaus & Seidenberg, 1981, 1989; Burgess, Tanenhaus & Seidenberg, 1989). One

such source of information is prior linguistic context, and for our purposes we are interested in whether the system utilizes contextual cues to guide its interpretation of an ambiguous string. Such contextual influence on ambiguity resolution has been demonstrated in the lexical (Rayner, Pacht, & Duffy, 1994; Tabossi, Colombo & Job, 1987; Swinney, 1979; Tanenhaus, Leiman & Seidenberg, 1979), syntactic (Altman, Garnham & Henstra, 1994, Trueswell, Tanenhaus & Garnsey, 1994) and lexical-syntactic domains (MacDonald, 1993).

It is thus not terribly surprising that there is also evidence that such contextual information is available and utilized to navigate literal/idiomatic ambiguity of the sort we are interested in (Tabossi & Cacciari, 1988; Bobrow & Bell, 1973; Cacciari & Tabossi, 1988; Tabossi et al, 2009; Titone & Connie, 1999). Hence there is reason to believe that given appropriate sentential context, we can influence individual's interpretations of phrases ambiguous between a literal and idiomatic sense. However given the models reviewed in the previous chapter, this possibility raises several questions.

First the models predict different behaviors with respect to the influence of such a contextual bias. Hybrid models (Cutting & Bock, 1997; Sprenger et al, 2006) and the Configuration Hypothesis (Cacciari & Tabossi, 1988) both predict a certain priority for literal processing, thus while it may be possible to use contextual bias to reduce consideration of the literal representation, these sorts of models predict that some amount of literal consideration will proceed anyway. With respect to the idiomatic interpretation,

however, these models both indicate that contextual influences could, at least in principle, stop consideration of the idiomatic interpretation altogether. This is in contrast to the Direct Access Hypothesis (Gibbs, 1980), and Lexical Representation Hypothesis (Swinney & Culter, 1979). The Direct Access Hypothesis predicts that idiomatic interpretation can be made available without any priority of literal processing, and indeed claims that this is the normal route, while the Lexical Representation Hypothesis predicts the parallel processing of both meanings, but does not rule out in principle that sufficient contextual information could restrict the parser to one or the other interpretations. In the following experiment we will examine the question of whether literal processing has priority over idiomatic access, and will specifically explore the consequences of recovery when contextual bias is incongruent with the correct interpretation in globally unambiguous sentences.

2.3 Motivation

In this experiment, we explore the dynamics of idiom access and interpretation during comprehension, and the relationship between the literal and idiomatic meanings of ambiguous strings. By looking at the processing load induced by violating comprehenders' expectations about the idiomaticity/compositionality of ambiguous strings during comprehension, we aim to gain new insights into the time-course of idiom processing and investigate the activation and inhibition of competing parses during disambiguation.

Experiment 1 is an off-line sentence-completion study designed to provide crucial baseline information that is necessary for selecting the stimuli for the main experiment, and also provides us with detailed information regarding the inherent biases in our idioms.

Experiment 2, the main experiment in this chapter, is a self-paced reading study that explicitly examines what happens when the comprehender is required to change their expected interpretation of an ambiguous string. In this study, strings that are ambiguous between literal and idiomatic interpretations (e.g. *rush into*, *look up*) were embedded into sentences whose subject nouns were designed to bias either a literal or idiomatic interpretation. By manipulating the bias and resolution of these sentences (see Example 10), we are able to both examine the differences between idiom and literal processing during smooth comprehension, and also during recovery from misplaced expectations. Examination of this recovery process can help shed light upon whether idiomatic and literal processes proceed in parallel or serially, as well as provide details regarding the dynamics at work during the comprehension process.

(10a) Literal Bias - The daring fireman … rushed into

(10b) Idiomatic Bias - The foolish entrepreneur … rushed into

(10c) Literal Resolution - …rushed into the building…

(10d) Idiomatic Resolution - …rushed into the decision…

For the purposes of our study we will be using phrasal verbs, which are particularly useful in this endeavor. As mentioned in chapter 1, the literature in support of the Configuration Hypothesis (Cacciari & Tabossi, 1988) suggests that literal processing may stop at some point once the idiomatic nature of the string is recognized. Phrasal verbs such as *look up* provide a test case in which the idiomatic meaning is plausible. Additionally, these phrases are short enough that by the time the sentence is disambiguated, it is less likely that the participant will have had enough time to fully decide upon one interpretation. This is valuable, as it allows us to examine whether the processing of the literal expression is indeed primary. In longer idioms, for example, it may be the case that by the end of the idiom the parser has already had sufficient time to decide upon the idiomatic meaning and to fully suppress its literal interpretation. Using phrasal verbs may allow us to see into this process before such suppression occurs.

2.4 Experiment I: Sentence Completion

Experiment 1 is a sentence-completion study that provides foundational data regarding our stimuli for Experiment 2. This study has two main aims: First, Experiment 1 was used to ensure that the critical biasing contexts necessary for our design are indeed powerful enough to create expectations towards idiomatic vs. literal interpretations. Second, this experiment provides a detailed measure of how strong these idiomatic and literal biases are for each particular item. This is important given that there is no reason to assume that the strength of the literal bias of one of our items will be equivalent to the strength of the idiom bias of that item. Furthermore, there is no reason to assume that our

items could be equi-biased (i.e. with our literally biased condition biasing a literal interpretation equally strongly as our idiomatically biased condition biases the idiomatic interpretation) *even in principle*. Certain idioms may resist literal interpretation even in the best scenarios, while others may be more flexible. Thus the results of this experiment provide a metric of bias in our items which will be incorporated into the regression-based data analyses conducted for the Experiment 2.

Thus by collecting data regarding the strength of the bias in each of our items we not only get a metric of how balanced our items are with respect to their bias, but can also then use this metric of bias strength later by factoring it into further analyses in Experiment 2. This will allow us to regress out effects due to unbalanced bias and examine effects of our contextual bias manipulation while controlling for potential item-specific effects that may confound our results. As mentioned in chapter 1, there is significant difficulty in defining idioms as a coherent class, and while some more frozen idioms (e.g. *kick the bucket*) seem to be generally accepted as belonging to that class, phrasal verbs, such as those used here, are less prototypical members of the idiom class. Thus this careful norming procedure, and the addition of bias strength information to our statistical models is critical in order to help mitigate concerns that these sorts of expressions may behave differently to other idioms.

2.4.1 Method

2.4.1.1 Participants

24 adult native speakers of English participated in this experiment over the Internet.

2.4.1.2 Procedure & Design

Participants completed the study on the Internet. Each item was displayed on a separate webpage, as shown in Figure 5. Participants were instructed to provide a natural completion for each sentence fragment, and to click the "Next Trial" button to move onto a new trial.

This is trial 14 of 40.
Please continue the sentence in the space provided and then click the "Next Trial" button.

The athletic runner, who wore contact lenses, ran across [   ]

(NEXT TRIAL)
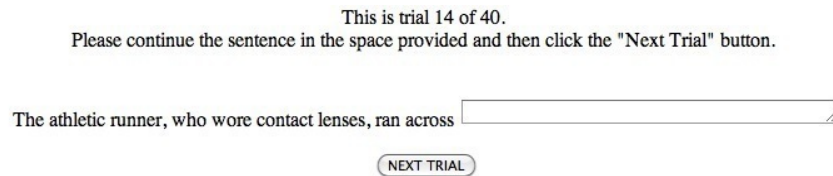
Figure 5: A typical trial in Exp 1, sentence completion study.

As shown in example (11a), we manipulated the region of the sentence that preceded the ambiguous target string (e.g. *ran across*), in order to create a bias toward an idiomatic interpretation (11a) or a compositional interpretation (11b).

	Bias	Lead-In	Verb+Preposition
(11a)	[The frustrated programmer] Idiom bias	[who wore contact lenses]	[ran across] …
(11b)	[The athletic runner] Compositional bias	[who wore contact lenses]	[ran across] ….

45

The sentences were designed such that each sentence consisted of a *Bias*, a *Lead-In* and the critical string, consisting of a verb followed by a preposition. As shown in (11), the two conditions differ only in their *Bias* component.

The idioms used in this experiment as well as Experiment 2 are all phrasal verbs, e.g. *ran across*, *rushed into*, *fit in*. Phrasal verbs were used for several reasons. First, they are highly frequent in both their literal and idiomatic usage. Second, they allow us to better control for string length and syntactic structure in our self-paced reading stimuli. Finally, as mentioned earlier, the brevity of these items provides advantages with respect to how quickly the sentence can be disambiguated after their appearance. The *Biases* were designed to bias the participant toward either the idiomatic (11a) or literal interpretation (11b) of the *Verb+P* component.

This experiment included 36 targets and 22 fillers (see Appendix I). Each participant saw exactly half of the target sentences and all of the fillers. Within each list exactly half of the target sentences were biased idiomatically and half were biased literally. Because no participant saw all of the targets, four lists were constructed that partially overlapped to control for possible item effects. Four counterpart lists were created which differed only in that the conditions of targets were switched. Finally, to control for potential order effects, reverse lists were created for each of these eight base lists for a total of 16 lists.

2.4.2 Experiment 1 Results

2.4.2.1 Coding

Because our aim is to assess how strongly the contexts bias particular ambiguous strings towards idiomatic and literal interpretations, participants' continuations were coded for whether the critical verb+preposition string was interpreted idiomatically (ex. 12) or compositionally (ex. 13). Cases where this was unclear were coded as 'ambiguous' (ex. 14). Overall, less than 5% (4.59%) of responses were coded as ambiguous. These trials were excluded from subsequent analyses.

(12) Idiomatic Continuations (critical string marked in italics)

a. The creepy detective, who was wearing a cool hat, *looked into* **the case of the jade falcon.**

b. The bored intellectual, who had won many contests, *dove into* **the book.**

c. The worried mother, whose husband was in the army, *ran up* **her Visa bill on postage.**

(13) Compositional Continuations

a. The happy astronomer, who was wearing a cool hat, *looked into* **the night sky with his telescope.**

b. The bored intellectual, who had won many contests, *dove into* **the crystal clear water beneath him.**

c. The worried mother, whose husband was in the army, *ran up* **the stairs**.

(14) Ambiguous Continuations

a. The graceful ballerina, who worked very hard, *turned in* **her tutu.**

b. The rich investor, who knew the value of everything, *traded in* **the stock market.**

c. The foolish entrepreneur, who liked living on the edge, *rushed into* **it**.

In order to assess the bias strength of the contexts for each of the target strings, we computed a *congruent trial proportion* for each item in each condition by dividing the number of trials in which the continuation was congruent with the intended bias (i.e., the number of idiomatic continuations after idiomatic-bias nouns, and the number of compositional continuations after compositional-bias nouns) by the total number of non-ambiguous tokens for that target. This resulted in two values for each target item: (i) *Literal bias*: the proportion of literal continuations when the item is biased literally, (ii) *Idiom bias*: the proportion of idiomatic continuations when the item is biased idiomatically. For later use we also computed a *Sentential Bias* which was the difference between the *Idiom Bias* and *Literal Bias*.

2.4.2.2 Target Selection

Experiment 1 contained 36 target items. Sixteen of these were selected to be used in the self-paced reading study, Experiment 2. The aim of Experiment 2 is to investigate the effects of misplaced expectations on processing, and thus we wanted to ensure that the contexts that are used in that study are powerful enough to reliably bias the ambiguous strings towards an idiomatic or compositional interpretation, but that the strings nevertheless remain ambiguous. Thus, items in which individuals did not provide any

unambiguous continuations congruent with the expected bias were removed. After this initial culling, we also removed any items in which the success of our biases were extremely asymmetric.

These criteria excluded 14 of the 36 targets, leaving us with 21 items as valid targets. 16 of these were then hand-selected for use in Experiment 2 based on the ease with which suitable items could be constructed.

2.2.2.3 Analysis

In this section, we report in more detail on the Literal-Bias and Idiom-Bias proportions for the 16 selected targets.
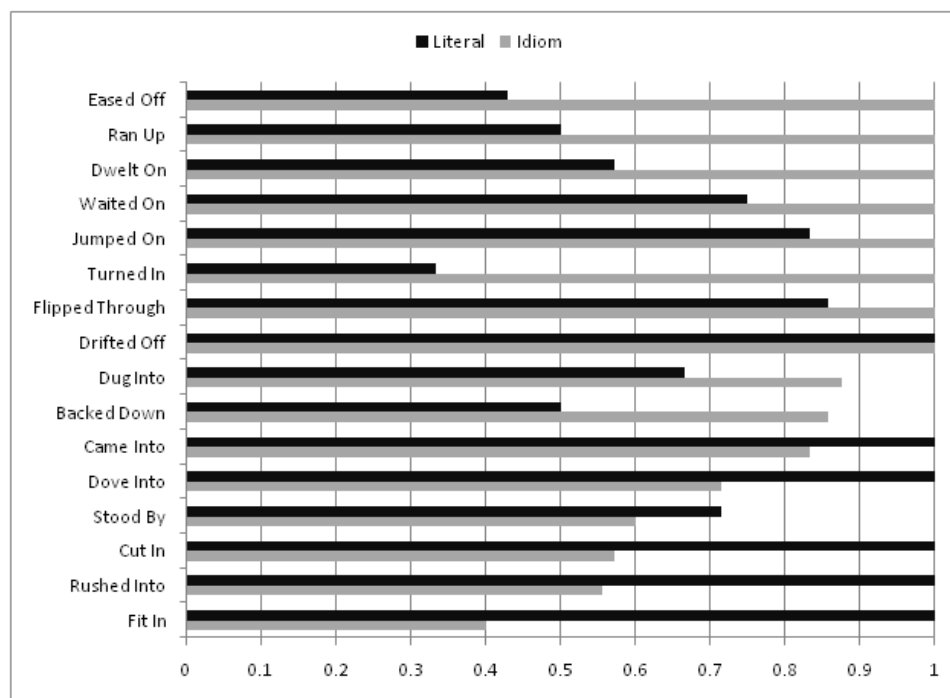
Figure 6: Absolute bias for each of our verb + p stimuli as proportion of sentence continuations in which participants provided the expected (e.g. congruent with the intended literal or idiomatic bias) sentence continuation.

Figure 6 shows the raw bias by verb + p. The dark bars indicate the proportion of literally biased trials in which participants provided a literal continuation. The light bars indicate the proportion of idiomatically biased trials in which participants provided an idiomatic continuation. Thus, for each verb we can see the results for both conditions and the length of each bar represents the proportion of trials for a particular condition in which participants provided continuations congruent with the bias. We can see here that while there are serious differences between our items, the overall trend is for continuations to be congruent with the intended biases.
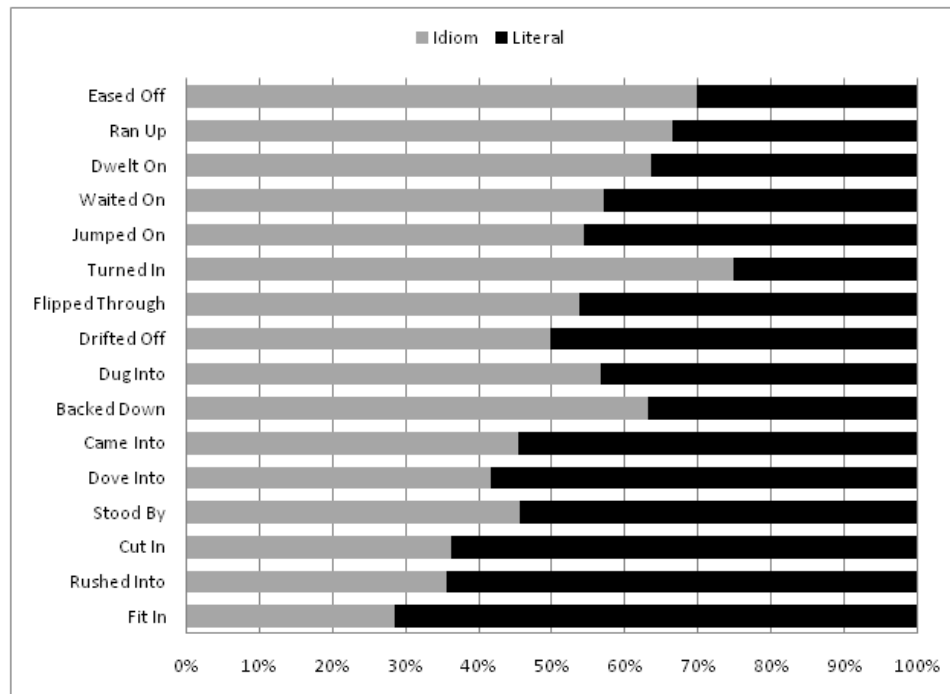
Figure 7: Relative biases of our verbs presented as percentages of sentence continuations congruent with our idiomatic (light bars) and literal (dark bars) out of the total number of unambiguous sentence continuations.

Figure 7 shows the relative bias strengths by verb over all trials for that item. The data is the same as in Figure 6, but this presentation of it is useful as it gives us a clearer notion of how balanced each item is with respect to the two biasing conditions. Again we can see that our items vary considerably. In the case of *eased off* we can see that the relative bias is in favor of the idiomatic interpretation. This means that in our items for *eased off*, participants are more likely to provide an idiomatic continuation when biased idiomatically than they are to produce a literal continuation when biased literally. For an item such as *fit in* we see the opposite situation, with *drifted off* being our only perfectly

balanced item. As with the previous figure, we see a large amount of variability in our items.

In addition to giving us a measure of how variable the items, a key aim of Experiment 1 is to ensure that the biasing contexts are indeed sufficiently biasing in the intended directions (towards idiomatic or towards compositional interpretation). To test this statistically, we ran a one-way ANOVA examining the effect of condition [Idiom-Bias, Literal-Bias]. The test revealed a significant effect of condition (F=81.3, df=1, p<.001), confirming that our contexts are indeed effective at triggering the intended bias.

The same data were also analyzed using a linear mixed-effects model (Jaeger, 2008; Baayen, Davidson & Bates, 2008), testing for effects of Condition and adding random effects of Subject and Item to remove potential variability due to individual differences in participants or items. These analyses replicated the main effect of condition (β = -.64, t = -11.19, p<.001). Further investigation of this model, however, revealed significant deterioration in the model fit (χ^2 = 7.36, df = 1, p < .01) if the random effect of items is removed. This confirms our suspicion that while the overall trend is for congruency, there are significant differences between the different targets. Thus we must be careful to include Items as a random factor in later experiments. To this end, later investigations will rely upon mixed-effect model testing, as it more easily allows us to account for by-item variability in our analyses without losing the advantages gained by having a direct metric of bias strength.

Investigating the effect of subjects reveals that removing the random effect of subjects from our model does not result in a significantly lower fit ($\chi^2 = 0$, df = 1, p = .99), thus we can conclude that individual differences between participants is not a significant factor driving these results.

2.4.3 Experiment 1 Discussion

Generally this experiment confirmed that the participants can be biased solely by the meaning of the sentential subject, and provided a set of normed stimuli for use in the self-paced reading experiment. Our results also confirmed our suspicions our idioms do not behave as a unified class with respect to their default interpretations and their behavior in biased contexts. The results of this experiment are thus beneficial, as they will allow us to control for this variable behavior in our next experiment.

2.5 Experiment 2: Self-Paced Reading

The 16 target items, selected and evaluated in Experiment 1 were used to examine how readers' expectations affect online processing during the comprehension of idiomatic and literal expressions. Additionally we aimed to test how quickly readers recover from incorrect expectations, and if there are any differences in their recovery for idiomatic vs. literal expressions.

Of critical interest in this experiment is whether or not literal processing has priority over idiomatic interpretation as predicted by the Configuration Hypothesis (Cacciari & Tabossi, 1988) and Hybrid Representation Theory (Cutting & Bock, 1997; Sprenger et al, 2006), and also what this priority means for recovery processes. We expect to replicate the finding in the literature that access to literal meaning is slower than access to idiomatic meaning during normal processing. However, by examining what happens when processing is perturbed we hope to gain insights into the process of activation and inhibition that occurs during idiom/literal ambiguity resolution.

In general, if literal processing has priority over idiomatic access then we might expect that in the processing of these short phrases, recovery to a literal interpretation would be easier than recovery to an idiomatic interpretation. The logic here is that the parser must do some amount of literal processing regardless of whether it believes the upcoming string is idiomatic or literal. Thus if the parser mistakenly attempts to parse a literal structure as idiomatic and must then revise its interpretation, it may be able to make use of the (perhaps partial) literal processing that it has already accomplished and thus be able to recover from such mistakes quickly. Thus while we expect to replicate the findings in the literature that idiomatic interpretation is faster than literal interpretation in normal processing, we might expect the inverse during recovery.

If, however, idiomatic processing has priority, as proposed by the Direct Access Hypothesis (Gibbs, 1980) then we would expect the opposite finding, where recovering

the literal meaning after mistakenly interpreting an expression as idiomatic would be difficult, while recovery of the priority (and hence already considered) idiomatic meaning would be easier. Finally, if both idiomatic and literal processes proceed in parallel, as predicted by the Lexical Representation Hypothesis (Swinney & Cutler, 1979), we expect recovery to be equally difficult or easy regardless of whether one needs to recover the idiomatic or literal interpretation.

2.5.1 Method

2.5.1.1 Participants

32 adult native speakers of American English participated. None of these participants participated in the previous experiment.

2.5.1.2 Materials

This experiment consisted of 16 targets, and 32 filler sentences. In the targets, we manipulated bias induced by the sentential subject (**Idiom Bias** vs. **Literal Bias**) and the resolution of the sentence (**Idiom Disambiguation** vs. **Literal Disambiguation**) – i.e., whether the ambiguous string turned out to be used idiomatically or compositionally.

Target items consisted of *Bias*, a *Lead-In*, a *Verb+P*, and a *Resolution*. Within a particular target item the verb+P, lead-in and spillover were kept constant, so sentences differed only in their *Bias* component and *Resolution*. For example:

(15) [Bias] [Lead-In] [Verb+P] [Resolution] [Spillover]

(15a)   [The hungry waitress] [who had been working all day] [dug into] [the sandwich] [just after noon on Sunday]

(15b)   [The hungry waitress] [who had been working all day] [dug into] [the tomb] [just after noon on Sunday]

(15c)   [The daring archaeologist] [who had been working all day] [dug into] [the sandwich] [just after noon on Sunday]

(15d)   [The daring archaeologist] [who had been working all day] [dug into] [the tomb] [just after noon on Sunday]

The Bias + Lead-In + Verb+P sequences came from Experiment 1, which showed that the biases are indeed effective at pushing comprehenders towards an idiomatic or compositional expectation. The *Resolutions* disambiguated the sentence toward either an idiomatic (Idiom-Resolution) (15a,c) or literal (Literal-Resolution)(15b,d) interpretation, and the spillover region was added to create a buffer between the disambiguating resolution and the end of the sentence. Thus a given item had identical *Lead-In*, *Verb+P*, and *Spillover* with the *Bias* and *Resolution* varying by condition. We generated four lists using a Latin Square design and four reverse lists to control for possible order effects.

Fifteen comprehension questions were interspersed among the target and filler items. Five of the questions probed basic information from the beginning of the immediately preceding sentence, 5 from the middle, and 5 from the end. This was done to encourage participants to attend to all parts of the sentence.

2.5.1.3 Procedure

Participants completed four practice trials before the main experiment. We used a standard moving-window self-paced reading task, where the words of a sentence are initially masked with hyphens (-). Participants pressed a key to unmask the first word of the sentence, and when they were ready, pressed the key again which re-masked the first word and unmasked the second word. In a given trial participants would continue this procedure until the final word of the sentence was unmasked, and would then press the key again to proceed to the next trial. Occasionally, participants would be presented with a yes/no comprehension question based upon the content of the sentence in the previous trial. The experiment itself was designed in E-Prime (Psychology Software Tools). Responses were recorded using a button-box.

2.5.2 Predictions

Analyses were performed over three regions: the pre-verbal region, verb+p region, and disambiguation region as shown for our stimuli in Table 1. For each region results are reported both for the average over the entire region as well as for each individual word within the region. We performed both region-level and word-level analyses primarily due concerns that observed effects may not manifest immediately on our critical words. This effect is known as spill-over in self-paced reading analyses, and given the brevity of our items we were concerned that processing difficulty, as measured by reaction time in

pressing a button, might not manifest immediately upon encountering the verb or the adjacent disambiguating term.

		Pre-Verb	Verb		Disambiguation		
The hungry waitress	who had been	working all day	dug into	the	sandwich	just after	noon on Sunday
The hungry waitress	who had been	working all day	dug into	the	tomb	just after	noon on Sunday
The daring archaeologist	who had been	working all day	dug into	the	sandwich	just after	noon on Sunday
The daring archaeologist	who had been	working all day	dug into	the	tomb	just after	noon on Sunday

Table 1: Target sentences with critical analysis regions marked

In particular, we chose to have participants read each word individually (including functional words such as *the* and prepositions) rather than chunking words together to gain maximum insight into the timing of any observed effects and so that we needn't make any assumptions regarding the status of the prepositional component of our phrasal verbs (e.g. presenting *[look up][the number]* may create different expectations than presenting *[look][up the number]*). As seen in table 1 above, however, we were careful in choosing our regions for analysis such that they included minimal additional content beyond the targets of interest.

Our specific predictions vary by region, as in each region new information is available to the parser. We will first describe the specific predictions for each of our three regions below, and then turn to our results and statistical analyses. Except where noted all analyses were conducted using linear-mixed effects models as in experiment 1, modeling fixed effects of Bias as well as our metric of Bias Strength obtained in experiment 1 and random effects of subjects and items. Fixed effects of Congruence were also modeled during the disambiguation region, but as sentences differed only in their initial Bias

before this region, this condition was collapsed for analyses of the pre-verb and verbal regions.

For the pre-verbal region participants have not yet encountered the critical phrasal verb nor has the sentence been disambiguated. Since the disambiguation has not yet occurred, the sentences differ only in their Bias. While we predict that Bias will affect the interpretation of the verb + preposition, this has also not yet been encountered in this region. thus we do not expect any significant effect of bias in this region.

At the verb+p region, participants first encounter the ambiguous phrasal verb, but again the sentences differ only in their Bias. Over this region we expect to replicate findings in the literature that idiomatic expressions are processed more rapidly than literal expressions. Hence, we expect to find faster reading times over this region when in Idiomatically Biased conditions as compared to Literally Biased conditions. However we remain cautious regarding this prediction. First, as noted earlier, we are using phrasal verbs in part because of their brevity, however given our self-paced reading task, this brevity may work against us for replicating the processing speed advantage for idiomatic expressions as our idioms may be too short to observe any speed advantage during this region. However, if we do still find a speed advantage over this region it would be a particularly strong replication of previous findings.

The primary region of interest is the disambiguation region, which includes the disambiguating noun from the *resolution* region, as well as the first two words of the spillover region. For this region we proceed with two analyses. The first examines the effects of Bias and Congruence over the region. We expect a main effect of congruence, with congruent trials being faster than incongruent trials. Essentially, individuals should be faster to process the sentence when the sentence is disambiguated congruently with their expectations. We may also see a main effect of bias, with literal trials being generally slower than idiomatic trials overall as result of the aforementioned processing advantage for idioms over literals.

To specifically examine the effects of recovery over the disambiguation region we exploit the fact that Literal-Congruent and Idiomatic-Incongruent trials are, with the exception of their sentential subject, identical sentences, and similarly for Idiomatic-Congruent and Literal-Incongruent trials. Crucially our examination will focus upon how the trial is resolved (either Literally or Idiomatically) and whether or not the preceding bias translates into faster processing times over this region. If literal processing is obligatory and can thus be relied upon regardless of expectations, then we would expect that processing will *not* be perturbed regardless of bias for literally resolving trials, but will be perturbed for idiomatically resolving trials if the preceding bias lead participants to expect a literal resolution. If literal processing is not obligatory then we expect either no effect or equal disturbances due to bias in both cases.

2.5.3 Results

2.5.3.1 Data Preparation

To prepare out data for analysis RTs below 100ms were removed. The removal of these data points affected less than 1% of the data (.68%). In addition, extreme outliers which were more than 5 trimmed standard deviations from the trimmed mean were removed. Overall there were no outliers on the low end. On the high end, several outliers were detected and removed, affecting 3.7% of the data. The resulting dataset was then log-transformed to adjust for the positive skew typical of reaction time data, and all further analyses were performed on these log-transformed values.

2.5.3.2 Pre-Verb

The pre verb region is defined as the three word region preceding (and not including) the verb. Crucially, we do not expect any effects during this region. It is possible we may see some effects of bias, as this factor manifests as different sentential subjects, and in all cases this region is the final three words of a phrase which modifies that subject. The interest in performing this analysis is that this region of our sentence provides a useful test-case for establishment of our model for the more interesting regions.

The overall pre-verb results by condition are given in Figure 8 below. As effects of the resolution manipulation have not yet occurred at this stage, those conditions have been collapsed.

Reading times over the pre-verb region
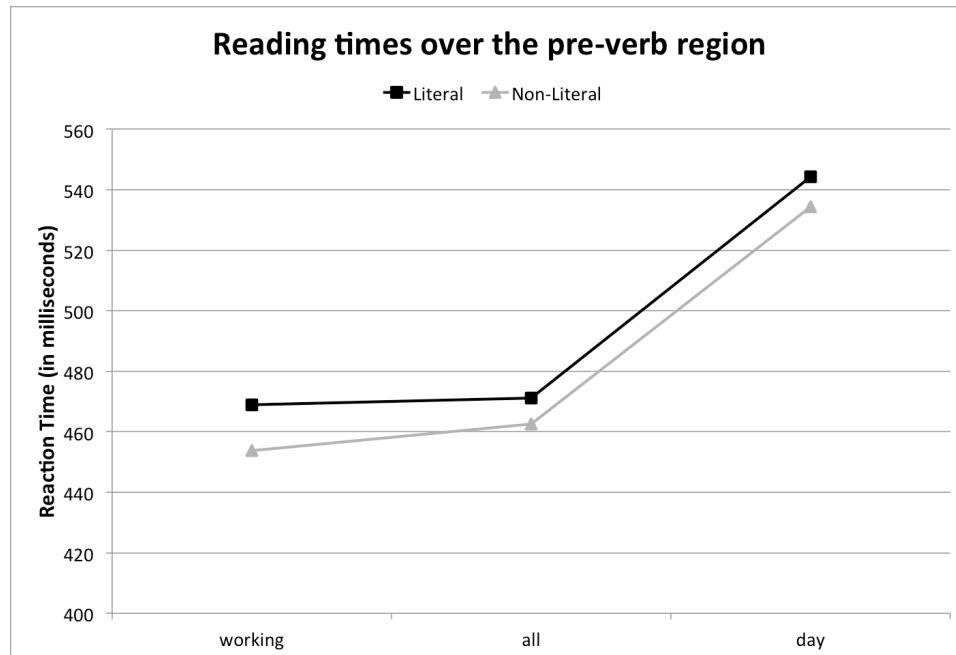
Figure 8: Reading times for literally and idiomatically biased conditions over the three word region preceding the verb. Reading times for literally biased trials are represented by the dark line, while reading times for idiomatically biased trials are represented by the light line.

Analysis of the Pre-Verbal Region				
	Beta	CI	t	p
Region	-0.038	(-0.08 , 0.011)	-1.66	.
V-3	-0.038	(-0.089 , 0.017)	-1.43	0.152
V-2	-0.033	(-0.087 , 0.023)	-1.16	0.246
V-1	-0.045	(-0.045 , -0.112)	-1.38	0.168

Table 2: Effects of bias over the pre-verbal region and over each word in that region. P values less than .1 are marked with a period (.), and p values less than .05 are marked with an asterisk (*).

Figure 8 suggests a slight speed advantage for the non-literal biasing sentences even before the onset of the verb. However statistical analyses (see Table 2) reveal this effect to be marginal, and only when we average over the entire region. For individual points

there is no significant effect of either bias or our metric of bias strength obtained in experiment 1.

In general no effects were expected in this region, as the words involved are controlled across conditions and the effect of our manipulation is not expected to have any serious impact until the verb. These results are in line with those predictions. Model comparisons over this region do, however, reveal significant model deterioration if we remove the random effects of subjects or items from our analysis, thus going forward we will continue to model these effects on consecutive regions.

2.5.3.3 Verb

Figure 9 shows the reaction times for the verbal region, which is the three-word region which includes the critical verb, preposition, and the word immediately following the preposition (typically the definite determiner). Again, since the disambiguation has not yet occurred sentences differ only in whether their subjects bias participants literally or idiomatically. Hence, Literally Biased and Idiomatically Biased conditions are still identical at this point (e.g. *The hungry waitress / daring archaeologist who had been working all day dug into the…*) and so the different congruency conditions have been collapsed.
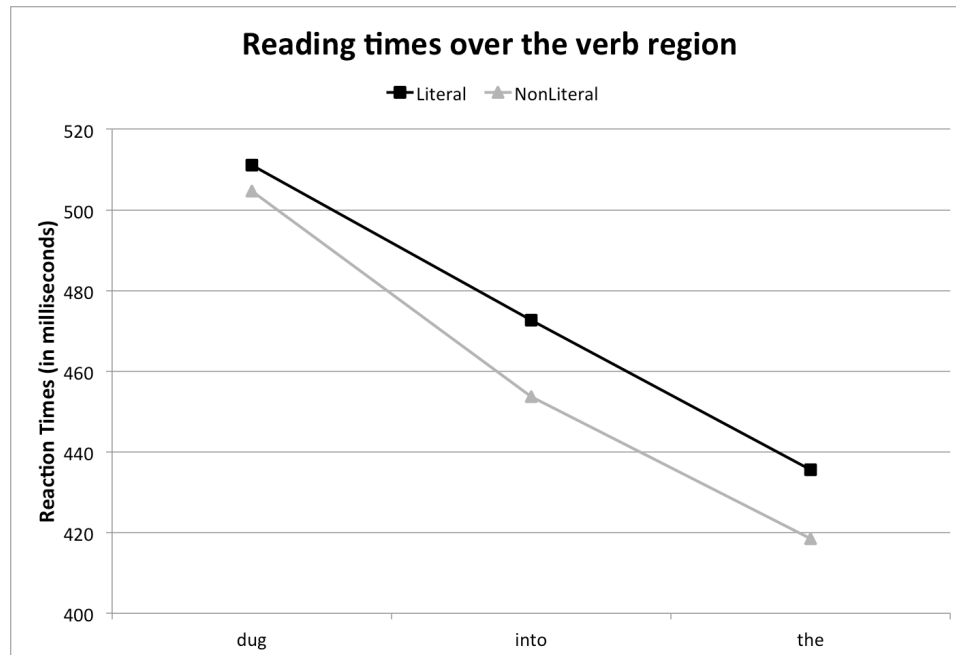
Figure 9: Reading times for literally and idiomatically biased conditions over the three word region starting with the verb. Reading times for literally biased trials are represented by the dark line, while reading times for idiomatically biased trials are represented by the light line.

As can be seen in Figure 9, reading times in the literal bias condition are longer than reading times in the idiomatically biased condition, i.e. participants read the verb region faster when they expect an idiomatic interpretation of the phrasal verb than when they expect a literal interpretation. This observation is confirmed by statistical analyses: An analysis that collapses across the three words shows that the Idiomatically Biased condition is indeed read significantly faster than the Literally Biased condition (significant effect of bias, see Table 3).

Analysis of the Verbal Region				
	Beta	CI	t	p
Region	-0.045	(-0.087 , -0.005)	-2.2	*
V	-0.022	(-0.084 , 0.036)	-0.72	0.47
V+1	-0.042	(-0.095 , 0.008)	-1.63	0.104
V+2	-0.048	(-0.101 , 0.001)	-1.91	.

Table 3: Effects of bias over the verbal region and over each word in that region. P values less than .1 are marked with a period (.), and p values less than .05 are marked with an asterisk (*).

Analyses of the individual words in this region fail to reveal any significant effects on either the verb or the preposition, but show a marginal effect on the word following the preposition. Given how brief our particular expressions are it is surprising that we find any effects at all this early on. Recall from our predictions, that while we generally expected to replicate the findings in the literature that idioms are processed faster than literal expressions, we were concerned that the use of phrasal verbs rather than larger idioms might result in this speed advantage not manifesting on the verbal region itself. Our results here are consistent with both the processing advantage for idiomatic expressions (as we find that idiomatically biased trials to be significantly faster overall when we average over the region) and also consistent with our concern regarding their brevity (as analyses on the individual words reveal only a marginal effect, and only later in the region).

2.5.3.4 Disambiguation

Figure 10 shows the results by condition for the disambiguating region, which consisted of the three word region immediately following the verbal region.
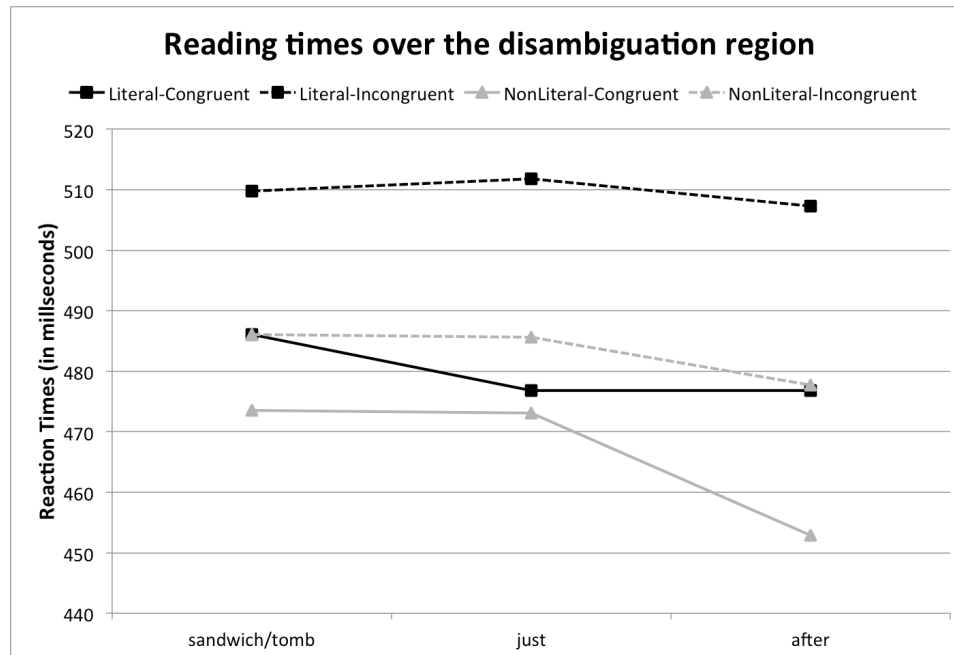
Figure 10: Reading times for literally and idiomatically biased conditions over the three word

disambiguation region. Reading times for literally biased trials are represented by the dark line, while

reading times for idiomatically biased trials are represented by the light line. Congruent trials are

represented with solid lines, and incongruent trials with dashed lines.

As can be seen in Figure 10, overall reading times on congruent trials (e.g. where the resolution supports the expectations created by the bias) appear to be faster than incongruent trials (e.g. where the resolution goes against the expectations created by the bias). Indeed the fastest condition in general appears to be for Idiomatically Biased - Congruent trials. Additionally, echoing the patterns seen in the preceding region, we still see effects of the initial bias, with Literally biased trials being generally slower than their idiomatically biased counterparts.

Analysis of the Disambiguation Region				
Region	Beta	CI	t	p
Bias	-0.06	(-.102 , -.019)	-2.86	*
Congruence	0.41	(-.001 , .0834)	1.97	*
Bias x Congruence	-0.01	(-.092, .077)	-0.21	0.83
V+3				
Bias	-0.0371	(-0.094 , 0.018)	-1.33	0.1833
Congruence	0.047	(-0.01 , 0.1)	1.69	.
Bias x Congruence	-0.0177	(-0.135 , 0.091)	-0.32	0.7513
V+4				
Bias	-0.0375	(-0.089 , 0.014)	-1.44	0.1495
Congruence	0.0448	(-0.006 , 0.099)	1.73	.
Bias x Congruence	-0.0143	(-0.118 , 0.088)	-0.27	0.7839
V+5				
Bias	-0.0594	(-0.113 , -0.005)	-2.17	*
Congruence	0.0545	(0.002 , 0.111)	2	*
Bias x Congruence	-0.0199	(-0.125 , 0.096)	-0.36	0.7157

Table 4: Effects of bias, congruency and their interaction over the disambiguation region and over each word in that region. P values less than .1 are marked with a period (.), and p values less than .05 are marked with an asterisk (*).

The statistical results are reported in Table 4. Our analyses reveal a significant main effect of both bias and congruence with no interaction when averaging over the full region. Analyses of the individual words in this region reveal marginal effects of congruence for the early portion of the region with the main effects being driven by the final word in the region. At this point we see a significant main effect of both bias and congruence. We also see a significant main effect of our metric of bias strength from experiment 1.

The main effect of congruence can be interpreted in line with the standard findings in the ambiguity literature. Participants are faster when they get what they expect. Note that at

the verbal region we interpreted the main effect of bias as confirming the pervasive *idiom-advantage* reported in the literature. This is because we expect from experiment 1, that during the verb region participants are generally interpreting the incoming language to be congruent with the bias.

This is not the case in this region. For most items the initial word of this region *is* the word which globally disambiguates the sentence one way or another. Additionally, regardless of how participants are interpreting the verb+p, the actual content of this region is *always* compositional. Thus any effect of bias in this region cannot be directly due to *literal-advantage*. Additionally, since we already see significant effects of bias on the word immediately before this region, *but not on the first word of this region*, the results suggest that this bias effect, and the one seen in the previous region, represent distinct processing incidents rather than some sort of spillover.

Thus we interpret this effect as a cost of *recovery*, broadly defining recovery as the (re)activation/re-ranking of the correct parse. To specifically evaluate the processing costs on *recovery* we conducted t-tests on the reaction times for our sentences over the averaged disambiguation region. The results are presented in Figure 11.

Averaged Reading Times over the Disambiguation Region
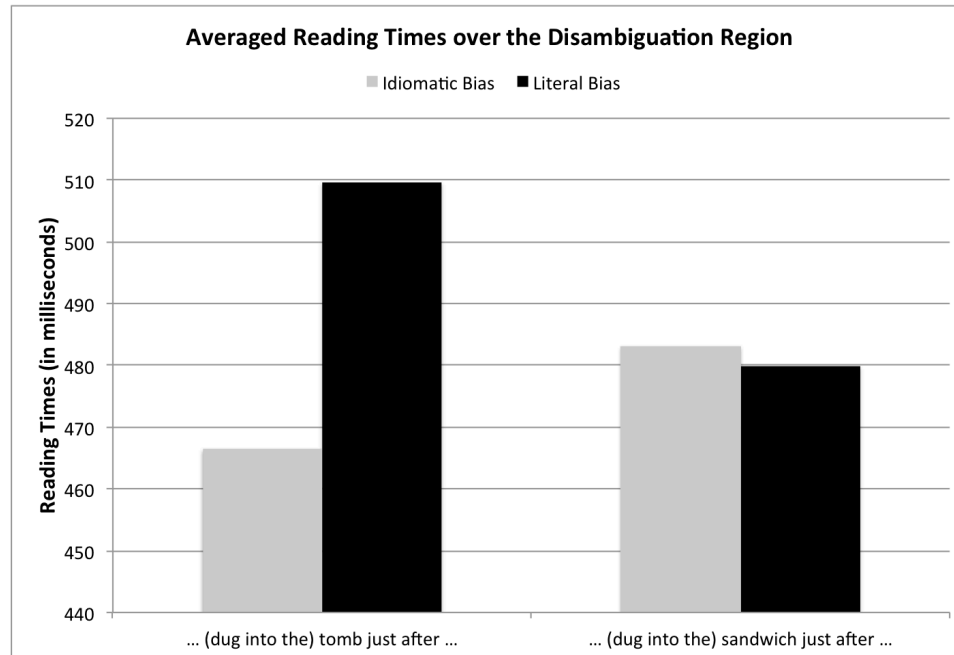
Figure 11: Average reading times over the disambiguation by resolution with literally resolving trials on the left and idiomatically resolving trials on the right. Light bars represent prior idiomatic bias while dark bars represent prior literal bias.

First note that we've altered the way we talk about our conditions slightly. Previously we were interested in *bias* and *congruence* with that bias. In this case we primarily interested in the *correct interpretation*, as this is what the participant should be constructing at this stage, rather than whether or not that interpretation is congruent with the bias. Thus we recode *congruence* (incongruent vs. congruent) as *resolution* (literal vs. idiomatic). Crucially what we are interested in here is whether, at this reconstruction stage, participants seem to be affected by the sentential bias. Figure 16 shows the absolute difference in reaction times over the averaged disambiguation region for literally and non-literally resolving trials.

Analyses were performed by calculating difference scores for Literal Resolving and Non-Literal Resolving trials by subjects and items, and using a two-tailed t-test to determine whether the difference scores differed significantly from 0 (i.e. does the previous bias have any significant effect on the reading times in the disambiguation region). For literally resolving trials the results demonstrate no significant advantage of bias type [Subjects: t(31) = .625, p = .54; Items: t(15) = -.01, p = .99]. For non-literally resolving trials there is a significant advantage for Non-Literally biased trials (mean 466ms) over Literally Biased trials (mean 509ms) by both subjects and items [Subjects: t(31) = 3.87, p < .001, Items: t(15) = 3.03, p < .01].

The results show that for literally resolving trials, that is those trials in which the final correct interpretation was the literal one, we find no significant effect of bias over the disambiguating region. Thus at the point participants are constructing the *correct* literal meaning, it does not appear to matter whether they were biased to expect the meaning to be literal or not. For idiomatically resolving trials, however, we find a significant effect due to bias. Thus participants constructing or re-constructing an idiomatic interpretation are strongly affected by their expectations.

2.6 Discussion

The main aim of this experiment was to examine the processing effects during recovery from incorrect expectations regarding the idiomaticity of an ambiguous phrasal verb. Our prediction was that if literal processing is obligatory, then individuals should be able to

rely upon this literal processing even in cases in which contextual bias leads them to incorrectly interpret an ambiguous string as idiomatic, but that no such fall-back mechanism should be available when incorrectly interpreting an ambiguous string as literal. The main results in this experiment are in line with this prediction.

Looking at reading times over the disambiguation region, we find that during real-time processing, when a comprehender is expecting an idiomatic interpretation, they find it easier to abandon and switch to a literal interpretation, rather than abandoning a literal interpretation in favor of an idiomatic interpretation. In short, participants seem to be able to process strings like *looked up the stairs* equally rapidly regardless of whether contextual bias supported the literal or idiomatic interpretation of *look up*. However, when processing strings such as *look up the number* participants are very fast when contextual bias supports the correct idiomatic interpretation of *look up* and much slower when contextual bias fails to support this interpretation. This result has several implications.

First, as discussed earlier, there is a wealth of evidence that idiomatic strings are processed more quickly than matched literal strings, and indeed we found this to be the case even with our stimuli. Recall that we found that processing of the verbal region was faster the string was biased idiomatically as compared to when the string was biased literally. This result is expected given the literature, but it is somewhat surprising that we would see such an effect given the brevity of our strings. However, this becomes more

surprising when considering the apparent recovery effects that we observed in the disambiguation region. The processing profile we see during recovery suggests that the process of integrating the idiom into the rest of the sentence is indeed rapid when the idiom is expected. However, when it is unexpected, we see a severe slowdown. Critically we do not see this slowdown with unexpected literal expressions.

Applying this result to current models of idiom processing supports the view that some degree of literal processing is obligatory. The lack of an effect of bias on processing the literal structure suggests that the literal representation of these strings is either active or quickly recoverable. One way to explain this is to claim that upon realizing one's mistake, the literal representation can be quickly retrieved and integrated into the sentence. However, as mentioned before, this view is not compatible with the general finding that retrieval of literal meaning is generally a slower process than retrieval of idiomatic meaning. A more suitable explanation is that the literal meaning is at least partially active anyway, hence the process of recovering from an incongruent bias is rapid by virtue of being able to fall back on this partially active parse. This behavior is predicted under both the Configuration Hypothesis and the Hybrid Models of idiom representation. In the former case, some processing of literal meaning occurs prior to idiomatic recognition, and while the parser may not necessarily complete the task of computing the literal meaning of the expression, by the time it realizes that there is a problem, it has done some work toward the correct literal parse. In the latter case access

to the idiomatic expression is dependent upon activation of the literal lemmas, and by spreading activation, their literal representation.

Additionally, the processing failure when processing a non-literal structure can be explained as a consequence of the activation process. While both the configuration hypothesis and the hybrid model claim that literal activation is, at least partially, obligatory, neither claim that idiomatic access is obligatory. Thus given sufficient contextual bias, it is possible that the parser simply fails to activate the necessary idiomatic interpretation. If this idiomatic meaning then turns out to be necessary to construct a valid sentential interpretation, then it must be retrieved after the fact, and this contributes to the slow down that we see in our data. Crucially these results are not predicted by either the Direct Access view, in which idiomatic interpretation has priority over literal interpretation, or by a parallel processing accounts such as the Lexical Representation Hypothesis.

One of the advantages of the hybrid model is that it provides a coherent model of how the idiomatic and literal representations of these phrases are related. Thus we can also explain these data by claiming that *both* the literal and idiomatic representations are activated in all cases with the literal representation having priority. After initial activation the parser may actively consider both possible interpretations in parallel. Given sufficient contextual support the parser then selects the idiomatic or literal parse depending upon the overall activation of the system. Given that the literal representation (which is

comprised of the literal lemmas and their associated grammatical and conceptual information) is much more distributed in the system than the idiomatic representation (which comprises the super-lemma and its associated conceptual information), we would predict that while activation of the literal parse may be slower, suppression should also be more difficult with the opposite pattern obtaining for the idiomatic representation.

Thus under a this view, it may be the case that our bias manipulation is sufficient to limit the activation of the idiomatic interpretation sufficiently to cause problems with recovery of the idiomatic representation, but is not sufficient to limit the more distributed activation of the literal interpretation. Likewise, under a more serial, activation and inhibition view, our biases may provide sufficient reason in both cases to inhibit the incongruent interpretation, however given the short time available before disambiguation, the system is capable of suppressing the idiomatic representation more thoroughly than the associated literal representation, hence making it more difficult to recover in the event of incongruence.

Unfortunately this experiment does not allow us to decide between these possibilities, as it was not designed with such a question in mind. We will explore the question of the time-course of literal and idiomatic activation in more detail in the following chapters, and will return to the question of the effects of contextual bias on idiom processing more fully in chapter 4. For now we can only claim that the apparent idiom processing speed advantage obtains during processing, but not recovery. During recovery, this experiment

suggests that access and retrieval of idiomatic meaning is actually slower than the process

of literal interpretation.

Chapter 3: Eye-Tracking: Syntactic & Lexical Effects

In the previous chapter we examined the online processing temporarily ambiguous idiomatic strings and specifically explored the effects of recovering idiomatic or literal meaning when contextual bias led comprehenders toward an incorrect interpretation. The results of that experiment primarily focused on the process of ambiguity resolution during idiom processing. In this chapter we turn our attention to how idiomatic expressions are represented, and how they are related to other entries in the mental lexicon. Critically, as discussed in chapter 1, models of idiom representation differ greatly in their proposed relationships between idioms and their component words, and between idioms and their apparent structural and grammatical properties. In this chapter we present the results of two experiments designed to examine whether comprehenders are sensitive to structural context when interpreting ambiguous idiomatic expressions during on-line processing, and to explore the relationship between idioms such as *kick the bucket* and their component words (i.e. *kick* and *bucket*).

3.1 Motivation

While proposed primarily in the domain of idiom production, the Hybrid Representation Hypothesis of idiom representation (Cutting & Bock, 1997; Sprenger et al, 2006) also makes coherent predictions regarding the process of idiom comprehension. Experiments 3a and 3b, presented in this chapter, focus on testing these comprehension-based predictions. A better understanding of idiom comprehension is important given that current models of idiom representation make vastly different predictions regarding i) the

relationship between an idiomatic expression and its literal components and ii) how idioms are represented structurally. The Hybrid Representation Hypothesis suggests that idioms are associated directly with their literal component lemmas, and are represented as structural phrasal units. In contrast, words-with-spaces approaches suggest no relationship between the idiom and its components and claim that idioms are represented as structureless units.

These experiments are also designed to contribute to our understanding of the time-course of idiom processing, an issue that was also addressed in experiment 2 in the previous chapter. Unlike experiment 2, however, the experiments presented in this chapter utilize the visual-world eye-tracking paradigm. The critical difference here is that while self-paced reading can give us information regarding processing difficulty when parsing ambiguous sentences, it does not allow us to directly examine *which interpretation* participants are considering at different points in time. Other work in idiom comprehension has relied upon cross-modal lexical decision (Cacciari & Tabossi, 1988) as a means of examining which interpretations are active at various time-points, however this too is limited in that it can only probe particular time-points of interest. Eye-tracking allows us to gain insights into the real-time consideration of literal and idiomatic interpretations over the time course of sentence comprehension without having to decided *a priori* which time-points to examine.

3.1.1 Syntactic Effects

One of the claims made by the Hybrid Representation Hypothesis (Cutting & Bock, 1997; Sprenger et al, 2006) is that idiomatic expressions are represented as inherently structural elements. Under the super-lemma view, discussed in more detail in Chapter 1, idiomatic expressions are represented as a set of constraints which define possible structural configurations over the set of component lexical items. During idiom comprehension these structural constraints should also be available, and may potentially be used to help the parser decide upon the correct interpretation of the ambiguous string.

As an example, consider the idiom *kick the bucket*. This idiom is generally known to be highly constrained in its degree of syntactic flexibility. For example, it is generally accepted that it cannot passivize without losing its idiomatic meaning. Under the Hybrid Representation Hypothesis, this lack of compatibility with passivation would be encoded into the super-lemma representation of the idiom. Thus we expect that during comprehension, encountering the appropriate lemmas in a passive structure will signal incompatibility with the super-lemma representation and thus rule out the possibility of an idiomatic interpretation.
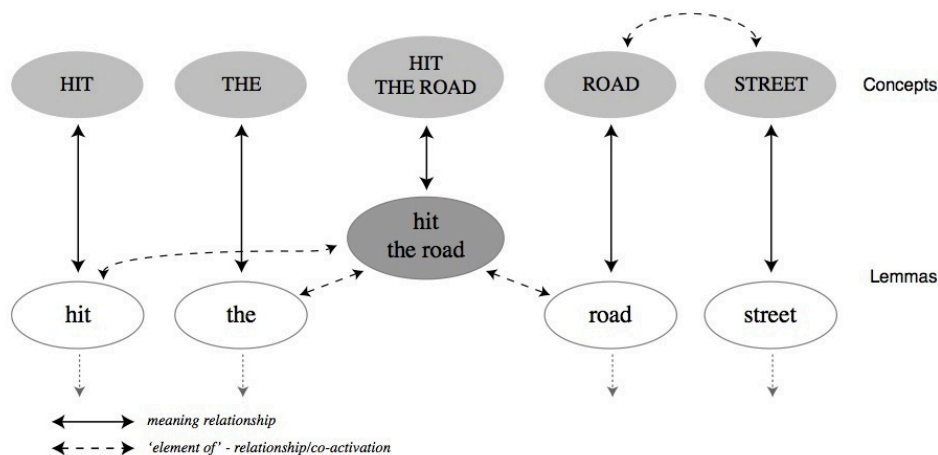
Figure 12: Super-lemma model of idiom representation from Sprenger et al, 2006 as presented in Chapter 1. Reproduced here for convenience.

This raises two relevant questions. The first regards the internal structural representation of the super-lemma itself. On the one hand we might predict that given an expression such as *the bucket was kicked*, interpretation of the idiomatic meaning would simply be ruled out, as it violates the structural constraints contained in the idiomatic super-lemma representation. If Sprenger et. al (2006) are correct in defining this representation as a function over the set of component lemmas (in this case a function in which *kick* and *bucket* combine into a verb phrase), then the passive expression should fail to elicit consideration of the idiomatic meaning by virtue of the fact that no such function exists (as the representation would crucially *not* include a function in which *kick* and *bucket* combine into a passive structure). The parser should then go about its business of computing the literal meaning of the phrase, perhaps only considering the idiomatic interpretation later as a sort of post-processing effect. Another possible interpretation of the model is that the super-lemma representation would become activated immediately

even when faced with the passive form by virtue of blind spreading activation from the component lemmas, and while further analysis would reveal the structure under review to be incompatible with its constraints, some activation may still spread to the conceptual level regardless.

This highlights the second question, which regards the time-course of this process. During comprehension the Hybrid Representation Hypothesis predicts an order of activation which proceeds from the literal lemmas, to the super-lemma representation and finally to the conceptual representation of the idiomatic interpretation. While it is unclear whether such distinctions in activation will be apparent, even in a measure as sensitive as eye-tracking, what is of central interest is whether the idiomatic interpretation is activated at all in syntactically incompatible contexts. If we do not see any activation of the idiomatic interpretation in syntactically incompatible contexts then this will provide evidence that the super-lemma representation may act as a sort of *gate* to the conceptual level. Thus upon hearing *kick* and *bucket* in close proximity, the parser will check to see if a valid phrasal function exists for these lemmas, and upon finding no such function, will simply process the expression literally. If we do find evidence of idiom activation in these syntactically-incompatible contexts, particularly if that activation occurs early in the time-course, then we will have evidence that these super-lemma representations are acting similarly to any other lemma in that they enter active competition with competitor lemmas during the processing ambiguous expressions (e.g. similar to parallel activation of the two senses of *bank*). Critically if these super-lemmas do not act as a sort of gate,

then further development of this model will require independent explanation regarding how the system integrates syntactic information to select the appropriate interpretation.

As previously discussed, idiomatic expressions vary with respect to how syntactically flexible they are, however attempts at classifying idioms based upon flexibility have been mixed. Fraser (1970) proposed six levels of syntactic frozenness, however later work by Swinney & Cutler (1979) failed to find any processing effects based upon these categories. Likewise, Nunberg (1978) proposed that idiomatic expressions were partially compositional, and suggested a correlation between semantic decomposability and syntactic flexibility (see also Nunberg et al, 1994). This approach, dubbed the *idiom decomposition hypothesis*, has been supported in large part by Gibbs and his colleagues. Gibbs & Nayak (1989) provide evidence that individuals have intuitions regarding the degree of semantic decomposability of an idiom and that this intuition is correlated with syntactic flexibility. Gibbs and colleagues have also demonstrated evidence that processing speed and memory for idiomatic expressions change as a function of syntactic flexibility (Gibbs & Nayak, 1989; Gibbs, Nayak & Cutting 1989; Gibbs & Gonzales, 1985).

However this categorization has also had mixed reception in the literature with several studies failing to find any effect of decomposability on idiom processing in both comprehension (Titone & Connie, 1994) and production (Cutting & Bock, 1997). Finally, recent work (Tabossi, Fanari & Wolf, 2008) investigating the *idiom decomposition*

hypothesis suggests that individuals' intuitions regarding the semantic decomposability of idioms are consistent only for some idioms, and that the correlation between decomposability and flexibility is rather tenuous

These findings complicate our goals, as our interest is in the structural representation of idioms. Part of the appeal of the super-lemma representation of idioms is that such representations can specify valid structural configurations, thus giving us a means of accounting for differences in structural flexibility on a case-by-case basis (Sprenger et al, 2006). However there is experimental evidence suggesting that this case-by-case approach may be too powerful, and fail to capture relevant facts about the acquisition and use of novel idiomatic expressions (Tabossi, Wolf & Koterle, 2009). One way to handle this is to posit that while super-lemma model can, in principle, specify structural possibilities on a case-by-case basis, in practice structural flexibility is a property of super-lemmas which is either specified by some other component, or possibly by analogy to other super-lemmas in the lexicon. Ideally one would be able to point to some valid typology of idioms based upon structural flexibility to explore this issue, however as discussed above there does not appear to be a decent candidate typology. Thus rather than attempt to explore the exact structural specification for our idioms, we chose to instead examine the bigger picture. Our syntactic manipulation asks *whether* interpretation of idioms is constrained by syntactic features, as opposed to asking *what* the syntactic representation looks like.

Thus to ensure that our syntactically unavailable condition was indeed structurally incompatible with an idiomatic interpretation of the relevant string we chose to use the presence or absence of a sentential break. This is indeed the biggest syntactic cue we could find and we know of no idioms that are capable of maintaining their meaning when divided across a sentential boundary. This will allow us to examine whether idiomatic interpretation is sensitive to syntactic features.

3.1.2 Lexical Effects

The Hybrid Representation Hypothesis also depicts a well-defined relationship between the literal components of an idiomatic string and its idiomatic representation in which access to the idiomatic representation is mediated via the literal components. However the dynamics of activation and inhibition during comprehension of idioms is less well understood. Critically, during production there is no ambiguity with respect to the intended message. In a situation where the speaker intends to convey the meaning *die suddenly*, if any competition is to occur it should be between the use of the idiom (e.g. *kick the bucket*) and other ways of expressing the same meaning (e.g. *die suddenly*). During comprehension, however, the comprehender has access to the linguistic expression, but the meaning is often ambiguous. If we apply the Hybrid Representation Hypothesis to comprehension, we predict competition between the literal and idiomatic meaning of the string, but it is unclear how this competition is resolved.

One of the interesting findings in Sprenger et al (2006) was that idiomatic representations are associated with semantically and phonologically related material in a way similar to what has been found for individual lemmas (Levelt, Schriefers, Vorberg, Meyer, Pechmann & Havinga, 1991; Dell & O'Seaghdha, 1991; Peterson & Savoy, 1998). For example, they found that sentence completion of incomplete idiomatic strings was facilitated by priming words phonologically or semantically related to the target word in the idiom. Thus one prediction is that strings such as *kick the pail* should also elicit idiomatic interpretation, by virtue of *pail* being semantically associated with *bucket*. Thus, the logic is that the lemma *pail* will partially activate the lemma *bucket* by virtue of the conceptual overlap between the two words. At this point the lexical system will have an activation pattern sufficient for activating the super-lemma *kick the bucket* and hence retrieving the idiomatic interpretation.

To test these predictions we will examine whether individuals consider the idiomatic meaning of expressions such as *kick the pail*, or *spill the vegetables* and compare activation in these expressions to ambiguous idioms such as *kick the bucket*, and *spill the beans*. We will refer to the former stimuli as lexically-unavailable, as expressions such as *kick the pail* are unambiguously literal, and to the latter as lexically-available. If access to the idiomatic representation is mediated by its literal components we expect activation of both the literal and idiomatic interpretations during comprehension of both the lexically-available (e.g. *kick the bucket*) and their lexically-unavailable (e.g. *kick the pail*) items. Furthermore, since we predict a priority of literal processing, we expect that our lexical

manipulation will primarily drive consideration of the idiomatic meaning, with greater consideration of idiomatic meaning in the Lexically Available condition than in the Lexically Unavailable condition. However we do not necessarily expect a similar modulation with respect to literal consideration (which is in this view obligatory), thus we expect Lexically Availability to have a weaker effect on literal consideration.

3.1.3 Task Demands

In addition to manipulations of the syntactic and lexical availability we also manipulated the kind of task that participants were asked to do. As I will explain in the remainder of this section, this was done to deepen our understanding of potential effects of lexical priming. Experiments in the visual-world paradigm differ greatly in the types of tasks involved with many studies requiring participants to manipulate or fixate a given object on the screen (Allopenna, Magnuson & Tanenhaus, 1998; Meyer 2005; Tanenhaus et al, 1995; Yee & Sedivy, 2006; Salverda, Dahan & McQueen, 2003, and others), while others simply require participants to look at the screen, possibly providing some feedback regarding whether or not the screen matches the auditory stimuli (Altmann, 2004; Altmann & Kamide, 2009; Kaiser, Runner, Sussman & Tanenhaus, 2009) and the results of these different tasks seem to converge. We chose to use the latter, and so participants were instructed only to maintain attention on the screen for the duration of the trial.

Our main concern, however was with weakness of semantic activation via orthographic stimuli. We decided to use orthographic stimuli rather than picture stimuli, as picture

stimuli depicting the idiomatic meaning of our expressions would be more complex and abstract than images relating to their literal interpretation, hence introducing a serious confound. We will discuss the motivation for using text-based stimuli rather than the typical picture stimuli in more detail in our procedure. However, it is known that looking behavior to text stimuli differs from looking behavior to visual stimuli (Meyer, 2005; Huettig & McQueen, 2007). Generally, looks to semantically related targets are either delayed or reduced when using text rather than images in the visual-world paradigm, mirroring views of reading comprehension which claim that semantic activation during reading is mediated via phonological activation (van Orden, Johnston & Hale; 1988).

Thus in order to test for potential effects of depth of semantic activation, we instructed half of our participants to read the four words on the screen silently to themselves before the onset of the audio stimuli. The other half were instructed to read the words aloud. The logic here is that if semantic activation via orthographic form is mediated by phonological access then we may see differences in these two tasks. Our expectations were that reading aloud would imply more reliable phonological, and hence more reliable semantic activation than reading silently, as there was no way to ensure that participants reading silently were fully processing the semantics of the words on the screen, however it may also be the case that deeper semantic activation of the words on the screen may lead to individuals more fully activating the conceptual representation of the idiom (e.g. they just read *death* and are now hearing *kick the bucket*) which may influence their behavior.

We will present the read-silently results in experiment 3a. These results allow us examine individuals processing of idiomatic expressions in the aforementioned syntactic and lexical conditions under task demands in which we expect overall weaker effects of semantic activation, but also predict weaker influence from the visual target words on participants' behavior. In experiment 3b we present the results of the read-aloud task in which we expect stronger semantic activation, but potentially more influence from the visual display on participants' looking behavior.

3.2 Experiment 3: Syntactic & Lexical Effects on Idiom Processing

In this experiment we examined the effects of syntactic and lexical availability on the processing and interpretation of idiomatic expressions. In short, we examined whether individuals consider the idiomatic and/or literal interpretations of a string in incompatible syntactic contexts and whether semantically associated strings (e.g. *kick the pail*) show similar looking behavior to well-formed idioms (e.g. *kick the bucket*). Participants' eye-movements were recorded as they listened to potentially idiomatic sentences and looked at four words shown on the computer screen. As mentioned earlier, we also modulated the task demands of the experiment with half of our participants reading the four words silently before the onset of the audio sentence (Experiment 3a) and half reading the four words aloud (Experiment 3b).

3.2.1 Method

3.2.1.1 Participants

32 undergraduate students at the University of Southern California participated in this study. All participants were native speakers of American English. None of the participants participated in any of the previous experiments.

3.2.1.2 Materials

The sentences that participants heard in this experiment contained strings that were ambiguous between a literal or an idiomatic interpretation (e.g. *kick the bucket*) as well as strings that could only be interpreted literally (e.g. *kick the pail*).

3.2.1.2.1 Selecting the Idioms

Given the goals of this experiment it was essential that the idiomatic expressions used in our stimuli were (i) known to our participants and (ii) were maximally structurally similar to each other such that our eye-tracking time analyses would be valid. To handle the latter point, a list of 21 potential idiomatic candidates were constructed which all had the form *verb x noun* where x was either the definite or indefinite article (*kick the bucket, cut the mustard*) or a possessive pronoun (*hold your horses, find her feet*). To ensure that all of the idioms used in the eye-tracking study were familiar to the participants, an off-line norming study was conducted in which participants (n=16, none of these individuals participated in the eye-tracking study) did four norming tasks: For each idiom they (i) indicated whether or not they were familiar with the expression, (ii) provided a paraphrase of the idiomatic meaning, (iii) provided the first three words that come to mind when thinking of the idiomatic meaning and (iv) rated on a five-point scale how

common they thought the given idiom was (1 = very rare, 5 = very widely known). Parts (i), (ii) and (iv) allowed us to see which idioms participants were familiar with and judged to be widely known. Part (iii) is important as it provided us with a set of semantic associates of the idiomatic meaning to utilize as the visual stimuli.

From the initial pool of 21, all of the idioms selected were known to a majority of participants. We selected the 12 most familiar idioms that allowed the construction of maximally similar stimuli (e.g. *shoot the breeze* and *pass the buck* were both highly familiar, but were excluded due to lack of a pragmatically reasonable literal interpretation and difficulty in finding semantic associates for the noun *buck* respectively). The idioms selected had a mean familiarity of 97% and a mean commonality intuition rating of 3.8 (on the 5 point scale). A full list of these idioms is given in Appendix II.

3.2.1.2.2 Audio Stimuli

To examine the relationship between idioms and their component lemmas we used a manipulation similar to those used previously in the production studies (see Cutting & Bock, 1997; Sprenger et. al, 2006). For each of 12 chosen idioms, I changed the final noun to a semantically-related word (e.g. *kick the pail*, *smell a mouse*). I will refer to the 12 original, unaltered idioms as being in the **Lexically Available** condition because they (e.g. *kick the bucket*) consist of lexical items that allow either a literal or idiomatic interpretation. I will refer to the altered versions (e.g. *kick the pail*) as being in the **Lexically Unavailable** condition, because the lexical items that comprise these altered

variants only have a literal interpretation and cannot be interpreted idiomatically. A full list of these items is given in Appendix II.

In addition to manipulating the lexical availability of an idiomatic interpretation, I also manipulated the syntactic context in which the potentially idiomatic strings occurred. In the **Syntactically Available** condition the unaltered lexically available or altered lexically unavailable string was inserted into a simple sentence containing only a proper name and a time phrase. In the **Syntactically Unavailable** condition the lexically available or lexically unavailable string was divided between two sentences such that the *verb* of the string occurred as the final word of the first sentence and the *x noun* component as the initial string of the following sentence. Crossing our two conditions yielded four versions of each target item as shown in Table 5. A full list is provided in Appendix III.

	Syntactically Available	Syntactically Unavailable
Lexically Available	John **kicked the bucket** last Thursday	It was surprising to see someone as skilled as John completely miss when he **kicked. The bucket** full of orange slices was destroyed when he accidentally missed the ball.
Lexically Unavailable	John **kicked the pail** last Thursday	It was surprising to see someone as skilled as John completely miss when he **kicked. The bucket** full of orange slices was destroyed when he accidentally missed the ball.

Table 5: Example Stimuli for experiment 3.

These sentences were then recorded by a single male native speaker of American English using a Logitech g35 USB headset. Recordings were performed in one sitting and each sentence was recorded as a whole (without splicing) to ensure maximal naturalness.

In addition to these 48 target sentences, 60 filler sentences were also recorded. To better mask the target stimuli half of all fillers were short simple sentences (similar to the **Syntactically Available** trials) and half were longer multi-sentence stories (similar to the **Syntactically Unavailable** trials).

3.2.1.2.3 Visual Stimuli

Visual stimuli consisted of a set of four words presented on the screen. For target items those stimuli consisted of an **Idiom Associate, Literal Associate** and two **Distractors**. **Idiom Associates** were selected based upon the results of the aforementioned off-line norming study in which participants listed the first three words that came to mind when considering the idiomatic interpretation. These words were aggregated into a list, and the most frequently reported word was used as the idiom associate. **Literal Associates** were associated with the literal meaning of the expression and was associated with either the noun (e.g. *bucket/pail*) or the verb when a noun associate was difficult to find (e.g. *kick*). Associates were determined using the University of South Florida free association norms (Nelson, McEvoy & Schreiber, 1998).[2] Overall 7 of the target items had noun associates,

---

[2] Association was defined as having either a forward or backward association in the USF database with the exception of two cases. The literal associate for *hold one's horses* was *reins* which was decided to be a good associate for both *horse* and *pony* but does not occur in the USF database (the other alternatives e.g. *hay, straw, barn* etc. were confounding with other items). Additionally, *skinny* was chosen for the literal associate of *tighten one's belt* as it captures the overall literal interpretation of this phrase.

and 5 had verb associates. **Distractors** were chosen such that they were not associated with any of the words in the expression or with any of the other words in the display. This was also assessed using the USF free association norms, and operationally defining lack of association as a lack of either forward or reverse association between any of the words in the expression (e.g. *kick, bucket*) or any of the other words in the display. To better mask our target stimuli, our fillers also contained a word semantically related to one of the words in the sentence or a word which directly matched a given word in the sentence.

Visual stimuli were presented on the screen with one word in each screen corner. Position of the associates and distractors was balanced both within the target items and overall (pooling targets and fillers). A sample display for *kick the bucket* is presented in Figure 13.

Idiom Associate	**Literal Associate**

<table>
<tr><td>Death</td><td>Foot</td></tr>
<tr><td>Triangle</td><td>Animal</td></tr>
</table>
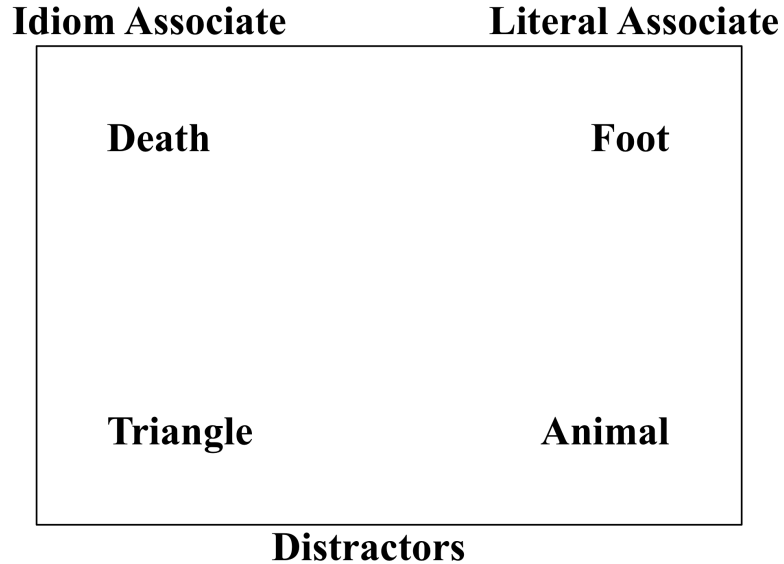
Distractors

Figure 13: Sample display with labels pointing out the Idiom Associate, Literal associate and Distractors.

(These labels were not present on the actual display seen by participants)

We chose to use a text-based variant of the visual-world paradigm because we were concerned that using images would add serious confounds to our procedure. Our decisions to use words rather than images in our task came from two main considerations. First, it proved to be incredibly difficult to provide image associates for the literal and idiomatic conditions that did not also add confounds to our results. In general the meaning of our idiomatic expressions is complex, abstract and has low imageability (e.g. *kick the bucket* meaning *to die suddenly*, *hold one's horses* meaning *to restrain oneself from performing a hasty action*, *smell a rat* meaning *to know instinctively that something is amiss*). Associates to our literal targets, however, are generally concrete and simple. Thus, if we had used pictures for our idiom and literal associates, they would have differed significantly in both their visual and conceptual complexity, and in many cases it would have been very difficult to provide a clear picture-based representation of the

idiomatic meaning (e.g. *find her feet*, *hold one's horses*). Using words allowed us much better control over our visual stimuli, and allowed us to create both idiom and literal associates for all targets.

In addition, as mentioned in the motivation for our task manipulation, orthographic stimuli are known to provide weaker semantic activation than image stimuli. Since our results are based upon the strength of semantic activation, the use of orthographic visual targets will only serve to act against our results, thus the choice is maximally conservative.

3.2.1.3 Procedure

Participants' eye-movements were tracked with an Eyelink II system (SR Research) as they listened to sentences and looked at the words on the screen. At the start of each trial, the four visual stimuli were displayed before the critical sentence(s) were presented over headphones. In the **Read Silently** condition, participants were instructed to read each of the four words to themselves. After five seconds of preview the audio stimulus started to play. In the **Read Aloud** condition, participants were instructed to read each of the four words aloud, and then press a button to indicate that they had finished. 500ms after the button press the audio stimuli began. From there all participants were instructed simply to maintain their attention on the screen for the duration of the trial. Aside from the initial procedure the two task conditions did not differ in any way.

Stimuli were presented to participants using a modified latin-square design. Each participant saw two versions of each target, but never saw any target in the same condition twice. Order of presentation was pseudo-randomized such that the first occurrence of a given target always occurred in the first half of the list and the second in the second half. Additionally, half of all participants saw lists with reversed order to control for potential order or learning effects. Thus the factors of **Syntactic Availability** and **Lexical Availability** were within-subjects factors, with **Task** as a between-subjects factor.

3.2.2 Experiment 3a: Read Silently Task

In this section we present the predictions, results and discussion for the within-subjects manipulations of **Syntactic Availability** and **Lexical Availability** for the **Read Silently** experiment (Experiment 3a). We will then present the predictions, results and discussion for the related **Read Aloud** experiment (Experiment 3b). Within each experiment, I will first present the results for the **Syntactically Unavailable** conditions in which the idiomatic interpretation was syntactically unavailable due to the presence of the sentential boundary, and then turn to the **Syntactically Available** conditions, where both the idiomatic and literal interpretations are, at least in principle, possible. After reviewing these two experiments independently, I will then analyze the two experiments together treating task (Read Silently vs. Read Aloud) as a between-subjects factor in our analysis and provide the general discussion for this chapter.

3.2.2.1 Experiment 3a: Predictions

We are primarily interested in qualitative differences in looking behavior in our different conditions, and on the timing of changes in looking behavior over the course of a trial. Our primary focus will be on the difference between the proportion of looks to the literal target and looks to the idiomatic target, and whether our lexical, syntactic and task manipulations have any significant effects on looking behavior.

With respect to our manipulation of **Syntactic Availability**, we predict that participants will *not* consider the idiomatic interpretation of the string in the syntactically-unavailable condition. As discussed earlier in this chapter, the Hybrid Representation Hypothesis claims that idiomatic expressions are represented as phrasal super-lemmas, which may act as a kind of gate to the conceptual representation of the idiomatic expression. Thus while we do expect activation of the individual literal lemmas (e.g. *kick* and *bucket*/*pail*) we predict that the incompatibility of the syntactic context around these lemmas (i.e. the presence of a sentential break between the verb and noun phrase) will rule out consideration of the idiomatic meaning. In our eye-tracking results this would manifest as significantly more looks to the literal target than the idiomatic target. Further, as we predict no consideration of idiomatic meaning at all, we also expect that looks to the idiomatic target will not be significantly different from looks to the distractors. We predict these effects to obtain regardless of whether the given string is lexically available (e.g. *kick the bucket*) or lexically unavailable (e.g. *kick the pail*).

For syntactically available trials, however, we expect to see a different pattern of results. Namely, we predict that both lexically available and lexically unavailable trials will result in consideration of *both* the idiomatic and literal interpretations. In our eye-tracking results we expect this to manifest as competition between the literal and idiomatic interpretations at some point after the onset of the critical noun (i.e. *bucket* or *pail*). Furthermore, we predict that an actual lexically available idiom (e.g. *kick the bucket*) should be capable of activation the associated idiomatic representation earlier and/or more strongly than a lexically unavailable but semantically related string (e.g. *kick the pail*). This should result in looks to the idiomatic target rising earlier and/or being significantly higher in the lexically available trials as compared to the lexically unavailable trials.

3.2.2.2 Experiment 3a: Results

For each condition we computed average fixation proportions (by subjects and items) over a time interval extending from 200ms after the onset of the critical noun (e.g. *Bucket* or *Pail*) to 1000ms post onset. As we were interested in changes in behavior over this time window, the full time window was also partitioned into eight 100ms windows and further analyses were performed on each of those windows.

To get a sense of how much listeners were considering the literal vs. the idiomatic interpretation of each string, we computed Literal Advantage scores (by both subject and item) for the region and each of the eight smaller windows. These scores are computed by taking the difference between the average looks to the literal target and the average looks

to the idiomatic target. A positive value represents more looks to the literal target; a negative value represents more looks to the idiomatic target and a value close to 0 represents equal looks to both the literal and idiomatic targets.

In what follows we will first present explore the behavior in each of our four (Syntactic Availability X Lexical Availability) conditions. Results examine the difference scores over time, in particular examining whether the Literal Advantage scores deviate significantly from 0.

3.2.2.2.1 Syntactically Unavailable Condition

3.2.2.2.1.1 Graph & Basic Discussion of Trends
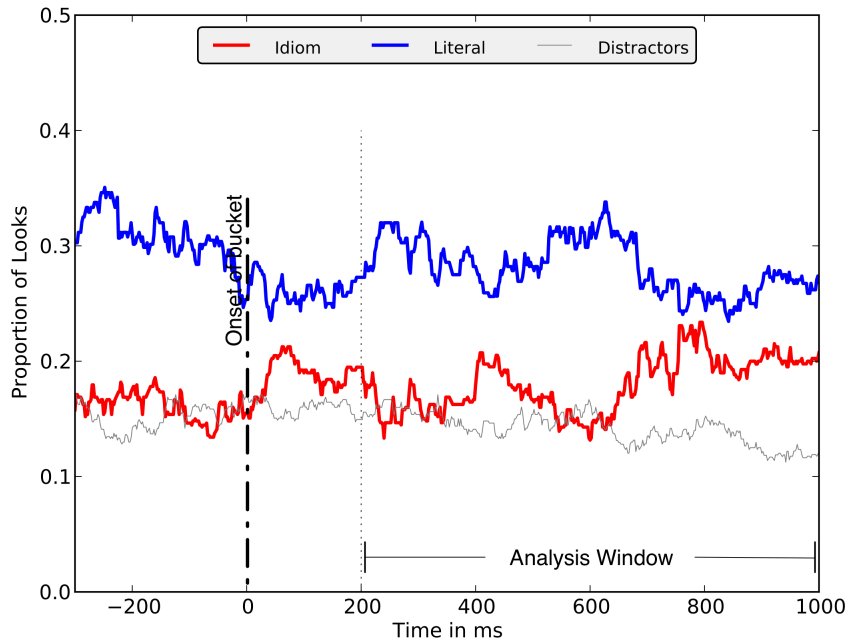
Figure 14: Looks to the idiom associate word (in blue/dark) vs. looks to the literal associate word (in red/light) as well as looks to the averaged distractor words (thin line in light grey). This figure shows looks in the Syntactically Unavailable - Lexically Available condition (e.g. kick. The bucket), for the read-silently task.

Syntactically Unavailable - Lexically Unavailable
(e.g. *kick. The pail*) [Read-Silently]
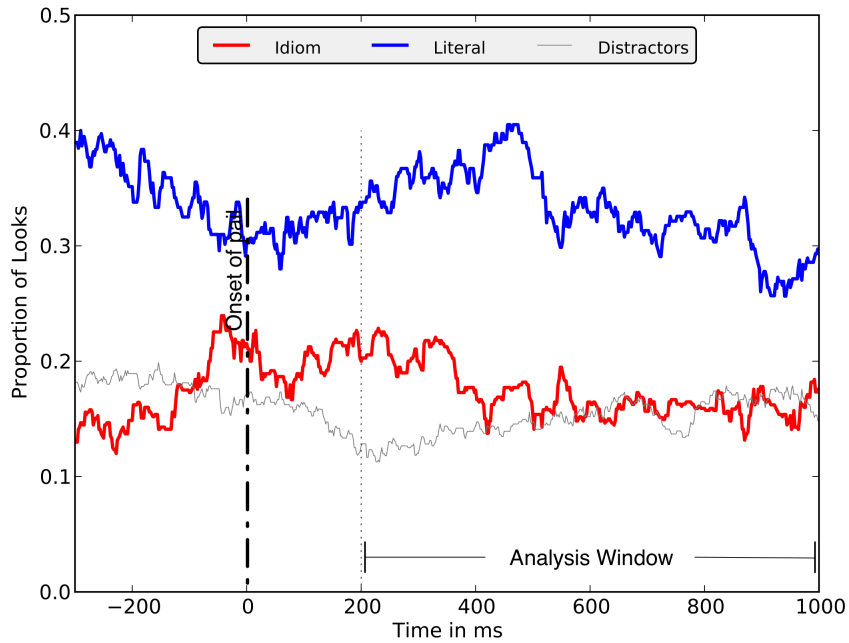
Figure 15: Looks to the idiom associate word (in blue/dark) vs. looks to the literal associate word (in red/light) as well as looks to the averaged distractor words (thin line in light grey). This figure shows looks in the Syntactically Unavailable - Lexically Unavailable condition (e.g. kick. The pail), for the read-silently task.

Figures 14 and 15 display the raw proportion of looks for the syntactically unavailable, read silently trials in the Lexically Available and Lexically Unavailable conditions respectively. In both the lexically available and lexically unavailable condition, we see more fixations to the literal associate word (e.g. *foot*) than to the idiom associate word (e.g. *death*) or the distractor words. (Each display contained two distractor words, proportion of fixations to distractor words is averaged for clarity). This literal advantage is already present before the onset of the critical noun and appears to persist for the duration of the critical window from 200ms until 1000ms post noun onset. Generally, this

100

pattern suggests that participants are largely focused upon the literal interpretation, with the proportion of fixations to the literal target being slightly higher in lexically unavailable (e.g. *kick. The pail*) trials.

We also observe that the proportion of fixations to the idiomatic target generally remain close to the distractor baseline, during our critical time-window, with a possible exception arising in the later 700ms-1000ms interval for lexically available trials (e.g. *kick. The bucket*). Taken together this suggests that participants are strongly considering the literal interpretation and are also not considering their potential literal interpretation (excluding the aforementioned exception).

3.2.2.2.1.2 Overall Statistical Analyses

To assess these patterns statistically, we conducted a series of two-tailed t-tests over eight 100ms time regions starting at 200ms post noun onset and extending to 1000ms (the full analyses are reported in Table 6.). Statistical analyses largely reveal what we observed visually. When the idiomatic meaning is lexically unavailable (e.g. *kick. The pail*), there is a significant literal advantage over the full region, with the literal advantage starting in the 300-400ms (significant by subjects only) and persisting through much of the trial. In contrast, for lexically available trials (e.g. *kick. The bucket*) we do not see a significant Literal Advantage in the difference scores, though we a marginal literal advantage in several early time windows up until 700ms, after which where we see no advantages for either the literal or idiomatic interpretation for the remainder of the trial.

Results of two-way t-test analyses on Literal Advantage by Time Window (Read-Silently)

	df	Syntactically Unavailable - Lexically Available (e.g. kick. The bucket)			Syntactically Unavailable - Lexically Unavailable (e.g. kick. The pail)		
		CI	t	p	CI	t	p
Region							
by Subject	15	(-0.039 , 0.238)	1.537	0.145	(0.032 , 0.291)	2.657	0.018 *
by Items	11	(-0.05 , 0.243)	1.448	0.175	(0.054 , 0.269)	3.317	0.007 **
200ms-300ms							
by Subject	15	(-0.011 , 0.305)	1.978	0.067 .	(-0.075 , 0.348)	1.373	0.19
by Items	11	(-0.03 , 0.31)	1.808	0.098 .	(-0.072 , 0.321)	1.395	0.19
300ms-400ms							
by Subject	15	(-0.042 , 0.285)	1.58	0.135	(-0.047 , 0.368)	1.651	0.12
by Items	11	(-0.022 , 0.298)	1.894	0.085 .	(0 , 0.308)	2.204	0.05 .
400ms-500ms							
by Subject	15	(-0.067 , 0.256)	1.248	0.231	(0.031 , 0.411)	2.481	0.025 *
by Items	11	(-0.08 , 0.256)	1.156	0.272	(0.064 , 0.407)	3.017	0.012 *
500ms-600ms							
by Subject	15	(-0.038 , 0.304)	1.662	0.117	(-0.017 , 0.356)	1.936	0.072 .
by Items	11	(-0.051 , 0.354)	1.647	0.128	(0.054 , 0.296)	3.175	0.009 **
600ms-700ms							
by Subject	15	(-0.029 , 0.32)	1.778	0.096 .	(0.009 , 0.343)	2.252	0.04 *
by Items	11	(-0.099 , 0.321)	1.159	0.271	(0.056 , 0.281)	3.307	0.007 **
700ms-800ms							
by Subject	15	(-0.132 , 0.221)	0.54	0.597	(0.016 , 0.3)	2.376	0.031 *
by Items	11	(-0.18 , 0.26)	0.4	0.697	(0.027 , 0.27)	2.698	0.021 *
800ms-900ms							
by Subject	15	(-0.092 , 0.199)	0.782	0.447	(0.007 , 0.287)	2.244	0.04 *
by Items	11	(-0.136 , 0.221)	0.521	0.613	(0.028 , 0.31)	2.643	0.023 *
900ms-1000ms							
by Subject	15	(-0.097 , 0.216)	0.809	0.431	(-0.004 , 0.248)	2.067	0.056 .
by Items	11	(-0.118 , 0.239)	0.747	0.471	(-0.044 , 0.281)	1.605	0.137

Table 6: Results of two-way t-tests comparing the literal advantage scores to 0 for our region of interest from 200ms to 1000ms post noun onset and for each of eight 100ms windows in that region.

3.2.2.2.1.3 Interim Discussion

We find that participants do strongly consider the literal interpretation in the syntactically-unavailable condition. This suggests that syntactic constraints guide online processing of idioms, but our results are not as strong as we had predicted. Recall that our prediction was that placing an idiomatic expression into an incompatible syntactic frame would cause the lexical access system to completely rule out the idiomatic possibility. This does appear to be what is happening in the lexically unavailable condition. In the lexically available condition, however the effect seems to be much less pronounced and

we also see what appears to be late competition between the idiomatic and literal interpretations. While these patterns of looks do support the notion that syntactic compatibility plays a role in deciding between an idiomatic and literal interpretation, the results of the lexically unavailable condition suggests that even in the presence of a sentential boundary there may be some delayed activation of potential idiomatic interpretation.

3.2.2.2.2 Syntactically Available Conditions

3.2.2.2.2.1 Graph & Basic Discussion of Trends

Figure 16: Looks to the idiom associate word (in blue/dark) vs. looks to the literal associate word (in red/light) as well as looks to the averaged distractor words (thin line in light grey). This figure shows looks in the Syntactically Available - Lexically Available condition (e.g. kick the bucket), for the read-silently task.

Syntactically Available - Lexically Unavailable
(e.g. *kick the pail*) [Read-Silently]
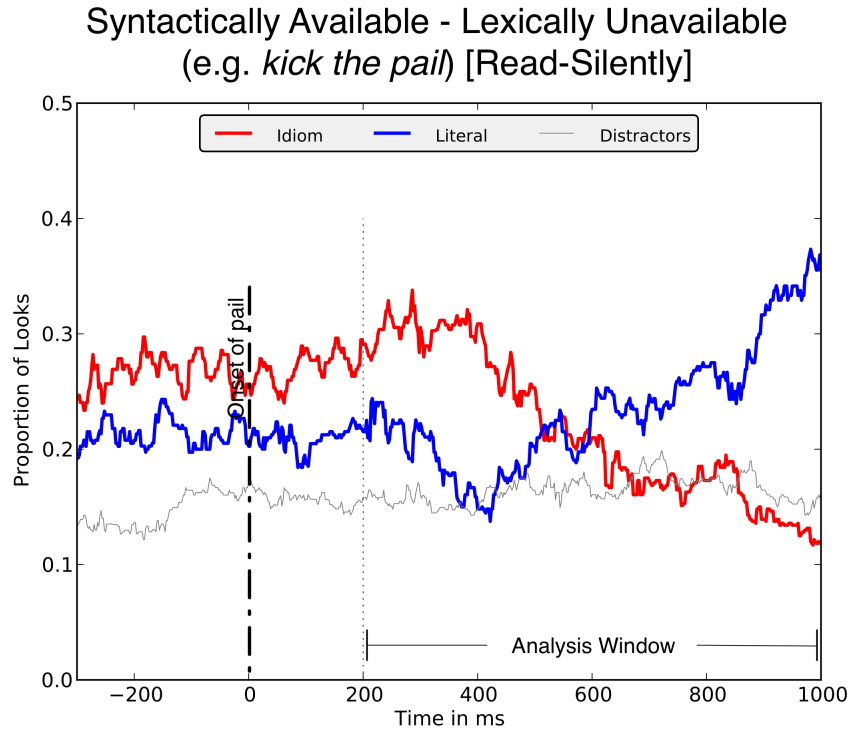
Figure 17: Looks to the idiom associate word (in blue/dark) vs. looks to the literal associate word (in red/light) as well as looks to the averaged distractor words (thin line in light grey). This figure shows looks in the Syntactically Available - Lexically Unavailable condition (e.g. kick the pail), for the read-silently task.

Now let us consider what happens when both the idiomatic and literal interpretations are licensed by the syntactic context. Figures 16 and 17 display the raw proportion of looks for the syntactically unavailable, read silently trials in the Lexically Available and Lexically Unavailable conditions respectively. In the lexically available condition (e.g. *kick the bucket*) we see that the proportion of fixations to the literal associate word and idiom associate word are very similar during the critical time window from 200ms - 1000ms post noun onset. Visual examination suggests that looks to the associate words begin to deviate from the distractors around 600ms post noun onset, and remain higher

than the averaged distractors for the rest of the time-window. This pattern suggests competition between the idiomatic and literal interpretations.

For the lexically unavailable condition (e.g. *kick the pail*) we see more fixations to the idiomatic target than the literal target early in the time-window shifting to more fixations to the literal target around 600ms post noun onset. The pattern here suggests an early advantage for the idiomatic interpretation followed by a brief period of competition in the 500ms - 600ms window, which is then followed by a growing advantage for the literal interpretation for the remainder of the trial.

Generally these results suggest activation of both literal and idiomatic interpretations in both cases, however in the lexically available (e.g. *kick the bucket*) condition this manifests as persistent long-lasting competition between the two interpretations, while in lexically unavailable (e.g. *kick the pail*) trials it manifests as a shift from the idiomatic to literal interpretation.

3.2.2.2.2 Overall Statistical Analyses

We again use a series of two-tailed t-tests over eight 100ms time regions starting at 200ms post noun onset and extending to 1000ms (full results are presented in Table 7). The persistent long-lasting competition between the literal and idiomatic interpretations that we observed in the lexically available condition is as verified by our statistical analyses revealing that literal advantage scores fail to significantly differ from 0 for the duration of the trial. For the lexically unavailable condition, our analyses reveal no

significant Literal or Idiomatic Advantage over the full region, which is expected given the change in behavior visible around 600ms. Our analyses by time-window generally reveal the same sort of competition apparent in the lexically available trials. However closer inspection shows a marginal advantage for the idiomatic interpretation in the 300ms - 400ms window (by subjects only), and a late advantage for the Literal interpretation in the final time window from 900ms - 1000ms.

Results of two-way t-test analyses on Literal Advantage by Time Window (Read-Silently)

	df	Syntactically Available - Lexically Available (e.g. kick the bucket) CI	t	p	Syntactically Available - Lexically Unavailable (e.g. kick the pail) CI	t	p
Region							
by Subject	15	(-0.083 , 0.131)	0.473	0.643	(-0.112 , 0.149)	0.299	0.769
by Items	11	(-0.101 , 0.145)	0.391	0.703	(-0.104 , 0.139)	0.321	0.754
200ms-300ms							
by Subject	15	(-0.145 , 0.053)	-0.983	0.341	(-0.268 , 0.096)	-1.008	0.329
by Items	11	(-0.227 , 0.156)	-0.407	0.692	(-0.254 , 0.138)	-0.652	0.528
300ms-400ms							
by Subject	15	(-0.098 , 0.168)	0.565	0.58	(-0.282 , 0.024)	-1.796	0.093 .
by Items	11	(-0.143 , 0.206)	0.402	0.696	(-0.324 , 0.069)	-1.427	0.181
400ms-500ms							
by Subject	15	(-0.123 , 0.217)	0.589	0.565	(-0.264 , 0.094)	-1.015	0.326
by Items	11	(-0.102 , 0.2)	0.716	0.489	(-0.295 , 0.072)	-1.337	0.208
500ms-600ms							
by Subject	15	(-0.099 , 0.187)	0.655	0.522	(-0.206 , 0.18)	-0.141	0.89
by Items	11	(-0.116 , 0.182)	0.487	0.636	(-0.159 , 0.131)	-0.216	0.833
600ms-700ms							
by Subject	15	(-0.093 , 0.196)	0.758	0.46	(-0.095 , 0.209)	0.803	0.435
by Items	11	(-0.16 , 0.246)	0.465	0.651	(-0.074 , 0.196)	0.99	0.344
700ms-800ms							
by Subject	15	(-0.131 , 0.152)	0.158	0.877	(-0.04 , 0.224)	1.482	0.159
by Items	11	(-0.22 , 0.288)	0.292	0.776	(-0.059 , 0.233)	1.308	0.217
800ms-900ms							
by Subject	15	(-0.152 , 0.155)	0.022	0.983	(-0.034 , 0.249)	1.617	0.127
by Items	11	(-0.254 , 0.264)	0.042	0.967	(-0.028 , 0.231)	1.721	0.113
900ms-1000ms							
by Subject	15	(-0.135 , 0.228)	0.545	0.594	(0.038 , 0.369)	2.614	0.02 *
by Items	11	(-0.251 , 0.281)	0.125	0.903	(0.043 , 0.364)	2.788	0.018 *

Table 7: Results of two-way t-tests comparing the literal advantage scores to 0 for our region of interest from 200ms to 1000ms post noun onset and for each of eight 100ms windows in that region.

3.2.2.2.2.3 Interim Discussion

The competition observed in the lexically available trials strongly supports our prediction that individuals access both the literal and idiomatic interpretations of potentially idiomatic strings, at least in unbiased contexts. The general statistical patterns for the lexically unavailable trials also support the view that literal and idiomatic interpretations are accessed during the processing of semantically associated strings such as *kick the pail*, however the behavioral trends suggest a more complicated process when parsing these strings. Particularly we see signs of early attention to the incorrect, idiomatic interpretation as shown by the high proportion of looks to the idiomatic target early in the time-course and full consideration of the correct literal interpretation only later.

3.2.3 Experiment 3b: Read Aloud Task

We now turn to Experiment 3b, in which we examine the results of the **Read Aloud** task. Results and analyses are presented exactly as they were in Experiment 3a. Recall that this experiment is identical to Experiment 3a in every way except that instead of reading the four visual stimuli silently to themselves, participants were instructed to read each of the four words aloud before pressing a button to start the audio stimuli.

3.2.3.1 Experiment 3b: Predictions

In general, we predict the same overall pattern of results that obtained in the previous experiment. Thus for syntactically unavailable trials we expect to see strong consideration of the literal interpretation and little consideration of the idiomatic interpretation regardless of whether the target string is lexically available (e.g. *kick. The*

bucket) or lexically unavailable (e.g. *kick. The pail*). Similarly, for syntactically available trials we expect consideration of both the idiomatic and literal interpretations for both lexically available and unavailable conditions.

With respect to the difference in task, we discussed previously that we expect the act of reading the words aloud to have two potential effects that may modulate these results as compared to the task of reading silently in Experiment 3a. First, we expect that the act of reading the words aloud will result in deeper semantic activation of the visual stimuli, which may serve to enhance participants' behavior during the audio presentation. Thus if we obtain comparable effects to those obtained in the previous experiment, we may see those effects starting earlier in the time-course and/or see effects that are greater in magnitude. In terms of our eye-tracking data this would translate into our observed proportion of fixation curves either shifting leftward, or exhibiting 'stronger' effects, which may manifest as either greater overall magnitude of proportions or decreased variance in our proportion of looks curves.

3.2.3.2 Experiment 3b: Results

These results again rely upon computing literal advantage scores over eight 100ms time windows beginning 200ms post noun onset and extending to 1000ms. Analyses are performed and reported for this experiment in the same way and same order that they were presented in Experiment 3a.

3.2.3.2.1 Syntactically Unavailable Trials

3.2.3.2.1.1 Graph & Basic Discussion of Trends
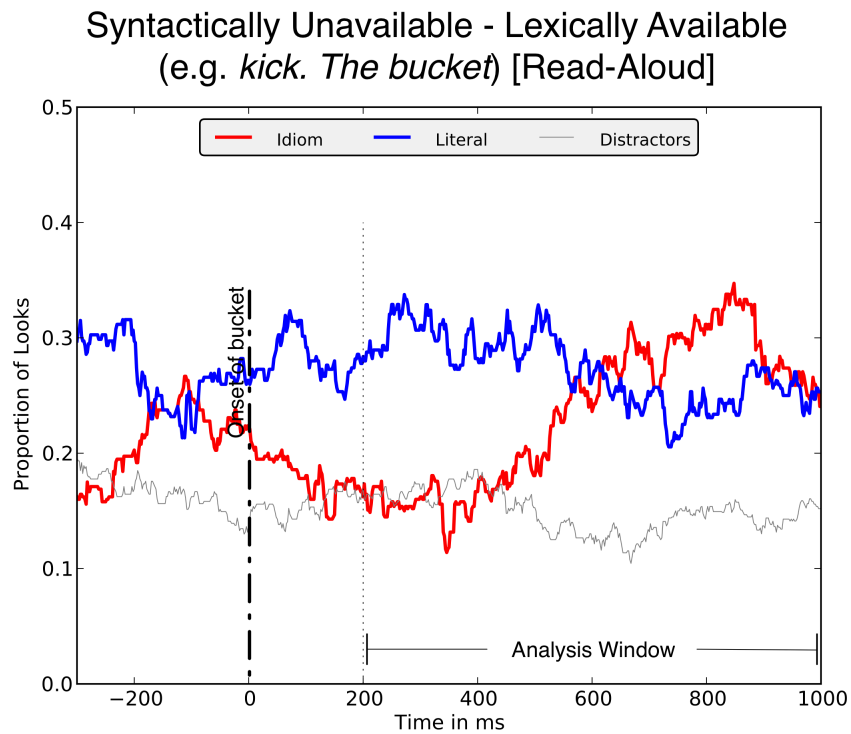
Figure 18: Looks to the idiom associate word (in blue/dark) vs. looks to the literal associate word (in red/light) as well as looks to the averaged distractor words (thin line in light grey). This figure shows looks in the Syntactically Unavailable - Lexically Available condition (e.g. kick. The bucket), for the read-aloud task.

Syntactically Unavailable - Lexically Unavailable
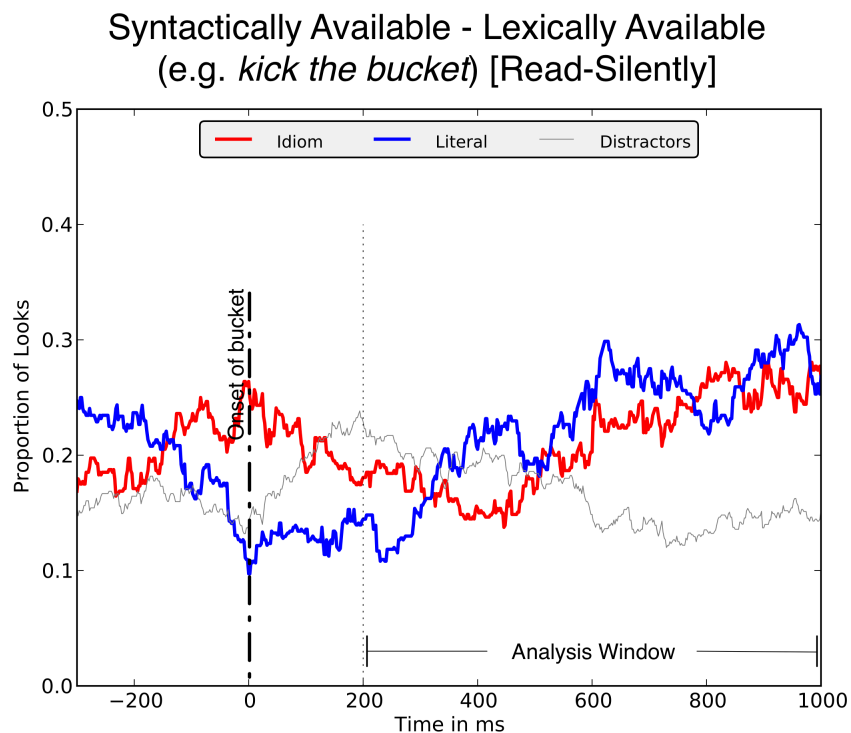### (e.g. *kick. The pail*) [Read-Aloud]
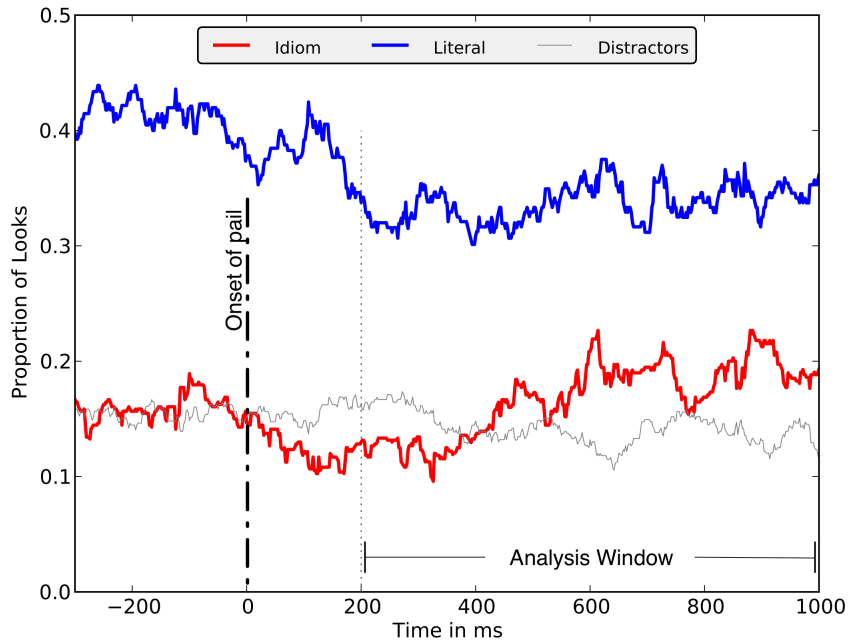
Figure 19: Looks to the idiom associate word (in blue/dark) vs. looks to the literal associate word (in red/light) as well as looks to the averaged distractor words (thin line in light grey). This figure shows looks in the Syntactically Unavailable - Lexically Unavailable condition (e.g. kick. The pail), for the read-aloud task.

First we examine the proportion of fixation curves for the Syntactically Unavailable condition. Figures 18 and 19 display the raw proportion of looks for the syntactically unavailable, read aloud trials in the Lexically Available and Lexically Unavailable conditions respectively. In both the Lexically Available and Lexically Unavailable conditions, we see more fixations to the literal associate word (e.g. *foot*) than the idiom associate word (e.g. *death*) or the averaged distractor words. This literal advantage is most clear in the Lexically Unavailable (e.g. *kick. The pail*) condition where we see a strong preference for the literal associate starting even before the onset of the noun itself

(i.e. after hearing *kick* during the pause between the two sentences). Additionally, we again observe that the proportion of fixations to the idiom associate remains close to the distractor baseline for the Lexically Unavailable Condition. This suggests that participants are focusing primarily on the literal interpretation of this string, and are not considering the idiomatic interpretation.

For the Lexically Available (e.g. *kick. The bucket*) condition we also see a general preference for the literal interpretation for the first 500ms, as the proportion of looks to the literal associate appear to be greater than the proportion of looks to the idiom associate (which remains close to the distractor baseline) over this early region. We also see what appears to be competition between the literal and idiom associates starting after 500ms and continuing on for the remainder of our time window.

3.2.3.2.1.2 Overall Statistical Analyses
We conducted a series of two-tailed t-tests over eight 100ms time regions starting at 200ms post noun onset and extending to 1000ms to evaluate these observations statistically (full results are presented in Table 8). For lexically available trials we see no significant literal advantage over the full region. Analyses by region reveals confirms our observation of an early literal advantage with our statistical results revealing a marginal literal advantage in the 200ms-300ms time window, becoming fully significant in the following time window. This is followed by competition for the remainder of the trial.

For Lexically Unavailable trials our statistical results confirm our observations of the data with a literal advantage over the full 800ms time window (significant by items, marginal by subjects). Looking at individual time windows we see a general trend of significant and marginal literal advantage starting from 200ms and extending until 800ms post noun onset.

Results of two-way t-test analyses on Literal Advantage by Time Window (Read-Aloud)

	df	Syntactically Unavailable - Lexically Available (e.g. kick. The bucket) CI	t	p	Syntactically Unavailable - Lexically Unavailable (e.g. kick. The pail) CI	t	p
Region							
by Subject	15	(-0.087 , 0.165)	0.657	0.521	(-0.022 , 0.338)	1.869	0.081 .
by Items	11	(-0.089 , 0.161)	0.641	0.535	(0.015 , 0.304)	2.437	0.033 *
200ms-300ms							
by Subject	15	(-0.048 , 0.313)	1.567	0.138	(0.004 , 0.404)	2.176	0.046 *
by Items	11	(-0.019 , 0.293)	1.935	0.079 .	(0.002 , 0.386)	2.222	0.048 *
300ms-400ms							
by Subject	15	(0.017 , 0.279)	2.41	0.029 *	(-0.038 , 0.408)	1.769	0.097 .
by Items	11	(0.01 , 0.278)	2.363	0.038 *	(0.021 , 0.388)	2.456	0.032 *
400ms-500ms							
by Subject	15	(-0.034 , 0.229)	1.573	0.136	(-0.074 , 0.385)	1.444	0.169
by Items	11	(-0.088 , 0.279)	1.146	0.276	(0.032 , 0.321)	2.683	0.021 *
500ms-600ms							
by Subject	15	(-0.086 , 0.215)	0.915	0.375	(-0.052 , 0.368)	1.603	0.13
by Items	11	(-0.11 , 0.224)	0.755	0.466	(0.005 , 0.321)	2.277	0.044 *
600ms-700ms							
by Subject	15	(-0.195 , 0.131)	-0.419	0.681	(-0.056 , 0.354)	1.552	0.142
by Items	11	(-0.162 , 0.123)	-0.306	0.765	(-0.028 , 0.321)	1.847	0.092 .
700ms-800ms							
by Subject	15	(-0.218 , 0.145)	-0.428	0.675	(-0.024 , 0.319)	1.832	0.087 .
by Items	11	(-0.204 , 0.114)	-0.623	0.546	(-0.019 , 0.307)	1.947	0.078 .
800ms-900ms							
by Subject	15	(-0.243 , 0.11)	-0.8	0.436	(-0.04 , 0.294)	1.626	0.125
by Items	11	(-0.256 , 0.115)	-0.841	0.418	(-0.077 , 0.318)	1.345	0.206
900ms-1000ms							
by Subject	15	(-0.16 , 0.168)	0.05	0.961	(-0.051 , 0.33)	1.561	0.139
by Items	11	(-0.151 , 0.136)	-0.115	0.911	(-0.092 , 0.343)	1.269	0.231

Table 8: Results of two-way t-tests comparing the literal advantage scores to 0 for our region of interest from 200ms to 1000ms post noun onset and for each of eight 100ms windows in that region

3.2.3.2.1.3 Interim Discussion

These results are strikingly similar to the results obtained in these same conditions in experiment 3b. Again we find evidence that individuals strongly consider the literal meaning in syntactically-unavailable contexts. For lexically unavailable strings such as *kick the pail*, participants strongly favor the literal interpretation and we see no evidence of consideration of the idiomatic interpretation at all. For lexically available strings, however, we replicate the findings from Experiment 3a. Recall that in experiment 3a we saw early preference for the literal interpretation of strings such as *kick the bucket*, but later observed competition between the literal and idiomatic interpretation. Crucially we also observe this same effect here though it appears to begin roughly 100-200ms earlier and be somewhat stronger in magnitude.

Recall that we had predicted that our Read Aloud task would in some way enhance the results obtained in the Read Silently version. The logic was that the deeper semantic activation resulting from reading our associate words aloud as opposed to reading them silently would cause the effects observed in the previous Experiment to manifest earlier and/or more robustly in this version of the task. Indeed this appears to be what we are seeing. Generally we observe the same qualitative pattern of results, with stronger statistical results overall and observe the late competition observed in Experiment 3a shifting earlier in the time course.

3.2.3.2.2 Syntactically Available Trials

3.2.3.2.2.1 Graph & Basic Discussion of Trends

Syntactically Available - Lexically Available
(e.g. *kick the bucket*) [Read-Aloud]
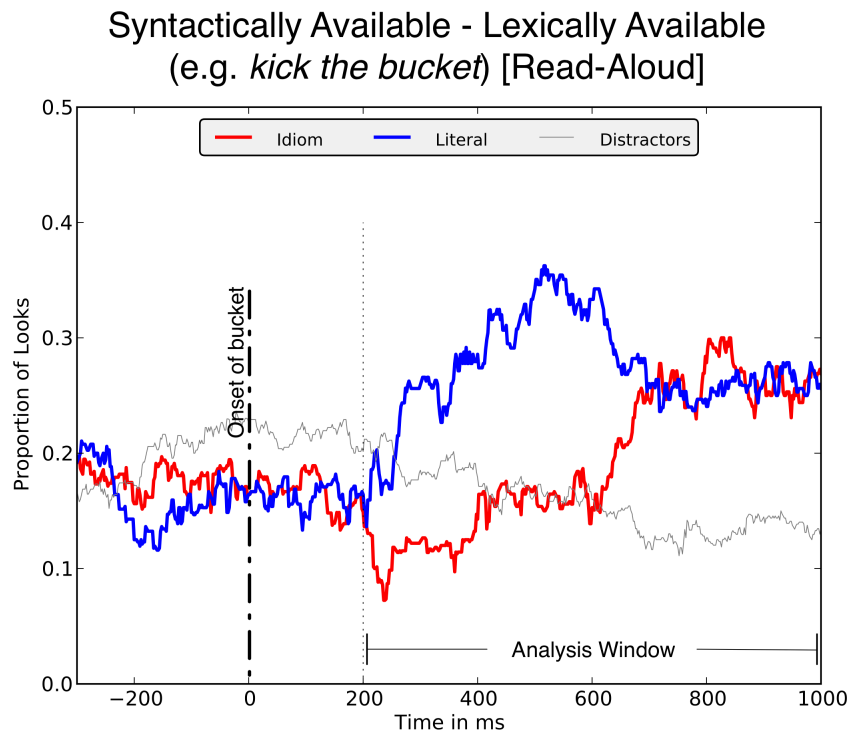
Figure 20: Looks to the idiom associate word (in blue/dark) vs. looks to the literal associate word (in red/light) as well as looks to the averaged distractor words (thin line in light grey). This figure shows looks in the Syntactically Available - Lexically Available condition (e.g. kick the pail), for the read-aloud task.

Syntactically Available - Lexically Unavailable
(e.g. *kick the pail*) [Read-Aloud]
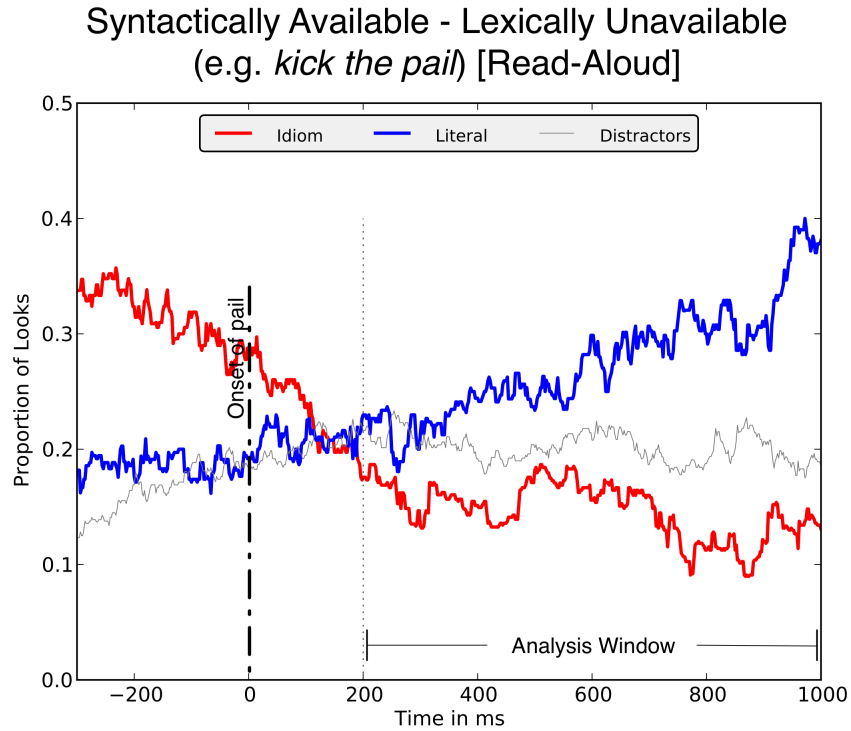
Figure 21: Looks to the idiom associate word (in blue/dark) vs. looks to the literal associate word (in red/light) as well as looks to the averaged distractor words (thin line in light grey). This figure shows looks in the Syntactically Available - Lexically Unavailable condition (e.g. kick the pail), for the read-silently task.

We turn now to the Syntactically Available condition. Figures 20 and 21 display the raw proportion of looks for the syntactically unavailable, read aloud trials in the Lexically Available and Lexically Unavailable conditions respectively. For Lexically Available (e.g. *kick the bucket*) trials we observe higher proportion of looks to the literal associate word than the idiomatic associate word or the averaged distractor words in the early time windows from 200ms-700ms followed by roughly equal proportion of looks to both the idiomatic and literal associates, both of which appear to be higher than the distractor baseline. This pattern suggests that participants are first considering only the literal

interpretation of the string, turning to also consider the idiomatic interpretation only later in the time-course.

For the Lexically Unavailable (e.g. *kick the pail*) trials we observe generally higher proportion of fixations to the literal associate as compared to either the idiom associate or averaged distractors. We do see evidence of strong consideration of the idiomatic meaning much earlier, shortly after the onset of the verb (e.g. *kick*) but that consideration seems to rapidly shift to consideration of the literal interpretation only after the onset of the critical noun (e.g. *pail*)

3.2.3.2.2.2 Overall Statistical Analyses

We again conducted a series of two-tailed t-tests over eight 100ms time regions starting at 200ms post noun onset and extending to 1000ms to evaluate these observations statistically (the full analyses are reported in Table 9). For the Lexically Available condition we find no significant literal or idiom advantage over the full time window. Examination of our individual time-windows confirm our visual observations that this literal advantage is focused upon the early time windows from 300ms to 600ms post noun onset, followed by competition between the idiomatic and literal associate.

For the Lexically Unavailable condition our analyses also suggest a literal advantage over the full time window. Our visual observations suggest that this literal advantage begins small in magnitude and slowly increases over the time-course, and our statistical analyses suggest that this is indeed the case. Analyses of the individual time windows reveals no

significant literal advantage over the time regions from 200ms to 700ms, and a significant literal advantage in all time windows from 700ms to 1000ms.

Results of two-way t-test analyses on Literal Advantage by Time Window (Read-Aloud)

	df	Syntactically Available - Lexically Available (e.g. kick the bucket) CI	t	p	Syntactically Available - Lexically Unavailable (e.g. kick the pail) CI	t	p
Region							
by Subject	15	(-0.048 , 0.205)	1.32	0.207	(0.017 , 0.231)	2.479	0.026 *
by Items	11	(-0.03 , 0.201)	1.631	0.131	(0.019 , 0.247)	2.571	0.026 *
200ms-300ms							
by Subject	15	(-0.025 , 0.218)	1.691	0.112	(-0.086 , 0.23)	0.967	0.349
by Items	11	(-0.048 , 0.254)	1.505	0.161	(-0.066 , 0.204)	1.127	0.284
300ms-400ms							
by Subject	15	(0.013 , 0.284)	2.34	0.034 *	(-0.085 , 0.212)	0.907	0.379
by Items	11	(-0.03 , 0.318)	1.816	0.097 .	(-0.052 , 0.205)	1.313	0.216
400ms-500ms							
by Subject	15	(-0.056 , 0.31)	1.48	0.16	(-0.045 , 0.233)	1.439	0.171
by Items	11	(-0.032 , 0.342)	1.824	0.095 .	(-0.029 , 0.252)	1.749	0.108
500ms-600ms							
by Subject	15	(-0.004 , 0.382)	2.087	0.054 .	(-0.058 , 0.19)	1.131	0.276
by Items	11	(-0.012 , 0.364)	2.06	0.064 .	(-0.085 , 0.254)	1.097	0.296
600ms-700ms							
by Subject	15	(-0.083 , 0.285)	1.175	0.258	(-0.043 , 0.252)	1.507	0.153
by Items	11	(-0.075 , 0.264)	1.23	0.244	(-0.05 , 0.296)	1.564	0.146
700ms-800ms							
by Subject	15	(-0.183 , 0.158)	-0.157	0.878	(0.032 , 0.308)	2.632	0.019 *
by Items	11	(-0.116 , 0.118)	0.027	0.979	(0.003 , 0.359)	2.24	0.047 *
800ms-900ms							
by Subject	15	(-0.207 , 0.151)	-0.332	0.745	(0.058 , 0.349)	2.983	0.009 **
by Items	11	(-0.129 , 0.121)	-0.068	0.947	(0.017 , 0.356)	2.423	0.034 *
900ms-1000ms							
by Subject	15	(-0.139 , 0.151)	0.088	0.931	(0.053 , 0.388)	2.802	0.013 *
by Items	11	(-0.099 , 0.129)	0.291	0.776	(0.046 , 0.416)	2.752	0.019 *

Table 9: Results of two-way t-tests comparing the literal advantage scores to 0 for our region of interest from 200ms to 1000ms post noun onset and for each of eight 100ms windows in that region

3.2.3.2.2.3 Interim Discussion

As with the results of the previous experiment, the competition observed in the Lexically Available condition strongly support our prediction that individuals access both the literal and idiomatic interpretations during the processing of unbiased potentially idiomatic strings such as *kick the bucket*. Additionally, we find an early focus on the literal interpretation in these trials, which supports the view that literal processing has priority

during the interpretation of these strings. However we fail to find any significant evidence that participants consider the idiomatic interpretation for strings such as *kick the pail*, which seems to go against our predictions.

Considering the task effects however may allow us to explore this further. The critical difference in the Lexically Available condition is that we observed early focus on the literal interpretation when participants read the associate words aloud, but did not observe this when they read the words silently. With regards to the Lexically Unavailable condition, we see what appears to be a qualitatively different result. Participants in the read silently task exhibited early consideration for the idiomatic interpretation which then shifted to a focus on the literal interpretation. In the read aloud task we see a slowly growing advantage for the literal interpretation, however we *also* see a suggestion of a preference for the idiomatic interpretation *before* the noun around the onset of the verb. Thus this looking behavior may represent a qualitatively different pattern or the same qualitative result shifted much earlier in the time-course.

3.2.4 Condition Comparisons

Now that we have a grasp of the general data patterns present in our conditions we can turn our attention to the full factor analysis of or three manipulations. We perform three sets of analyses to further examine our results. The first set of analyses examine the effect of our Syntactic, Lexical and Task manipulations on looks to the idiom and literal associates separately. These analyses are important, as we expected that Syntactically

Availability would play a large role in looks to the literal associate, while Lexically Available would play a large role in looks to the idiom associate. In short, these independent analyses will allow us to examine (i) which of our manipulations are driving looks to the literal and idiomatic associates respectively, and (ii) whether we have evidence of baseline preferences already at the onset of our critical time window.

Answering the first question allows us to better understand the way in which our syntactic, lexical and task manipulations are driving our results. Answering second gives us a sense of whether individuals are anticipating the idiomatic or literal interpretation at the outset of our time window. This is particularly valuable information, as at the start of our time window participants have already heard the verb (e.g. *kick*), and hence individuals may already be preparing to consider the idiomatic interpretation in these cases, These analyses were performed using a multilevel empirical-logit regression analysis (see Barr, 2008) over the full time windows. This analysis is advantageous for answering these questions for two reasons. First, since time can be included as a factor in this analysis, 'main' effects of our conditions indicate effects at the intercept (i.e. where time = 0). Hence this analysis provides a means of evaluating baseline effects existing in our data. Secondly, looks to the idiomatic and literal targets are not independent. To handle this in the ANOVA and t-tests we chose to examine Literal Advantage scores rather than attempt to compare the two curves directly. However if we wish to examine our curves separately this is not possible.

There are several suggestions in the literature that suggest a way of handling this dependency. These methods primarily involve the transformation of the data into difference scores (Huettig & Altman, 2007), probability of fixating a given target (Allopenna, Magnusson & Tanenhaus, 1998), log-ratios (Ferguson & Breheny, 2011, Huettig & Altmann, 2005) or other ratio measures (Dahan & Tanenhaus, 2005). This analysis addresses our dependencies by computing the empirical-logit, which is an approximation of the log-odds of fixating a given target in a given time window. Essentially this metric computes the log-odds of fixating a given item as compared to *not* fixating that target (which includes possibly fixating on something else) in consecutive 50ms bins. Use of the log-odds scale has been shown generally to provide a better approximation for examining effects over proportional categorical data (Jaeger, 2008).

While this method also provides information regarding the interaction between our factors and time (interpreted in the model as measures of the slope of our data), there is currently a heated debate in the field regarding the reliability of slope measures in eye-tracking analyses. Thus to analyze the effects of our conditions over our time-course we will then go back to our previously discussed Literal Advantage scores and evaluate the effects of our factors on these scores over our time region and over our eight consecutive 100ms bins.

We will first examine the effects of our factors on looks to the literal associate word (e.g. *foot*) and then turn our attention to the effect of our factors on looks to the idiom

associate word (e.g. *death*). Finally we will turn to a full analysis of our literal advantage metric using ANOVA over our 100ms bins.

3.2.4.1 Looks to the Literal Associate

Examining looks to our literal associate word reveals a robust main effect of Syntactically Availability and an interaction between time and Syntactically Availability, full analyses are reported in Table 10. Examining the main effect of bias reveals a negative parameter estimate for in our model. As we coded Syntactically Unavailable as negative in our model and Syntactically Available as positive, this effect translates into a significantly higher chance of fixating the literal associate word when in Syntactically Unavailable conditions as compared to Syntactically Available conditions at the intercept (i.e. from the earliest bin). This result confirms what we have already observed in the data. Participants are more likely to fixate on the literal associate when the potentially idiomatic string is placed in a syntactically incompatible context.

It is also worth noting that the main effect of Syntactically Availability obtains at the outset of our trials (from the intercept at 200ms post noun onset), thus this effect also suggests that consideration of the literal interpretation in the Syntactically Unavailable condition obtains as a baseline effect, carrying over from before our critical time-window. Examining our results suggests that this is also the case, with participants already settling on the literal interpretation even before the onset of the noun. This is reasonable considering that participants already have at this point already heard the verb and sentential break, thus ruling out any consideration of idiomatic meaning.

Generally, the results also suggest that the syntactic manipulation is largely responsible for modulating look to the literal associate, with our lexical manipulation having little effect on this behavior.

Analysis of Baseline Looks to the Literal associate word

		Estimate	Error	t	p	
SynAvail						
	Subject	-1.1998	0.3405	-3.524	0	***
	Items	-1.32617	0.38258	-3.466	0.001	**
LexAvail						
	Subject	-0.1626	0.3405	-0.477	0.633	
	Items	-0.39769	0.38258	-1.039	0.299	
Task						
	Subject	-0.1093	0.4625	-0.236	0.813	
	Items	0.04377	0.38258	0.114	0.909	
SynAvail x LexAvail						
	Subject	-0.6881	0.6809	-1.01	0.312	
	Items	-0.15374	0.76517	-0.201	0.841	
SynAvail x Task						
	Subject	0.7647	0.6809	1.123	0.261	
	Items	0.7245	0.76517	0.947	0.344	
LexAvail x Task						
	Subject	0.1187	0.6809	0.174	0.862	
	Items	0.3374	0.76517	0.441	0.659	
SynAvail x LexAvail x Task						
	Subject	0.1157	1.3618	0.085	0.932	
	Items	0.11833	1.53034	0.077	0.938	

Table 10: Analyses of baseline effects for our looks to the literal associate word over our analysis window.

3.2.4.2 Looks to the Idiom Associate

Looks to the Idiomatic associate were examined independently using the same method as presented in the previous section. Overall the results are more complex than the results obtained for the literal associate. We see hints of a weak main effect of Syntactically

Availability (marginal by subjects) which suggests that looks to the idiom associate at the onset of our time-window are somewhat higher in the Syntactically Unavailable condition than the Syntactically Available condition as well as a marginal baseline effect of Lexically Availability (marginal by items) suggesting that looks to the idiom associate are higher in the Lexically Unavailable condition as compared to the Lexically Available condition. The former result suggests that participants have a slightly higher base-line preference for the idiomatic interpretation in the Syntactically Unavailable condition, which may be due to the longer time that they have to consider the words on the screen (due to the pause between sentences) in that condition. The latter effect relates to the observed early preference for the idiomatic expression in Lexically Unavailable trials. We also see a marginal baseline effect of our task, indicating that participants are more likely to fixate the idiom associate word in the no-task condition as compared to the read-aloud condition.

Generally these results are weaker than those obtained when examining looks to the literal associate, however they do suggest that looks to the idiom associate word are being modulated by all of our factors from the onset of our analysis window.

Analysis of Baseline Looks to the Idiom associate word

		Estimate	Error	t	p	
SynAvail						
	Subject	0.6089	0.3636	1.675	0.094	.
	Items	-0.0154	0.43396	-0.035	0.972	
LexAvail						
	Subject	-0.4594	0.3636	-1.263	0.206	
	Items	-0.80536	0.43396	-1.856	0.063	.
Task						
	Subject	-0.7242	0.5221	-1.387	0.165	
	Items	-0.76951	0.43396	-1.773	0.076	.
SynAvail x LexAvail						
	Subject	-1.1388	0.7272	-1.566	0.117	
	Items	-1.18898	0.86791	-1.37	0.171	
SynAvail x Task						
	Subject	-0.2333	0.7272	-0.321	0.748	
	Items	0.54904	0.86791	0.633	0.527	
LexAvail x Task						
	Subject	0.64	0.7272	0.88	0.379	
	Items	0.5453	0.86791	0.628	0.53	
SynAvail x LexAvail x Task						
	Subject	-0.692	1.4543	-0.476	0.634	
	Items	-0.92148	1.73582	-0.531	0.596	

Table 11: Analyses of baseline effects for our looks to the idiom associate word over our analysis window.

3.2.4.3 Literal Advantage Analysis

The results of our analyses over the Literal Advantage scores are presented in Table 12.

Over the full time window we see an effect of Lexical Availability and an interaction

between Syntactic Availability and Task, both significant by subjects only. The effect of

Lexical Availability indicates significantly higher literal advantage scores in lexically

unavailable trials as compared to lexically available trials, meaning that participants were

more likely to consider the literal meaning over the idiomatic meaning in trials in *kick the*

pail trials as compared to *kick the bucket* trials. For the interaction, we see larger literal

advantage scores for Syntactically Unavailable trials than Syntactically Available trials when participants read silently, but no significant difference in Literal Advantage when participants read aloud (and a trend in the opposite direction). Thus reading the words aloud seems to generally wash out the distinction between our syntactic conditions.

Examining our factors by time window reveals a significant effect of our lexical availability on our results in the time windows extending from 700-1000ms. In earlier windows we see a significant main effect of Syntactic Availability extending from 200ms - 500ms and a marginal effect in the 500-600ms window. This effect manifests as significantly lower literal advantage scores for syntactically available trials than for syntactically unavailable trials. Thus overall participants do tend to utilize syntactic information in the way we predicted with more consideration of the literal interpretation when the idiomatic string is in a syntactically invalid configuration. We also see an interaction between Syntactic and Lexical Availability in the 300-500ms time region, and an interaction between Syntactic Availability and Task in the 400-600ms time window. The former is a result of significantly lower literal advantage scores in *kick the pail* trials as compared to *kick the bucket* trials. Recall that participants showed a strong early preference for the idiomatic interpretation in the *kick the pail* trials. In *kick. The bucket/pail* trials the effect is reversed as we had expected, with more consideration of the literal meaning for *pail* as opposed to *bucket.* The latter interaction is similar to the one we saw overall, with larger differences in Literal Advantage scores between Syntactically Available and Syntactically Unavailable trials when subjects read silently

than when they read aloud. Finally we see a marginal main effect of task in the 300-400ms time window, with higher literal advantage scores in the read aloud task as compared to the read silently task.

Analysis of Variance of our Factors (Syntactic Availability x Lexical Availability x Task) over our region and over consequitive 100ms time windows

Region	Syntactic Availability			Lexical Availability			Task			Syntactic x Lexical			Syntactic x Task			Lexical x Task			Syntactic x Lexical x Task		
	F	p		F	p		F	p		F	p		F	p		F	p		F	p	
Subject	1.073	0.322		8.304	0.015	*	0.496	0.496		1.436	0.256		6.451	0.027	*	0.334	0.575		0.002	0.967	
Items	1.283	0.266		1.802	0.19		0.252	0.62		0.805	0.377		1.432	0.241		0.437	0.514		0.002	0.968	
200ms-300ms																					
Subject	6.407	0.028	*	0.008	0.93		2.091	0.176		0.477	0.504		1.009	0.337		0.072	0.794		0.147	0.709	
Items	8.493	0.007	**	0	0.987		2.02	0.166		0.326	0.573		1.514	0.228		0.204	0.655		0.089	0.767	
300ms-400ms																					
Subject	5.284	0.042	*	0.79	0.393		3.587	0.085	.	9.719	0.01	*	1.849	0.201		0.593	0.457		0.048	0.831	
Items	6.184	0.019	*	0.626	0.435		2.243	0.145		2.237	0.145		1.626	0.212		0.122	0.73		0.139	0.711	
400ms-500ms																					
Subject	3.139	0.104		0.029	0.868		2.288	0.159		6.954	0.023	*	6.192	0.03	*	0.036	0.852		0.772	0.398	
Items	3.011	0.093	.	0.006	0.937		0.752	0.393		2.121	0.156		2.1	0.158		0.017	0.898		0.481	0.493	
500ms-600ms																					
Subject	1.981	0.187		0.002	0.961		0.76	0.402		2.356	0.153		7.265	0.021	*	0.016	0.902		0.64	0.441	
Items	0.938	0.341		0.058	0.812		0.443	0.511		1.568	0.22		1.519	0.227		0.002	0.966		0.251	0.62	
600ms-700ms																					
Subject	0.234	0.638		3.575	0.085	.	0.025	0.877		1.098	0.317		2.509	0.142		0.142	0.714		0.279	0.608	
Items	0.348	0.56		1.306	0.262		0.213	0.647		0.653	0.425		2.026	0.165		0.585	0.45		0.375	0.545	
700ms-800ms																					
Subject	0.006	0.941		12.591	0.005	**	0.013	0.912		0.108	0.749		0.45	0.516		0.621	0.447		0.043	0.839	
Items	0.057	0.813		7.816	0.009	**	0.026	0.872		0.032	0.859		0.426	0.519		0.735	0.398		0.027	0.869	
800ms-900ms																					
Subject	0.012	0.916		18.708	0.001	**	0.081	0.781		0.028	0.869		1.259	0.286		0.549	0.474		0.018	0.895	
Items	0.009	0.924		11.439	0.002	**	0.096	0.759		0.097	0.757		0.743	0.395		1.494	0.231		0.027	0.871	

Table 12: ANOVA analyses of the effects of our conditions and their interactions on our region of interest

from 200ms to 1000ms post noun onset, as well as on each of eight individual 100ms windows within that

region.

3.3 Experiment 3: Discussion

In this Chapter we set out to investigate two main questions: (i) whether individuals consider the idiomatic interpretation of a given string in syntactically incompatible contexts, and (ii) whether semantically related, but non-idiomatic strings such as *kick the pail* still result in consideration of idiomatic meaning. To address these questions we examined participants' eye-movements over a time window starting 200ms post noun onset and extending to 1000ms. We analyzed these eye-movements in several ways to address different aspects of our questions. First we examined the qualitative patterns of looking behavior by examining literal advantage scores over eight 100ms time slices. These analyses gave us an overall picture of individual's behavior in each of our conditions which can be compared qualitatively. We then performed a quantitative analysis of our conditions examining the effects of our manipulations on looks to the idiomatic and literal associate words separately to which of our factors were driving the looking behavior related to each of these visual targets and to evaluate potential baseline effects. Finally we performed a series of ANOVAs over our regions and individual time windows to examine the quantitative effects of our factors on participants' looking behavior.

Our primary prediction regarding the manipulation of Syntactically Availability was that individuals would not consider the idiomatic interpretation of our strings if they were placed in a syntactically incompatible context. Recall that we suggested that the super-lemma phrasal representation proposed by the Hybrid Representation Hypothesis, may

act as a sort of *gate* to the conceptual level. Thus upon receiving an incompatible syntactic context, this gate would impede or prevent further spreading activation to the conceptual level, and hence prevent consideration of the idiomatic meaning. Our results largely support this prediction, as we find more consideration of the literal interpretation over the full time course in our Syntactically Unavailable condition, but do not find this pattern of behavior in our Syntactically Available condition. We also found evidence in our independent analysis of looks to the literal associate, that looks are driven primarily by this syntactic manipulation, and that participants appear to have a general base-line preference for the literal interpretation even at the beginning of our time-window.

Taken as a whole these results suggest that syntactic information *is* used online help guide individuals' interpretations of these strings. Additionally, the base-line effects suggest that this information is rapidly utilized to constrain the possible meaning of these strings. However, we also found some evidence for late consideration of the idiomatic interpretation of ambiguous idioms like *kick the bucket* even when the syntactic context should have ruled out an idiomatic interpretation, and observed that this general pattern of later consideration of idiomatic meaning obtained in both of our tasks. This result deserves further scrutiny.

Of central interest is why individuals would consider the idiomatic meaning at all in our Syntactically Unavailable condition, and secondly why such consideration would only extend to the Lexically Available trials. The former point is possible and predictable from

the architecture of our model. While the super-lemma associated with *kick the bucket* would be incompatible with the syntactic structure in which the string occurs, this would not prevent the individual lemmas *kick* and *bucket* from passively activating the idiomatic representation anyway. Thus it is possible that the super-lemma representation is positively primed due to the consecutive activation of *kick* and *bucket* and also negatively primed due to the syntactic mismatch. This early negative priming may explain why the initial behavior is not to consider the idiomatic interpretation and occurs only later. This proposal would then have to explain why the same doesn't happen for *kick the pail*, and the easiest explanation would simply be that the overall positive activation to the super-lemma is weaker due to *pail* being indirectly associated with the component super-lemma.

Another, and perhaps more reasonable explanation would be that this late idiomatic consideration is some sort of post-processing effect caused by the close proximity of *kick* and *the bucket*. Under this explanation the general processing of the sentence would proceed normally until at some point roughly 600-700ms post noun onset some post-processing memory component recognizes the idiomatic string. It is difficult to determine which of these views are correct, however post-experimental debriefing of participants suggests that whatever the source, it is probably subconscious, as the vast majority of participants indicated that they did not notice the inclusion of idiomatic expressions in the study, and none self-reported that they'd noticed these expressions in the cross-sentential contexts.

Regardless, this result does indicate that the strong version of our prediction, namely that the super-lemma representation acts as a firm gate to the conceptual level, seems to be in need of revision. The late consideration of idiomatic meaning in these cases, taken with our potential explanations above suggest either that (i) the ability for the super-lemma to act as a gate to the conceptual level is not strict but is governed by a process of spreading activation and inhibition, which in certain cases may allow some activation through to the conceptual level even when syntactic factors preclude the idiomatic interpretation, or that (ii) syntactic incompatibility only prohibits activation spreading from the super-lemma to the conceptual level for a short time, and the effects we are noticing are post-processing considerations of the idiomatic sequence of words.

With regards to our second question, recall that we predicted consideration of the idiomatic interpretation for both Lexically Available (e.g. *kick the bucket*) and Lexically Unavailable (e.g. *kick the pail*) strings. For *kick the bucket* the model predicts that activation of the super-lemma idiomatic representation necessarily involves activation of component literal lemmas and, by extension, their literal interpretation and conceptual content. The model also predicts the same sort of activation for *kick the pail*, by virtue of indirect activation spreading from *pail* to the conceptually related *bucket* and our read-silently results also support this prediction. Again, our results generally confirm this prediction, as we observed the expected competition between the literal and idiomatic interpretations in all syntactically available conditions except for one. What was not expected, was that consideration of the idiomatic interpretation would be *stronger* for

kick the pail than *kick the bucket* in the read-silently task, and that we would observe no competition at all for *kick the pail* in the read-aloud task.

We suggest that the former result may be due, in part, to conventionality considerations. There is evidence that the idiomatic sense of a expressions such as *kick the bucket* is both more frequent and more familiar (Cronk et al, 1993; Popiel & McRae, 1988). In line with these findings several studies have provided evidence that the processing advantage apparent for figurative expressions is due in part to the high saliency and frequency of their figurative usage as compared to their literal usage (Giora, 1999; Gibbs, 1980; Tabossi et al, 2009; Giora & Fein, 1999; Schwiegert, 1986). Interestingly, Gibbs (1980) discovered that memory recall was better for unconventional literal usages for idioms than for their idiomatic usage. He attributed this result to what he termed a *double-take* effect, in which individuals first process the conventional idiomatic sense, then are forced to reject that interpretation and revise to the less conventional literal interpretation. This revision process essentially boosts the memory imprint of the incorrect, but more conventional interpretation.

Applying this to our result is a bit less straight-forward, as presumably the conventional usage of the phrase *kick the pail* is indeed the literal interpretation. However, it may be the case that upon hearing *kick the…* the parser is already entertaining the possibility of the idiom. This sort of forward prediction is indeed expected under a wide array of contemporary sentence processing models (see Jaeger & Snider, 2008; Levy, 2008;

Jurafsky, 2003 and others). Upon receiving the input *pail* is then surprised at the continuation, specifically at the failure for the sentence to continue with *bucket*. This sort of double-take reaction may temporarily boost activation of the conventional idiomatic meaning of *kick the bucket* by virtue of the fact that *kick the pail* is lexically similar enough to still activate the necessary super-lemma representation and is less frequent and hence may exhibit stronger priming effects (Bock, 1986; Masson & Bodner, 2003, Versace & Nevers, 2003; Forster & Davis, 1984). Regardless of the explanation, our time-course information shows that this boost is both rapid and early.

This explanation, of course, makes the complete lack of idiomatic consideration in the read-aloud task even more mysterious. A possible explanation for this discrepancy may be found in the general trend of the read-aloud task to boost the effects obtained in the read-silently task. Thus it is possible that the early consideration of idiomatic meaning observed for *kick the pail* in the read-silently condition is shifted earlier, to the onset of the verb in the read-aloud condition. Thus the increased depth of semantic activation caused by reading the idiomatic target aloud induces participants to strongly consider the idiomatic interpretation very early, just after the onset of *kick*. Upon hearing *pail* the parser then rapidly shifts to the literal interpretation over our critical time window. In the case of the read-silently condition, however, the shallower semantic activation results in the parser only settling on the idiomatic interpretation later, with this effect manifesting at the left edge of our time window, and then later in the time window we see the rapid shift to the literal interpretation. Of course, if this was the case we should see a similar pattern

of early idiomatic consideration in the *kick the bucket* cases, but such effects are not apparent in our data. In general, this effect remains somewhat mysterious and further experimentation is required to explore this phenomenon.

In conclusion this chapter explored the relationship between an syntactic compatibility and semantic association in the on-line interpretation of idiomatic expressions. We found evidence supporting a weak view of the super-lemma, in which the super-lemma serves as a robust but permeable gate to the conceptual level when interpreting idiomatic strings, and found robust evidence that syntactic information is used rapidly when determining the correct interpretation of potentially idiomatic strings. We also found evidence supporting the proposal that the representation of an idiomatic expression stands in a principled relationship to its literal components, and furthermore that standard semantic priming effects (e.g. *bucket* primes *pail*) can also actively modulate activation of idiomatic representations during comprehension. Finally, in examining our task effects, we found evidence for stronger semantic activation when individuals are required to speak the visual stimuli aloud than when they are only required to read them silently. This difference in depth of semantic activation manifested as both stronger effects in our analyses of looking behavior, and a general trend of these effects arising earlier.

Chapter 4: Eye-Tracking: Contextual Effects

In the previous chapter we focused on the structural properties of idiomatic representations and examined the relationship between idiomatic expressions and their literal components. In Chapter 2 we examined ambiguity resolution and explored the different effects of recovering idiomatic and literal information when contextual expectations lead to an incorrect interpretation. In this chapter we present a set of experiments which attempts to combine the exploration of how idioms are represented and related to other items in the lexicon begun in chapter 3 and the effects of contextual bias on idiom processing observed in chapter 2.

As discussed in chapters 1 and 2, current theories make divergent predictions regarding the priority of idiomatic and literal interpretation during the processing of idioms, and particularly with respect to the influence of context. We will first spend some time to remind ourselves of the various proposals in the literature and examine these proposals specifically as they apply to the role of contextual bias.

4.1 Introduction

4.1.1 Direct Access

Direct access theories, such as that proposed by Gibbs (1980, 1985), claim that individuals are capable of accessing the meaning of an idiomatic expression directly without computation of its literal components. There is considerable evidence for this approach, though the details vary from one proposal to another. Bobrow & Bell (1973)

proposed that access to idiomatic meaning was made available by a special idiom-processing mode, and claimed that given sufficient bias to initiate this mode one could directly access idiomatic meaning without the need of literal processing. Gibbs (1980) proposed a priority for conventional usage by which unconventional literal interpretations can be bypassed given sufficient context to do so (but see Giora, 1999). Gibbs et al (1989) go further, proposing that even when an individual processes an idiomatic expression 'compositionally', the process may not necessarily involve activation of the literal meaning of the phrase. In more recent work, Gibbs has adopted a conceptual metaphor version of idiom processing in the spirit of Lakoff & Johnson (1980), which includes the proposal that literal interpretation is unnecessary for access to idiomatic meaning (Gibbs et al, 1997).

In general the Direct Access view proposes either that literal processing is not necessary for access to idiomatic meaning, or that idiomatic meaning has priority over literal meaning in some sense. This view does not preclude the possibility that an individual might process the literal meaning of a phrase like *kick the bucket* given sufficient reason to do so. Crucially, however, this view proposes that the literal interpretation does not have any priority over the idiomatic interpretation, and that access to the literal meaning is not required to access idiomatic meaning.

In relation to context, the Direct Access view makes two possible predictions. First, it may be the case that idiomatic expressions have priority in processing regardless of

context. Hence given an ambiguous string such as *kick the bucket* the idiomatic interpretation should arise regardless of contextual bias, with the literal interpretation arising only if there is sufficient contextual support to warrant the string's literal meaning. Another possibility is that idiomatic processing can, in principle, have priority. This view is in line with Bobrow & Bell's (1973) proposal of an idiom processing mode. Hence under an idiomatically biasing context, individuals will adopt an idiomatic mode of processing and hence we would expect the idiomatic meaning, but not necessarily the literal meaning to be activated. Conversely, given a literally biasing context, individuals will adopt a literal mode of processing and fail to retrieve the idiomatic interpretation.

4.1.2 Literal Obligation

In contrast to the Direct Access view are proposals that literal meaning *does* have priority over idiomatic meaning. In general the notion of literal priority is the somewhat unfair conflation of two distinct proposals. The first of these we will call *literal obligation*, which is the claim that the processing of the literal interpretation of an idiomatic string is obligatory to a certain extent. Models differ with respect to how much literal processing is obligatory, and in the exact order of operations of these processes.

The Lexical Representation Hypothesis, for example, falls into this class of models (Swinney & Cutler, 1979), and proposes that literal processing is generally obligatory and occurs in parallel with idiomatic processing. Likewise the Configuration Hypothesis (Cacciari & Tabossi, 1988), discussed in Chapters 1 and 2, proposes that in unbiased

contexts literal processing proceeds obligatorily until such time as the parser accumulates sufficient evidence to recognize a given string as idiomatic. Likewise for the Hybrid Representation Hypothesis (Cutting & Bock, 1997; Sprenger et al, 2006), in which access to the idiomatic super-lemma during comprehension necessarily requires activation of the literal component lemmas.

These models, of course, differ a great deal in both their proposed time-course of literal processing as compared to idiomatic processing, and in the extent of literal processing expected. Under the Lexical Representation Hypothesis, for example, both literal and idiomatic processing proceed in parallel, thus for an expression such as *kick the bucket* both the idiomatic and literal interpretation would be computed. For the Configuration Hypothesis, literal processing must occur until the parser accumulates sufficient evidence for idiomaticity. The amount of literal processing required will vary depending on the idiom and contextual factors, with the extent of literal processing ranging from perhaps none (in the case of overwhelming contextual cues) to the entire string. Finally, the Hybrid Representation Hypothesis requires some degree of literal processing, at minimum the activation of the component literal lemmas, in order to activate the idiomatic interpretation. However, beyond this basic activation no further literal processing is necessarily required for the processing of idioms.

These models also differ with respect to the role of context in guiding the parser through the interpretation of ambiguous strings. Both the Configuration Hypothesis and the

Hybrid Representation Hypothesis predict some partial activation of the literal meaning, even in cases in which contextual biases suggest that the idiomatic interpretation is more likely. For the Hybrid Representation Hypothesis, this activation is a necessary condition for the activation of the conceptual meaning of the idiom. For the Configuration Hypothesis the argument is that in the normal case, some amount of literal processing precedes activation of the idiom, however the proponents of the Configuration Hypothesis leave open the possibility that sufficient context may result in a more direct access like retrieval of the idiomatic meaning (Tabossi & Zardon, 1993).

The predictions of the Lexical Representation Hypothesis are less clear. One possibility is that both the literal and idiomatic interpretations will be activated in parallel upon encountering the ambiguous string. Another option is that, similarly to the idiom processing mode view, prior contextual bias will serve to restrict the possible interpretations of the ambiguous string (see Swinney & Hakes, 1976). Thus in a literally biasing context we may only see activation of the literal interpretation, and in idiomatically biasing contexts only the idiomatic interpretation.

4.1.3 Literal Priority

In this section we will consider the related, but different notion, namely that literal processing is not simply obligatory, as discussed in the previous section, but is in fact *necessary* for access to the idiomatic meaning. We will refer to this proposal as *literal priority*. Turning our attention back to the three models just discussed, we can

demonstrate that this proposal is independent from the proposal of *literal obligation.* The Lexical Representation Hypothesis, for example, under a strong view of parallel processing, proposes a *literal obligation* in that the processing the literal interpretation of an idiomatic string is not optional. However it does not propose *literal priority*, in that the processing of the literal interpretation is not necessary for the idiomatic meaning to be available.

For the Configuration Hypothesis, the model does suggest a *literal priority* in that access to the idiomatic representation is only granted upon obtaining sufficient evidence for idiomaticity. In the normal, unbiased case this evidence takes the form of incremental literal processing of the expression until such evidence is obtained. However it is possible, as mentioned earlier, that given overwhelming contextual bias, such evidence may be available without having to process any literal structure at all.

For the Hybrid Representation Hypothesis the architecture of the model requires some level of literal activation before the idiomatic interpretation is even made available, thus these models are firmly *literal priority* models. Recall that under this view the activation of the idiomatic super-lemma is mediated by its component literal lemmas, hence to retrieve the idiomatic meaning requires that the comprehender at minimum activates the literal component lemmas.

4.1.4 Motivation

The experiments presented in this chapter are designed to further examine the predictions made by the Hybrid Representation Hypothesis of idiom representation (Cutting & Bock, 1997; Sprenger et al, 2006) as it applies to idiom comprehension and to contribute to our understanding of the time-course of idiom processing. Generally, we will continue the examination between the relationship between idioms and their literal components begun in Experiment 3 in chapter 3 by examining whether individuals still exhibit consideration for the impossible idiomatic interpretation of semantically related strings (e.g. *kick the pail*) under conditions in which narrative context biases individuals toward either a literal or idiomatic interpretation of the string.

In examining the effects of contextual bias on idiom processing, we will also further explore the behaviors we uncovered in Experiment 2 in chapter 2. Recall that our findings in Experiment 2 lead us to suggest that individuals obligatorily consider the literal interpretation of ambiguous strings, even when contextual bias pushes them toward an idiomatic interpretation. When context suggests a literal interpretation, however, we claimed that participants either do not consider the idiomatic representation, or do so only weakly. The methodological paradigm employed by Experiment 2 (self-paced reading) did not allow us to discern between these two possibilities, however the current methodology (eye-tracking) grants us a richer view of the time-course of individuals' consideration during the processing of these expressions which may allow us to shed some light upon this question. Given the results obtained in Experiment 2, we predict that

individuals *will* consider the literal and idiomatic interpretations of idioms (e.g. *kick the bucket*) and semantically associated strings (e.g. *kick the pail*) in idiomatically biased trials. However for literally biased trials, we predict strong consideration of the literal meaning with weak or no consideration of the idiomatic interpretation.

As we found that participants behaved differently depending upon whether they were required to read our visual target words aloud or silently in Experiment 3, we will also examine participant behavior in these two different tasks in this experiment. Recall that looking behavior appeared to be enhanced by the act of reading our stimuli aloud in Experiment 3. We attributed this effect to deeper semantic activation of the relevant word as a result of accessing, producing (and thus hearing the production) of its phonological form (van Orden, Johnston & Hale; 1988). We include this between-subjects manipulation in this experiment as well, as it is possible that this depth of activation may interact with sentential context. We expect this task manipulation to play largely the same role as it did in the previous experiment, with earlier or greater magnitude effects observed in the read-aloud task as compared to the read-silently task. Additionally, these effects may interact with context such that idiomatic activation under literally biasing contexts, if it occurs at all, may be enhanced in the read-aloud as compared to the read-silently task.

4.2 Experiment 4: Contextual & Lexical Effects

In this experiment we investigate the question of whether contextual bias influences comprehenders' online interpretation of potentially idiomatic ambiguous strings such as *kick the bucket*. Additionally, we found evidence in the previous chapter for idiomatic activation even for semantically related non-idiomatic strings such as *kick the pail*. Thus to further examine the effects of context, and to further explore the relationship between idiomatic expressions and their literal components we also examined lexical availability (e.g. *kick the bucket* vs. *kick the pail*) modulates the effect of context during sentence processing. As with the previous experiment, participants' eye-movements were recorded as they listened to potentially idiomatic sentences and looked at four words shown on the computer screen. Also as with the previous experiment, we also modulated the task demands of the experiment with half of our participants reading the four words silently before the onset of the audio sentence (Experiment 4a) and half reading the four words aloud (Experiment 4b).

4.2.1 Method

4.2.1.1 Participants

32 undergraduate students from the University of Southern California participated in this study. All participants were native speakers of American English. None of the participants participated in any of the previous experiments.

4.2.1.2 Materials

This experiment utilized the same idioms and semantic associates generated for the experiments presented Chapter 3. These items may be found in Appendix II. Unlike the previous experiment, however, I was interested in investigating (i) how contextual bias modulates consideration of idiomatic and literal interpretations of ambiguous idioms such as *kick the bucket* and (ii) whether contextual bias has similar effects on semantically related non-idiomatic strings such as *kick the pail*. The former question allows us to examine the effects of contextual bias on idiom processing in greater detail than in our self-paced reading study presented in Chapter 2, while the latter provides further insights into the relationship between idioms and their component words explored in the experiments in Chapter 3.

For our lexical manipulation we used the same unaltered idiomatic and altered semantic-associate pairs used in Chapter 2. Hence in the **Lexically Available** condition our stimuli consisted of ambiguous idioms such as *kick the bucket, find her feet* and *smell a rat*, while in the **Lexically Unavailable** condition we altered the idiom by changing to final noun to a semantically related word resulting in strings such as *kick the pail, find her toes* and *smell a mouse*.

Bias was manipulated by placing each idiom and semantic-associate into two possible sentential frames. All stimuli were placed into the second of a pair of carrier sentences that told a short story. The second sentence always contained the idiom or semantic-

associate as a contiguous string, and the first sentence was designed to bias the interpretation of the string toward either the idiomatic or literal interpretation, resulting in our **Idiomatically Biased** and **Literally Biased** conditions respectively.

	Idiomatically Biased	Literally Biased
Lexically Available	The department was in an uproar when top secret information was leaked to the press. Jack was not surprised, he had **smelled a rat** in the office ever since they hired the new interns.	Plumbers eventually get used to disgusting odors, but some are just too awful. Louis nearly gagged when he **smelled a rat** in the office which had crawled into a duct and died.
Lexically Unavailable	The department was in an uproar when top secret information was leaked to the press. Jack was not surprised, he had **smelled a mouse** in the office ever since they hired the new interns.	Plumbers eventually get used to disgusting odors, but some are just too awful. Louis nearly gagged when he **smelled a mouse** in the office which had crawled into a duct and died.

Table 13: Example Stimuli for Experiment 4.

Crossing **Bias** and **Lexical Availability** yielded four versions of each target item as shown in Table (13) above. A full list is provided in Appendix IV. Sentences were recorded in the same way as in the previous experiment (and were recorded in the same session as the stimuli for the previous experiment). Additionally, the same set of 60 fillers and the same set of visual stimuli used in the previous study were also used in this experiment.

Looking at the example sentences above we should note that the Idiomatically Biased, Lexically Unavailable trial feels intuitively to me to be somewhat pragmatically odd though by no means ungrammatical. Intuitively this seems to be due to the fact that the contextual bias pushes me to interpret the string as idiomatic, despite the fact that it isn't

an idiom. This was an unfortunate consequence of balancing our conditions, as we were unable to come up with a better version that held other factors constant. Before going forward, we wish to note that this oddness is not restricted merely to this example, but is a general feature of this condition of our stimuli as can be confirmed by examining our stimuli in Appendix IV. Generally, however, the intuitive oddness of these particular items is perhaps exactly the effect that we want (with contextual bias enhancing an impossible reading), though we will need to be somewhat cautious in our analyses as it is difficult to disentangle the intentional pragmatic oddness of the grammatical usage of the expression in an unexpected context from oddness that may be due to difficulty in constructing a coherent narrative of the event depicted by the sentence.

4.2.1.3 Procedure

The procedure was the same as in the experiment presented in Chapter 3: Participants saw four words on the screen and listened to the audio stimuli which contained our idiomatic strings. Because we found effects of task demands, which we interpreted as differences in depth of semantic activation, in the experiment presented in Chapter 3, we also manipulated the task in this experiment. As with the experiments in Chapter 3, task was manipulated as a between-subjects factor with half of our participants reading the four words silently (Experiment 4a) and the other half reading the four words aloud (Experiment 4b) before the onset of the audio stimulus.

4.2.2 Experiment 4a: Read Silently Task

In this section we present the predictions, results and discussion for the within-subjects manipulations of **Bias** and **Lexical Availability** for the **Read Silently** experiment (Experiment 4a). We will then present the predictions, results and discussion for the related **Read Aloud** experiment (Experiment 4b). Within each experiment, I will first present the results for the **Idiomatically Biased** conditions in which the narrative context supports the idiomatic interpretation of the relevant string, and then turn to the **Literally Biased** conditions, where the narrative context supports the literal interpretation. After reviewing these two experiments independently, I will then analyze the two experiments together treating task (Read Silently vs. Read Aloud) as a between-subjects factor in our analysis and provide the general discussion for this chapter.

4.2.2.1 Experiment 4a: Predictions

We are primarily interested in qualitative differences in looking behavior in our different conditions, and on the timing of changes in looking behavior over the course of a trial. Our primary focus will be on the difference between the proportion of looks to the literal target and looks to the idiomatic target, and whether our lexical, contextual and task manipulations have any significant effects on looking behavior. To analyze this statistically we computed Literal Advantage scores, which are the difference between looks to the literal associate and looks to the idiom associate within each display by subjects and items (see Chapter 3).

We predict a strong effect of **Bias** in our results. Recall that in Chapter 2 we found evidence which suggested that while participants may or may not consider the idiomatic interpretation of an ambiguous string depending on context, we did not find corresponding evidence that context modulates consideration of the literal interpretation. While that evidence was based primarily upon participants' processing difficulty in recovering unexpected meaning, in these data we expect to replicate this finding by observing consideration of literal meaning in both idiomatically and literally biased conditions, but observing no consideration (or weak consideration) of idiomatic meaning in the literally biased condition.

Additionally, the magnitude of idiomatic consideration in our Literally Biased condition will help us address the unanswered question from our experiments in Chapter 2 discussed earlier in the motivation for this study. Recall that our measure in Chapter 2 was reaction time, which was not sensitive enough to tell us whether individuals simply did not consider the possible idiomatic interpretation at all in literally biased contexts, or whether activation of the idiomatic meaning was simply much weaker. If, in these data, we find no evidence for idiomatic consideration in the Literally Biased condition then this would support the former view, that context is capable of restricting consideration to the literal interpretation only. If, however, we find some weak consideration of the idiomatic interpretation then this would support the latter view, that the idiomatic interpretation is active but largely suppressed by context.

For our manipulation of **Lexically Availability** we predict, as in Experiment 3, that participants will still consider the idiomatic interpretation even in Lexically Unavailable strings such as *kick the pail*. However we also predict an interaction between **Bias** and **Lexically Availability**. In the Literally Biased condition we predict little or no consideration of the idiomatic interpretation for either Lexically Available (e.g. kick the bucket) or Lexically Unavailable (e.g. kick the pail) conditions. However if it is the case that the idiomatic interpretation is weakly considered even in a literally biased context, then we expect this effect be stronger in the Lexically Available condition than in the Lexically Unavailable condition.

For the Idiomatically Biased condition, however, we predict an overall strong and early consideration of the idiomatic interpretation, as well as some consideration of the literal meaning. However, recall that in Experiment 3 we found an strong initial preference for the idiomatic interpretation for strings like *kick the pail* as compared to *kick the bucket*. We interpreted these effects as a sort of *double-take* reaction. If this view is correct we expect to see a similar pattern of behavior in this experiment with early consideration of the idiomatic interpretation stronger in the Lexically Unavailable condition than in the Lexically Available condition.

4.2.2.2 Experiment 4a: Results

Analyses were conducted in the same way as they were carried out in Experiment 3 in Chapter 3. For each condition we computed Literal Advantage scores (by both subject and item) for the region and each of the eight smaller windows. These scores are

computed by taking the difference between the average looks to the literal target and the average looks to the idiomatic target. A positive value represents more looks to the literal target; a negative value represents more looks to the idiomatic target and a value close to 0 represents equal looks to both the literal and idiomatic targets.

In what follows we will first present explore the behavior in each of our four (Bias X Lexical Availability) conditions. Results examine the difference scores over time, in particular examining whether the Literal Advantage scores deviate significantly from 0.

4.2.2.2.1 Idiomatically Biased Condition

4.2.2.2.1.1 Graph & Basic Discussion

Idiomatically Biased - Lexically Available
(e.g. *kick the bucket*) [Read-Silently]

Figure 22: Looks to the idiom associate word (in blue/dark) vs. looks to the literal associate word (in red/light) as well as looks to the averaged distractor words (thin line in light grey). This figure shows looks in the Idiomatically Biased - Lexically Available condition (e.g. kick the bucket), for the read-silently task.

Idiomatically Biased - Lexically Unavailable
(e.g. *kick the pail*) [Read-Silently]

Figure 23: Looks to the idiom associate word (in blue/dark) vs. looks to the literal associate word (in red/light) as well as looks to the averaged distractor words (thin line in light grey). This figure shows looks in the Idiomatically Biased - Lexically Unavailable condition (e.g. kick the pail), for the read-silently task.

Figures 22 and 23 display the raw proportion of looks for the Idiomatically Biased, read silently trials in the Lexically Available and Lexically Unavailable conditions respectively. In both the Lexically Available and Lexically Unavailable condition we see a greater proportion of fixations to the idiomatic associate word (e.g. *death*) than to the literal associate word (e.g. *foot*) or the averaged distractor words. Additionally this idiomatic advantage appears to be already present before the onset of our critical word, starting shortly after the onset of the verb. This pattern suggests that in these conditions participants are more strongly considering the idiomatic interpretation of our strings early on until around 500ms in both conditions.

152

After 500ms we see that the proportion of looks to our idiom associate word and the proportion of looks to our literal associate word begin to converge, with this competition lasting for the rest of our critical window from 500ms until 1000ms. This suggests that later in the time-course participants are considering both the idiomatic and literal interpretations of these strings.

4.2.2.2.1.2 Overall Statistical Analyses

We examined these patterns statistically using a series of two-tailed t-tests over our Literal Advantage scores for each of the eight 100ms time regions starting at 200ms post noun onset and extending to 1000ms (the full analyses are reported in Table 14). Results of our statistical investigations fail to reveal any significant literal or idiomatic advantage over our regions, and while there is some hint at an early trend toward an Idiomatic Advantage in both conditions, neither result is statistically significant. In essence, this indicates that participants are equally likely to look at the idiom associate (e.g. *death*) and the literal associate (e.g. *foot*) in an idiom-biasing context regardless of whether the idiomatic meaning is lexically available or not.

	df	Idiomatically Biased - Lexically Available (e.g. kick the bucket)			Idiomatically Biased - Lexically Unavailable (e.g. kick the pail)		
		CI	t	p	CI	t	p
Region							
by Subject	15	(-0.178 , 0.122)	-0.398	0.696	(-0.177 , 0.081)	-0.794	0.44
by Items	11	(-0.2 , 0.14)	-0.389	0.705	(-0.215 , 0.083)	-0.976	0.35
200ms-300ms							
by Subject	15	(-0.243 , 0.044)	-1.484	0.159	(-0.227 , 0.082)	-1.003	0.332
by Items	11	(-0.29 , 0.073)	-1.31	0.217	(-0.278 , 0.086)	-1.164	0.269
300ms-400ms							
by Subject	15	(-0.268 , 0.039)	-1.586	0.134	(-0.241 , 0.043)	-1.485	0.158
by Items	11	(-0.286 , 0.073)	-1.306	0.218	(-0.307 , 0.067)	-1.41	0.186
400ms-500ms							
by Subject	15	(-0.25 , 0.092)	-0.983	0.341	(-0.202 , 0.115)	-0.589	0.565
by Items	11	(-0.243 , 0.1)	-0.917	0.379	(-0.308 , 0.13)	-0.897	0.389
500ms-600ms							
by Subject	15	(-0.24 , 0.167)	-0.383	0.707	(-0.194 , 0.128)	-0.439	0.667
by Items	11	(-0.211 , 0.124)	-0.575	0.577	(-0.231 , 0.155)	-0.43	0.675
600ms-700ms							
by Subject	15	(-0.184 , 0.219)	0.186	0.855	(-0.217 , 0.14)	-0.462	0.651
by Items	11	(-0.198 , 0.207)	0.05	0.961	(-0.206 , 0.091)	-0.849	0.414
700ms-800ms							
by Subject	15	(-0.19 , 0.2)	0.055	0.957	(-0.257 , 0.104)	-0.902	0.381
by Items	11	(-0.189 , 0.218)	0.155	0.88	(-0.261 , 0.092)	-1.05	0.316
800ms-900ms							
by Subject	15	(-0.154 , 0.229)	0.419	0.681	(-0.21 , 0.14)	-0.423	0.679
by Items	11	(-0.167 , 0.276)	0.544	0.597	(-0.211 , 0.145)	-0.406	0.693
900ms-1000ms							
by Subject	15	(-0.154 , 0.244)	0.48	0.638	(-0.18 , 0.208)	0.155	0.879
by Items	11	(-0.211 , 0.242)	0.153	0.881	(-0.172 , 0.149)	-0.16	0.876

Table 14: Results of two-way t-tests comparing the literal advantage scores to 0 for our region of interest from 200ms to 1000ms post noun onset and for each of eight 100ms windows in that region.

4.2.2.2.1.3 Interim Discussion

Generally it is somewhat surprising that our difference score analyses failed to reveal any significant Idiom advantage in these trials. The raw proportions do seem to indicate a pattern of early attention on the Idiomatic Target, which should be revealed by a Literal Advantage score which is significantly different from 0 (and in particular is negative). While our overall analyses of our time regions fail to fully support this observation, they do confirm our prediction that participants do consider the literal interpretation of these strings even in idiomatically biased contexts. Thus this data is generally in line with the

results presented in the previous chapter: Participants *do* consider the idiomatic and literal interpretations of ambiguous strings, and appear to do so regardless of whether the string presented is *kick the bucket* or *kick the pail*.

Our analyses also failed to support a qualitative distinction between the Lexically Available and Lexically Unavailable conditions. Recall that we predicted that consideration of the idiomatic meaning should be stronger in the early regions of our time-window for the Lexically Unavailable condition. Given these results this prediction is not confirmed qualitatively, however we will return to this question later when we preform our statistical analysis which quantitatively compares these two conditions.

Recall that these trials are idiomatically biased, and as we mentioned earlier in our materials, that the Idiomatically Biased, Lexically Unavailable trials were somewhat odd. In that sense this competition is even more interesting, as one might predict that the semantic oddness associated with these sentences would induce literal consideration. It appears that this is not the case, with the Idiomatic Bias strongly overpowering any Literal bias that may have been induced by the oddness of these trials.

4.2.2.2.2 Literally Biased Condition

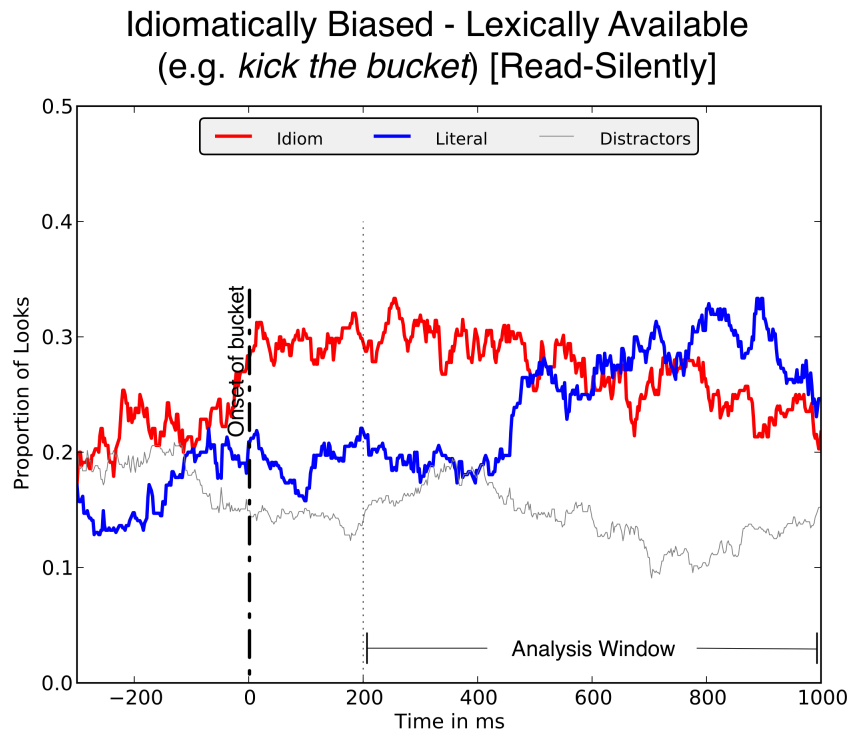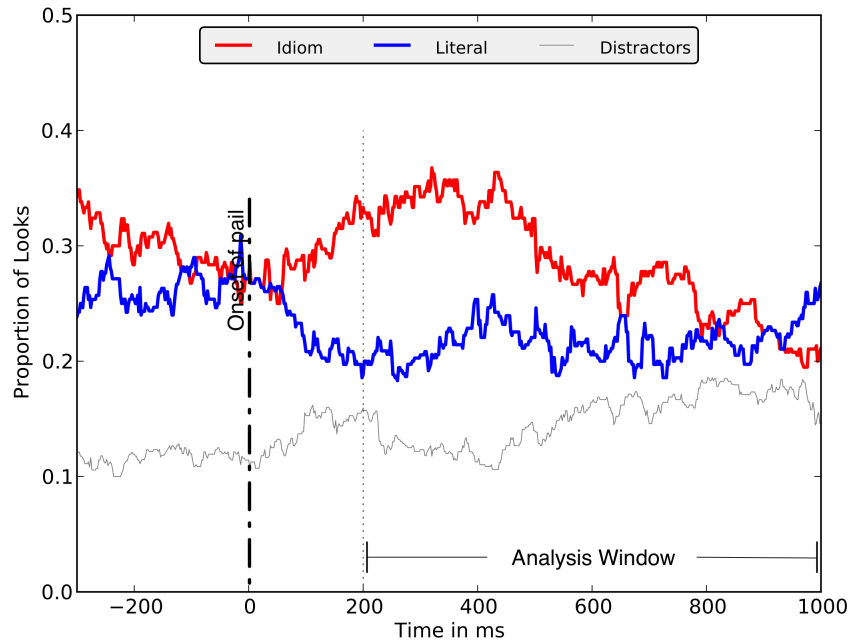*4.2.2.2.2.1 Graph & Basic Discussion copy*
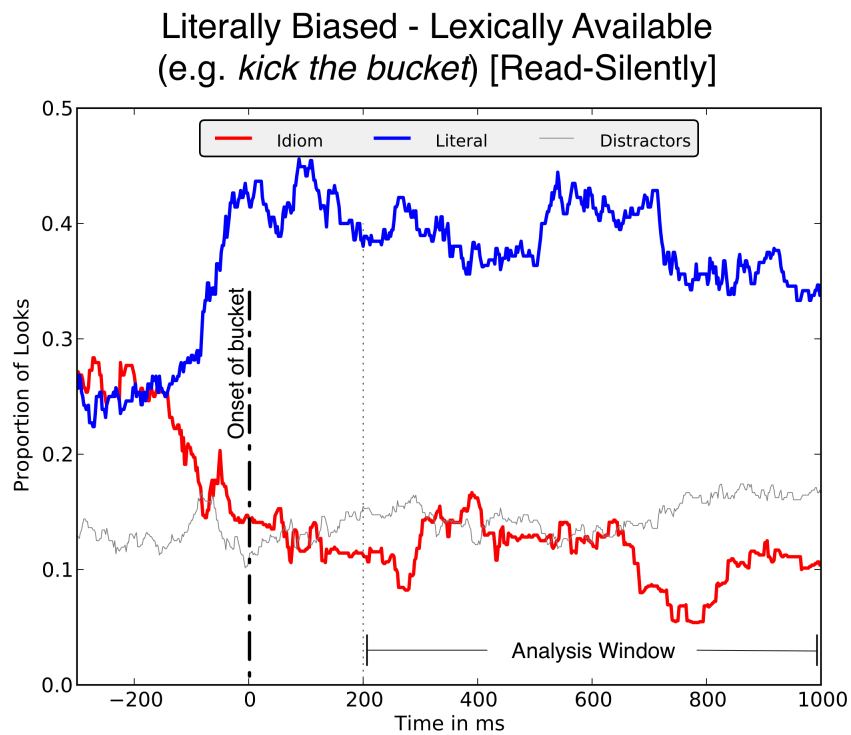
Figure 24: Looks to the idiom associate word (in blue/dark) vs. looks to the literal associate word (in red/light) as well as looks to the averaged distractor words (thin line in light grey). This figure shows looks in the Literally Biased - Lexically Available condition (e.g. kick the bucket), for the read-silently task.
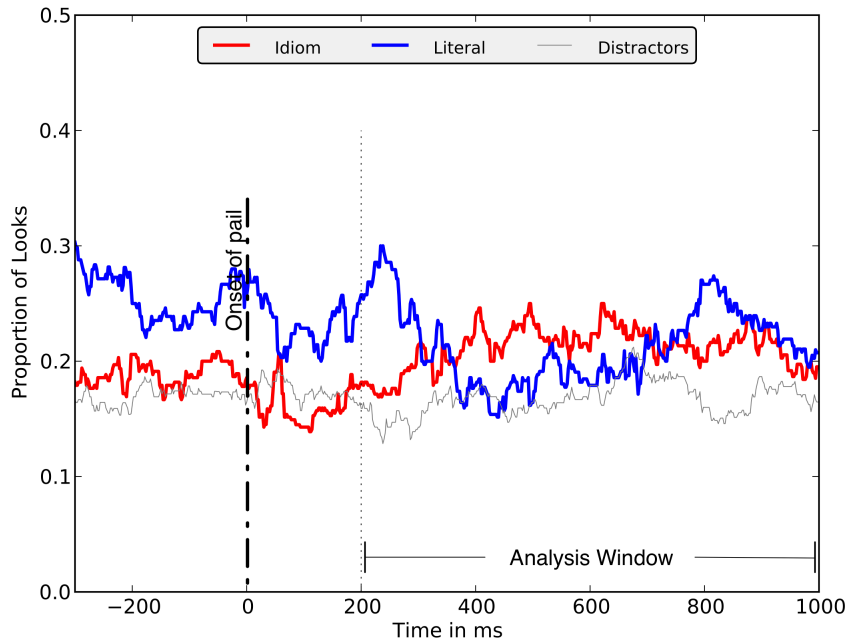
Figure 25: Looks to the idiom associate word (in blue/dark) vs. looks to the literal associate word (in red/light) as well as looks to the averaged distractor words (thin line in light grey). This figure shows looks in the Literally Biased - Lexically Unavailable condition (e.g. kick the pail), for the read-silently task.

Figures 24 and 25 display the raw proportion of looks for the Literally Biased, read silently trials in the lexically available and lexically unavailable conditions respectively. For Lexically Available trials we see much higher proportion of looks to the literal associate word than the idiom associate word or averaged distractor words. This effect begins early just after the onset of the verb and persists for the duration of our time-window. Additionally, during our time-window the proportion of looks to the idiomatic target remains low and similar to looks to the unrelated distractors. This pattern of results suggests strong consideration of the literal interpretation in this condition, with little evidence for consideration of the idiomatic interpretation.

For Lexically Unavailable trials we observe little difference between the proportion of fixations to the literal and idiom associate words, and furthermore little difference between the proportion of looks to the associates and the unrelated distractor items. This pattern suggests low-level competition between the two interpretations as well as general difficulty in forming an interpretation in this condition.

4.2.2.2.2.2 Overall Statistical Analyses
Two-way t-tests over our literal advantage scores fore each of our eight 100ms windows confirm our visual observations (full analyses are presented in Table 15). We find a significant literal advantage in the overall region and in each individual time window for the Lexically Available condition. For the Lexically Unavailable condition we see no significant literal or idiomatic advantage.

Results of two-way t-test analyses on Literal Advantage by Time Window (Read-Silently)

	df	Literally Biased - Lexically Available (e.g. kick the bucket) CI	t	p	Literally Biased - Lexically Unavailable (e.g. kick the pail) CI	t	p
Region							
by Subject	15	(0.082 , 0.386)	3.277	0.005 **	(-0.073 , 0.102)	0.351	0.731
by Items	11	(0.162 , 0.324)	6.591	0 ***	(-0.109 , 0.132)	0.214	0.834
200ms-300ms							
by Subject	15	(0.084 , 0.384)	3.323	0.005 **	(-0.059 , 0.184)	1.103	0.287
by Items	11	(0.108 , 0.38)	3.949	0.002 **	(-0.065 , 0.197)	1.111	0.29
300ms-400ms							
by Subject	15	(0.029 , 0.366)	2.499	0.025 *	(-0.123 , 0.105)	-0.168	0.869
by Items	11	(0.086 , 0.332)	3.743	0.003 **	(-0.169 , 0.121)	-0.365	0.722
400ms-500ms							
by Subject	15	(0.055 , 0.391)	2.835	0.013 *	(-0.144 , 0.09)	-0.493	0.629
by Items	11	(0.124 , 0.331)	4.836	0.001 **	(-0.198 , 0.115)	-0.585	0.57
500ms-600ms							
by Subject	15	(0.031 , 0.444)	2.456	0.027 *	(-0.14 , 0.14)	0.001	0.999
by Items	11	(0.155 , 0.384)	5.191	0 ***	(-0.17 , 0.122)	-0.357	0.728
600ms-700ms							
by Subject	15	(0.084 , 0.444)	3.132	0.007 **	(-0.145 , 0.119)	-0.212	0.835
by Items	11	(0.161 , 0.363)	5.702	0 ***	(-0.184 , 0.146)	-0.253	0.805
700ms-800ms							
by Subject	15	(0.083 , 0.449)	3.105	0.007 **	(-0.1 , 0.162)	0.498	0.626
by Items	11	(0.169 , 0.371)	5.9	0 ***	(-0.122 , 0.226)	0.654	0.527
800ms-900ms							
by Subject	15	(0.051 , 0.414)	2.737	0.015 *	(-0.059 , 0.168)	1.028	0.32
by Items	11	(0.129 , 0.357)	4.687	0.001 **	(-0.095 , 0.207)	0.818	0.431
900ms-1000ms							
by Subject	15	(0.042 , 0.387)	2.652	0.018 *	(-0.095 , 0.127)	0.316	0.757
by Items	11	(0.123 , 0.311)	5.074	0 ***	(-0.087 , 0.143)	0.538	0.601

Table 15: Results of two-way t-tests comparing the literal advantage scores to 0 for our region of interest from 200ms to 1000ms post noun onset and for each of eight 100ms windows in that region

4.2.2.2.2.3 Interim Discussion

These results are qualitatively quite different than the results obtained for the Idiomatically Biased trials presented previously. Whereas we noted competition between the Idiomatic and Literal targets in the Idiomatically Biased conditions, we see something quite different in the Literally Biased conditions. For *kick the bucket* trials we see a strong literal preference, with no hint of idiomatic consideration. This suggests that in this case the Literal Bias is overpowering any consideration of the idiomatic interpretation, suggesting a strong role for Literal Bias. Recall that in the Idiomatically Biased trials,

however, we did not see such strong bias effects. Our prediction was that providing individuals with a literally biasing context might restrict their consideration to the literal interpretation only, while we predicted idiomatically biasing contexts would *not* restrict consideration. These results support this prediction.

We were also interested in whether literal biasing context prevents or merely weakens consideration of the idiomatic interpretation. As we observe no suggestion of idiomatic consideration for the Lexically Available condition with a literally biasing context, our results support the view that participants do not consider the idiomatic interpretation at all in theses cases.

The results of the *kick the pail* trials are less clear. On the one hand our analyses seem to indicate competition between literal and idiomatic interpretations, however looks to the targets, while generally above the distractor baseline, are very low, thus we need to interpret this data cautiously. What we can say is that the behavior for *kick the pail* biased literally is quite different than for *kick the bucket*, which is again a result not found for Idiomatically Biased trials.

4.2.3 Experiment 4b: Read Aloud Task

We now turn to Experiment 4b, in which we examine the results of the **Read Aloud** task. Results and analyses are presented exactly as they were in Experiment 4a. Recall that this experiment is identical to Experiment 4a in every way except that instead of reading the

four visual stimuli silently to themselves, participants were instructed to read each of the four words aloud before pressing a button to start the audio stimuli.

4.2.3.1 Experiment 4b: Predictions

Generally or predictions are the same as for Experiment 4a. Thus we expect to see little or no consideration of the idiomatic interpretation in the Literally Biased condition, and consideration of both the idiomatic and literal interpretations in the Idiomatically Biased condition. Additionally, while we did not observe these patterns in the previous experiment, we still predict that if we do see weak activation of the idiomatic interpretation in the Literally Biased condition it should be *stronger* in the Lexically Available (e.g. kick the bucket) condition than in the Lexically Unavailable (e.g. kick the pail) condition. Finally we also predicted that in the Idiomatically Biased condition we should see stronger initial consideration of the idiomatic interpretation in the Lexically Unavailable condition as compared to the Lexically Available condition, as a result of the *double-take* effect observed in the experiments in Chapter 3.

With regard to the difference in task demands we again predict that the deeper semantic activation associated with speaking the visual targets aloud will increase the magnitude of our effects. This may have the effect of increasing (or decreasing in the case of competition) the Literal Advantage scores in our conditions or of shifting these effects earlier in our time window.

4.2.3.2 Experiment 4b: Results

These results are again rely upon computing literal advantage scores over eight 100ms time windows beginning 200ms post noun onset and extending to 1000ms. Analyses are performed and reported for this experiment in the same way and same order that they were presented in Experiment 4a. As before, we will first examine the results for our Idiomatically Biased conditions and then turn to examine the results for the Literally Biased conditions.

4.2.3.2.1 Idiomatically Biased Trials

4.2.3.2.1.1 Graph & Basic Discussion
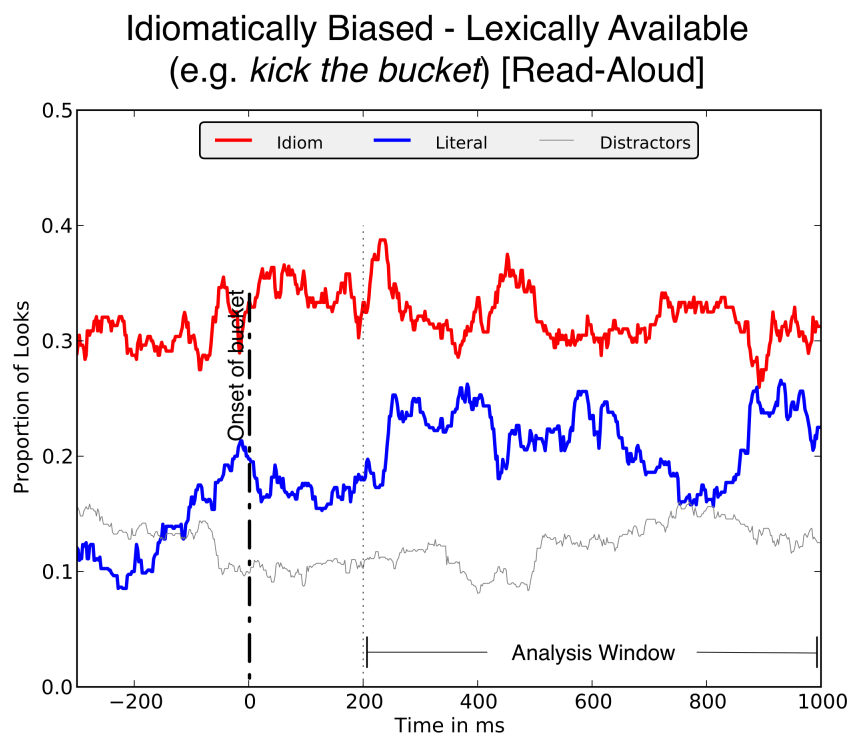
Figure 26: Looks to the idiom associate word (in blue/dark) vs. looks to the literal associate word (in red/light) as well as looks to the averaged distractor words (thin line in light grey). This figure shows looks in the Idiomatically Biased - Lexically Available condition (e.g. kick the pail), for the read-aloud task

Idiomatically Biased - Lexically Unavailable
(e.g. *kick the pail*) [Read-Aloud]
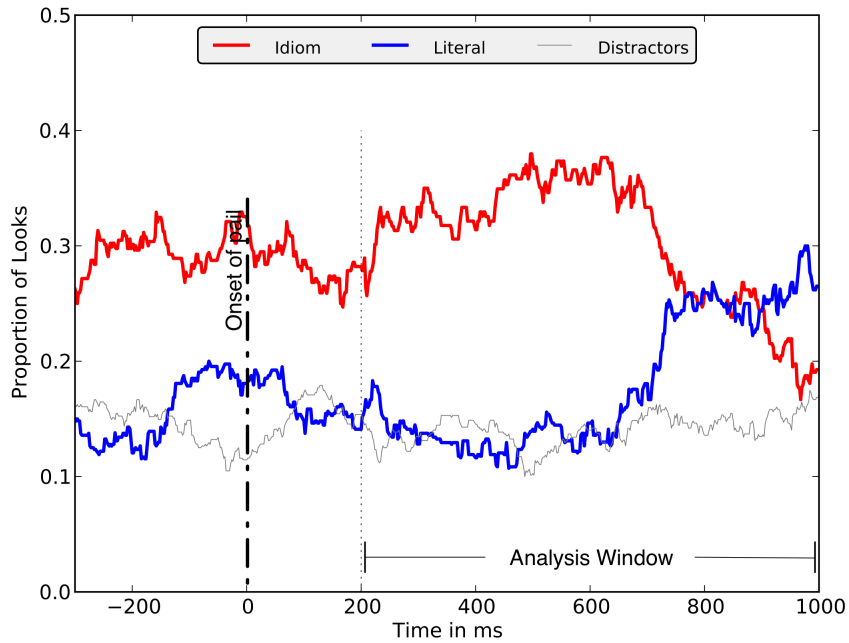
Figure 27: Looks to the idiom associate word (in blue/dark) vs. looks to the literal associate word (in

red/light) as well as looks to the averaged distractor words (thin line in light grey). This figure shows looks

in the Idiomatically Biased - Lexically Unavailable condition (e.g. kick the pail), for the read-aloud task

Figures 26 and 27 display the raw proportion of looks for the Idiomatically Biased, read

aloud trials in the lexically available and lexically unavailable conditions respectively. As

with the results of the read-silently task we a high early proportion of fixations to the

idiom associate word as compared to the proportions of looks to the literal associate word

or averaged distractors in both conditions. Additionally we also see evidence that the

proportions of looks to the idiom associate word and the literal associate word converge

later in the trial.

Looking specifically at the Lexically Available (e.g. *kick the bucket*) condition, we also see that the proportion of looks to the literal associate word remains above the baseline distractors for much of the time window. This pattern of results suggests that participants may be considering both the idiomatic and literal interpretations for the duration of these trials, with consideration of the idiomatic meaning being potentially somewhat stronger.

For the Lexically Unavailable (e.g. *kick the pail*) condition, however we do not see any evidence of the proportion of looks to the literal associate word deviating from the distractor baseline until around 700ms post noun onset (which is actually *delayed* as compared to the read-silently task), where looks to the idiom associate and literal associate converge. This suggests a strong initial bias for interpreting these strings idiomatically, with consideration of their literal meaning arising only later in the time-window.

4.2.3.2.1.2 Overall Statistical Analyses
We examined these patterns statistically using a series of two-tailed t-tests over our Literal Advantage scores for each of the eight 100ms time regions starting at 200ms post noun onset and extending to 1000ms (the full analyses are reported in Table 16). For lexically available trials, our analyses reveal a significant Idiomatic Advantage over the full region (by items only), and significant or marginal Idiomatic advantage over several of our time windows. For lexically unavailable trials we see a marginal idiomatic advantage overall, with the effect focused upon the 200ms - 700ms region, followed by the observed competition.

Results of two-way t-test analyses on Literal Advantage by Time Window (Read-Aloud)

	df	Idiomatically Biased - Lexically Available (e.g. kick the bucket)			Idiomatically Biased - Lexically Unavailable (e.g. kick the pail)		
		CI	t	p	CI	t	p
Region							
by Subject	15	(-0.234 , 0.037)	-1.549	0.142	(-0.232 , -0.008)	-2.291	0.037 *
by Items	11	(-0.209 , -0.008)	-2.369	0.037 *	(-0.251 , 0.014)	-1.966	0.075 .
200ms-300ms							
by Subject	15	(-0.293 , 0.061)	-1.391	0.184	(-0.309 , 0.026)	-1.799	0.092 .
by Items	11	(-0.237 , -0.026)	-2.742	0.019 *	(-0.312 , 0.015)	-1.998	0.071 .
300ms-400ms							
by Subject	15	(-0.221 , 0.087)	-0.924	0.37	(-0.323 , -0.067)	-3.238	0.006 **
by Items	11	(-0.179 , 0.029)	-1.594	0.139	(-0.335 , -0.029)	-2.618	0.024 *
400ms-500ms							
by Subject	15	(-0.254 , 0.042)	-1.532	0.146	(-0.363 , -0.098)	-3.708	0.002 **
by Items	11	(-0.283 , 0.011)	-2.032	0.067 .	(-0.375 , -0.043)	-2.772	0.018 *
500ms-600ms							
by Subject	15	(-0.218 , 0.08)	-0.988	0.339	(-0.351 , -0.095)	-3.705	0.002 **
by Items	11	(-0.245 , 0.074)	-1.183	0.262	(-0.361 , -0.076)	-3.373	0.006 **
600ms-700ms							
by Subject	15	(-0.236 , 0.063)	-1.238	0.235	(-0.318 , -0.06)	-3.121	0.007 **
by Items	11	(-0.211 , 0.026)	-1.725	0.112	(-0.329 , -0.058)	-3.154	0.009 **
700ms-800ms							
by Subject	15	(-0.292 , 0.003)	-2.092	0.054 .	(-0.198 , 0.134)	-0.407	0.689
by Items	11	(-0.282 , -0.035)	-2.817	0.017 *	(-0.22 , 0.115)	-0.69	0.504
800ms-900ms							
by Subject	15	(-0.286 , 0.021)	-1.837	0.086 .	(-0.162 , 0.134)	-0.201	0.843
by Items	11	(-0.252 , 0.028)	-1.764	0.105	(-0.177 , 0.165)	-0.076	0.94
900ms-1000ms							
by Subject	15	(-0.229 , 0.097)	-0.866	0.4	(-0.093 , 0.217)	0.848	0.41
by Items	11	(-0.247 , 0.096)	-0.966	0.355	(-0.084 , 0.212)	0.951	0.362

Table 16: Results of two-way t-tests comparing the literal advantage scores to 0 for our region of interest from 200ms to 1000ms post noun onset and for each of eight 100ms windows in that region

4.2.3.2.1.3 Interim Discussion

When participants are asked to read the visual targets aloud, we find that a context which biases the idiomatic interpretation results in similar behavior to the effects observed when the task is to read the visual targets silently. Our results here also mirror some of the task-difference effects we observed in the Experiment 3 and support our predictions that deeper semantic activation resulting from reading the visual targets aloud will result in stronger looking behavior. Recall that in many cases the effect of the read-aloud task served to magnify the differences seen in the same data in the read-silently condition

from Experiment 3, supporting our predictions regarding the depth of semantic activation in these two tasks. Our results suggest that these task differences observed in experiment 3 are also active in this study.

These results also support our prediction that idiomatic bias will *not* prevent consideration of the literal interpretation of these strings. In both Lexically Available and Lexically Unavailable conditions we observe consideration of the idiomatic and literal interpretations as evidenced by competition. We also observe a general preference for the idiomatic interpretation in early in our time-course for both trials, with some suggestion for a stronger idiomatic advantage in the 200ms to 700ms range for our Lexically Unavailable condition. Recall that in Experiment 3 we found evidence that participants have strong early consideration for the idiomatic interpretation for these strings. We suggested that this result may be due to a sort of *double take* effect, and we see evidence for the same behavioral pattern in these data.
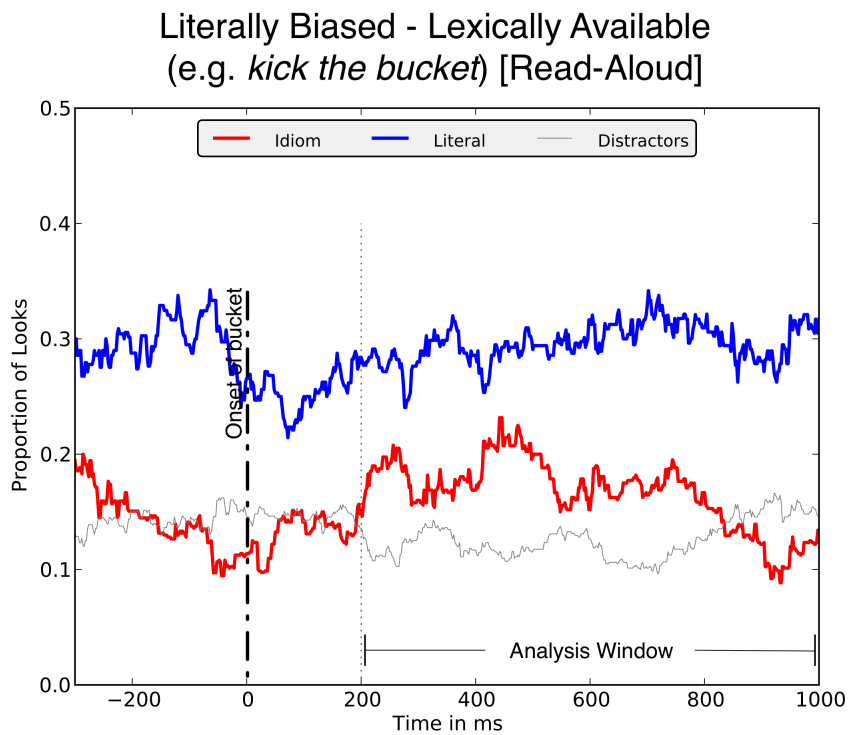
Figure 28: Looks to the idiom associate word (in blue/dark) vs. looks to the literal associate word (in red/light) as well as looks to the averaged distractor words (thin line in light grey). This figure shows looks in the Literally Biased - Lexically Available condition (e.g. kick the bucket), for the read-aloud task

Literally Biased - Lexically Unvailable
(e.g. *kick the pail*) [Read-Aloud]
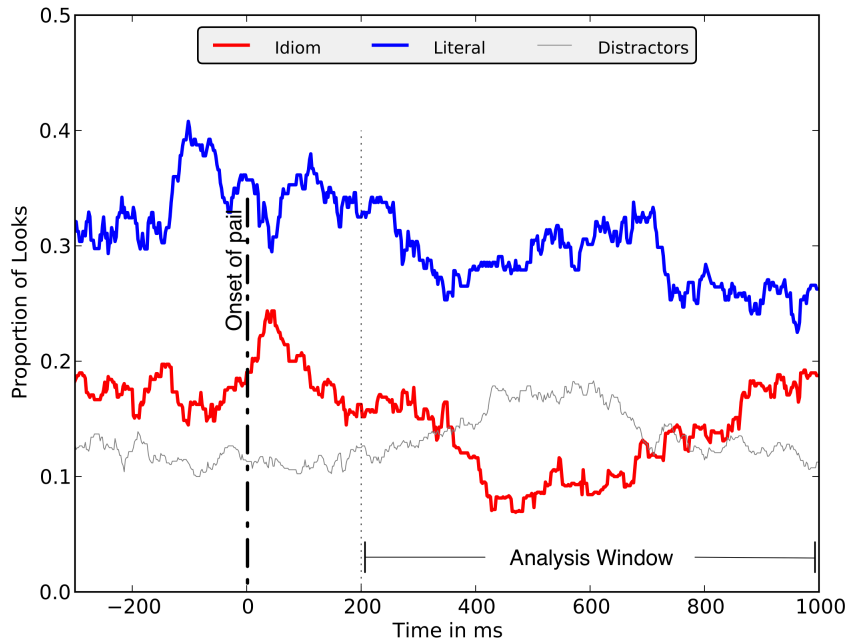
Figure 29: Looks to the idiom associate word (in blue/dark) vs. looks to the literal associate word (in red/light) as well as looks to the averaged distractor words (thin line in light grey). This figure shows looks in the Literally Biased - Lexically Unavailable condition (e.g. kick the pail), for the read-aloud task.

Figures 28 and 29 display the raw proportion of looks for the Literally Biased, read aloud trials in the lexically available and lexically unavailable conditions respectively. In both conditions we see what appears a strong overall preference for the literal associate word over either the idiom associate word or the averaged distractors. This is similar to the results obtained for the read-silently version of the Lexically Available condition.

However unlike the read-silently results we do see evidence for some consideration of the idiomatic interpretation in these conditions. In the Lexically Available condition we see the proportion of looks to the idiom associate deviating weakly from the distractors

starting at the beginning of our time window and remaining above the distractor baseline until around 800ms post noun onset. In the Lexically Unavailable condition we also see some evidence of this trend though it occurs much later, with the looks to the idiomatic associate word appearing to begin to move away from the distractor baseline around 800ms post noun onset.

4.2.3.2.2.2 Overall Statistical Analyses

We examined these patterns statistically using a series of two-tailed t-tests over our Literal Advantage scores for each of the eight 100ms time regions starting at 200ms post noun onset and extending to 1000ms (the full analyses are reported in Table 17). Analyses support our visual observations, revealing a significant literal advantage overall for both Lexically Available and Lexically Unavailable trials. Looking at our individual time-windows reveals the Literal Advantage to be focused in the later time windows for the Lexically Available condition and in the earlier time windows for Lexically Unavailable condition supporting our visual observations that some consideration of the idiomatic interpretation occurs early on in the Lexically Available condition and later in the 800-1000ms range for the Lexically Unavailable condition.

Generally, our analyses reveal an overall Literal Advantage. As compared to the same conditions in the read-silently task two things stand out. First the extreme Literal Advantage we saw in the *kick the bucket* trials seems to have closed somewhat, with the primary literal advantage seeming to occur later in the time-course. Second, the low-level

competition we observed in the *kick the pail* trials has been replaced by a sizable Literal Advantage, focused largely in the earlier time windows.

Results of two-way t-test analyses on Literal Advantage by Time Window (Read-Aloud)

	df	Literally Biased - Lexically Available (e.g. kick the bucket)			Literally Biased - Lexically Unavailable (e.g. kick the pail)		
		CI	t	p	CI	t	p
Region							
by Subject	15	(-0.03 , 0.296)	1.737	0.103	(0.036 , 0.274)	2.782	0.014 *
by Items	11	(0.016 , 0.247)	2.499	0.03 *	(0.045 , 0.268)	3.096	0.01 *
200ms-300ms							
by Subject	15	(-0.118 , 0.293)	0.904	0.38	(-0.019 , 0.345)	1.909	0.076 .
by Items	11	(-0.024 , 0.237)	1.79	0.101	(0.021 , 0.296)	2.531	0.028 *
300ms-400ms							
by Subject	15	(-0.043 , 0.316)	1.618	0.126	(-0.002 , 0.301)	2.108	0.052 .
by Items	11	(-0.05 , 0.303)	1.577	0.143	(0.016 , 0.265)	2.477	0.031 *
400ms-500ms							
by Subject	15	(-0.079 , 0.28)	1.191	0.252	(0.062 , 0.347)	3.066	0.008 **
by Items	11	(-0.062 , 0.214)	1.216	0.249	(0.094 , 0.314)	4.075	0.002 **
500ms-600ms							
by Subject	15	(-0.078 , 0.337)	1.329	0.204	(0.065 , 0.361)	3.062	0.008 **
by Items	11	(0.005 , 0.245)	2.283	0.043 *	(0.102 , 0.314)	4.337	0.001 **
600ms-700ms							
by Subject	15	(-0.093 , 0.325)	1.184	0.255	(0.048 , 0.358)	2.796	0.014 *
by Items	11	(-0.026 , 0.266)	1.802	0.099 .	(0.062 , 0.367)	3.088	0.01 *
700ms-800ms							
by Subject	15	(-0.069 , 0.38)	1.478	0.16	(-0.011 , 0.276)	1.971	0.067 .
by Items	11	(0.003 , 0.27)	2.249	0.046 *	(-0.017 , 0.307)	1.974	0.074 .
800ms-900ms							
by Subject	15	(-0.041 , 0.345)	1.679	0.114	(-0.049 , 0.242)	1.409	0.179
by Items	11	(0.03 , 0.31)	2.669	0.022 *	(-0.059 , 0.273)	1.415	0.185
900ms-1000ms							
by Subject	15	(0.015 , 0.357)	2.317	0.035 *	(-0.047 , 0.203)	1.333	0.202
by Items	11	(0.075 , 0.305)	3.639	0.004 **	(-0.14 , 0.295)	0.782	0.451

Table 17: Results of two-way t-tests comparing the literal advantage scores to 0 for our region of interest from 200ms to 1000ms post noun onset and for each of eight 100ms windows in that region

We are also particularly interested in whether the apparent looks to the idiomatic target word differ significantly from the distractor baseline. To evaluate this statistically we computed Idiom vs. Distractor difference scores by subtracting the average proportion of looks to the idiom target word from the average proportion of looks to our averaged distractors. We then computed two-tailed t-tests over these scores for each of our time windows. With the exception of using Idiom vs. Distractor difference scores this analysis

is identical in procedure to the one reported above for our Literal Advantage Scores. The

results are reported in Table 18.

Results of two-way t-test analyses on Idiom - Distractor Scores by Time Window (Read-Aloud)

	df	Literally Biased - Lexically Available (e.g. kick the bucket) CI	t	p	Literally Biased - Lexically Unavailable (e.g. kick the pail) CI	t	p
Region							
by Subject	15	(-0.047 , 0.132)	1.016	0.326	(-0.085 , 0.067)	-0.258	0.8
by Items	11	(-0.027 , 0.101)	1.266	0.232	(-0.119 , 0.095)	-0.246	0.81
200ms-300ms							
by Subject	15	(-0.037 , 0.192)	1.438	0.171	(-0.067 , 0.154)	0.835	0.417
by Items	11	(-0.083 , 0.182)	0.823	0.428	(-0.081 , 0.172)	0.794	0.444
300ms-400ms							
by Subject	15	(-0.064 , 0.155)	0.885	0.39	(-0.084 , 0.077)	-0.093	0.927
by Items	11	(-0.094 , 0.162)	0.583	0.572	(-0.139 , 0.144)	0.038	0.97
400ms-500ms							
by Subject	15	(-0.054 , 0.216)	1.277	0.221	(-0.157 , 0)	-2.12	0.051 .
by Items	11	(-0.024 , 0.221)	1.774	0.104	(-0.221 , 0.059)	-1.27	0.23
500ms-600ms							
by Subject	15	(-0.081 , 0.174)	0.78	0.447	(-0.168 , -0.001)	-2.159	0.047 *
by Items	11	(-0.017 , 0.106)	1.586	0.141	(-0.204 , 0.039)	-1.496	0.163
600ms-700ms							
by Subject	15	(-0.033 , 0.192)	1.511	0.152	(-0.153 , 0.054)	-1.025	0.322
by Items	11	(0 , 0.144)	2.201	0.05 .	(-0.179 , 0.068)	-0.991	0.343
700ms-800ms							
by Subject	15	(-0.038 , 0.145)	1.244	0.232	(-0.122 , 0.125)	0.022	0.983
by Items	11	(-0.001 , 0.11)	2.155	0.054 .	(-0.138 , 0.125)	-0.11	0.914
800ms-900ms							
by Subject	15	(-0.094 , 0.092)	-0.018	0.986	(-0.075 , 0.146)	0.686	0.503
by Items	11	(-0.074 , 0.057)	-0.295	0.774	(-0.091 , 0.146)	0.516	0.616
900ms-1000ms							
by Subject	15	(-0.108 , 0.022)	-1.408	0.179	(-0.048 , 0.172)	1.206	0.246
by Items	11	(-0.132 , 0.033)	-1.322	0.213	(-0.064 , 0.172)	1.006	0.336

Table 18: Results of two-way t-tests comparing the Idiom vs. Distractor scores to 0 for our region of

interest from 200ms to 1000ms post noun onset and for each of eight 100ms windows in that region.

Generally this analysis fails to find statistically robust differences between the averaged

distractors and the looks to the idiom target word.

4.2.3.2.2.3 Interim Discussion

The results obtained for the Literally Biased condition support our prediction that

participants will overwhelmingly focus upon the literal interpretation of our strings when

given a literally biasing context. Additionally, the effects observed here conform to our

prediction of stronger, more robust looking behavior as a result of the deeper semantic activation due to reading the words aloud when compared to the corresponding results in the read silently task.

We also observe hints of some consideration of the idiomatic interpretation even given the literally biasing context. While we failed to find robust statistical differences between consideration of the idiom and distractors in this case, we can say qualitatively that the observed pattern of looks was not present in the read silently versions of this condition. As predicted, the consideration of the idiomatic interpretation is generally weaker than the consideration of the literal interpretation. Additionally our statistical analyses suggest that this weak idiomatic consideration begins earlier in the Lexically Available condition than in the Lexically Unavailable condition, which supports our predictions regarding possible weak idiomatic consideration in literally biased contexts. Recall that we predicted that if consideration of the idiomatic interpretation occurred it would be weak overall, and further that lexically legitimate idioms such as *kick the bucket* would induce stronger effects than strings such as *kick the pail*. The pattern of results here does suggest that consideration of the idiomatic interpretation is weak overall and further we see evidence for our prediction of strength in that participants consider the idiomatic interpretation earlier for *kick the bucket* than for *kick the pail*.

4.2.4 Condition Comparisons

Now that we have a grasp of the general data patterns present in our conditions we can turn our attention to the full factor analysis of or three manipulations. As in experiment 3, we perform three sets of analyses to further examine our results. The first set of analyses examine the effect of our Bias, Lexical and Task manipulations on looks to the idiom and literal associates separately to examine which of our factors are driving looks to these items and evaluate potential base-line effects at the outset of our analysis window. We then examine our Literal Advantage scores directly using a series of ANOVA over each of our 100ms time windows starting at 200ms post noun onset and extending to 1000ms.

4.2.4.1 Looks to the Literal

Analysis of looks to our literal associate reveals a significant effect of Bias and a significant interaction between Bias, Lexically Availability and Task at the onset of our analysis window, full analyses are reported in Table 19. The baseline effect of Bias is expected given that these sentences are literally biased and at the onset of our analysis window participants have already heard the verb. Additionally this effect is in the expected direction; the positive model estimate translates into more looks to the literal associate word at the onset of our analysis window in the Literally Biased condition.

The interaction indicates that this baseline effect of Bias was modulated by both Lexically Availability and Task. Examining our data reveals a larger baseline preference for the literal associate effect for the Lexically Available condition as compared to Lexically Unavailable condition when biased idiomatically. However for Literally Biased

trials we see a greater preference for the literal associate word in the Lexically Available condition when the visual targets are read-silently, and the opposite effect when they are read aloud, with greater preference for the literal associate word in Lexically Unavailable conditions.

Analysis of Baseline Looks to the Literal associate word

		Estimate	Error	t	p	
Bias						
	Subject	1.07924	0.35757	3.018	0.003	**
	Items	1.00022	0.32514	3.076	0.002	**
LexAvail						
	Subject	0.4114	0.35757	1.151	0.25	
	Items	0.09985	0.32514	0.307	0.759	
Task						
	Subject	-0.28943	0.40148	-0.721	0.471	
	Items	-0.27282	0.32514	-0.839	0.401	
Bias x LexAvail						
	Subject	0.37644	0.71514	0.526	0.599	
	Items	-0.26325	0.65029	-0.405	0.686	
Bias x Task						
	Subject	0.53435	0.71514	0.747	0.455	
	Items	0.59094	0.65029	0.909	0.363	
LexAvail x Task						
	Subject	-0.20058	0.71514	-0.28	0.779	
	Items	-0.48324	0.65029	-0.743	0.457	
Bias x LexAvail x Task						
	Subject	-3.89686	1.43028	-2.725	0.006	**
	Items	-3.70889	1.30057	-2.852	0.004	**

Table 19: Analyses of baseline effects for our looks to the literal associate word over our analysis window.

4.2.4.2 Looks to the Idiom

Analysis of looks to the idiomatic target reveal a significant baseline effect of Bias. The negative model estimate in this case indicates that participants attend to the idiom associate more in the Idiomatically Biased trials, and that this effect is present at the

beginning of our analysis window. As these sentences are biased idiomatically, and participants have by the beginning of our time window already heard the verb this is expected. Essentially, these results indicate that participants are already anticipating an idiomatic reading even before our region of interest.

Analysis of Baseline Looks to the Idiom associate word

		Estimate	Error	t	p	
Bias						
	Subject	-1.31446	0.38121	-3.448	0.001	**
	Items	-1.51257	0.34214	-4.421	0	***
LexAvail						
	Subject	-0.02241	0.38121	-0.059	0.953	
	Items	-0.04652	0.34214	-0.136	0.892	
Task						
	Subject	-0.17567	0.40494	-0.434	0.664	
	Items	-0.04132	0.34214	-0.121	0.904	
Bias x LexAvail						
	Subject	-0.07427	0.76242	-0.097	0.922	
	Items	-0.18068	0.68428	-0.264	0.792	
Bias x Task						
	Subject	-0.45286	0.76242	-0.594	0.553	
	Items	-0.43316	0.68428	-0.633	0.527	
LexAvail x Task						
	Subject	0.20848	0.76242	0.273	0.785	
	Items	0.56821	0.68428	0.83	0.406	
Bias x LexAvail x Task						
	Subject	0.1931	1.52483	0.127	0.899	
	Items	-0.19294	1.36855	-0.141	0.888	

Table 20: Analyses of baseline effects for our looks to the idiom associate word over our analysis window.

4.2.4.3 Literal Advantage Analysis

The results of our analyses of our literal advantage scores by factor is given in table 21. Our results reveal a strong effect of Bias, both over the full region and over each of our individual 100ms time windows. This effect manifests as significantly higher literal

advantage scores in the Literally Biased condition as compared to the Idiomatically Biased condition. We also see a marginal main effect of Lexical Bias and a marginal interaction between Lexical Bias and Task in the overall time window. The main effect of task reveals significantly higher Literal Advantage scores in the *kick the bucket* trials as compared to *kick the pail* trials, and while this is the case generally, the interaction with task reveals this effect to be greater in magnitude in the read-silently condition than the read-aloud condition.

Looking at our individual time windows we see the that the main effect of bias is established throughout the time-course, while the lexical effect is focused upon the 500ms - 700ms region with the interaction between lexical availability and task focused only in the final 100ms of this window. We also see some suggestion of an interaction between bias and task focused in the 400ms - 700ms region and a full interaction focused in the 400ms - 700ms region. The former reveals that while Literal Advantage scores are significantly higher in Literally Biased trials as compared to Idiomatically Biased trials, the magnitude of difference is greater in the read-aloud condition as compared to the read-silently condition. The full interaction further reveals that this result is strongest in Lexically Unavailable, read-aloud trials and weakest in Lexically Unavailable, read-silently trials, with little task difference among the Lexically Available trials over the 400ms −700ms region.

Analysis of Variance of our Factors (Bias x Lexical Availability x Task) over our region and over consequitive 100ms time windows

Region	Bias F	Bias p	Lexical Availability F	Lexical Availability p	Task F	Task p	Bias x Lexical Availability F	Bias x Lexical Availability p	Bias x Task F	Bias x Task p	Lexical Availability x Task F	Lexical Availability x Task p	Bias x Lexical x Task F	Bias x Lexical x Task p
Region														
Subject	26.578	0 ***	2.286	0.159	0.298	0.596	1.587	0.234	2.279	0.159	3.46	0.09 .	1.986	0.186
Items	14.383	0.001 **	2.544	0.121	0.388	0.538	0.899	0.351	0.693	0.412	2.561	0.12	2.207	0.148
200ms-300ms														
Subject	30.879	0 ***	0.325	0.58	0.482	0.502	0.553	0.473	0.049	0.83	1.959	0.189	1.632	0.228
Items	15.073	0.001 **	0.204	0.655	0.385	0.539	0.234	0.632	0.023	0.881	0.881	0.355	2.244	0.145
300ms-400ms														
Subject	20.624	0.001 **	2.118	0.174	0.078	0.785	0.357	0.562	0.972	0.345	0.751	0.405	3.726	0.08 .
Items	17.357	0 ***	2.77	0.106	0.062	0.805	0.184	0.671	0.41	0.527	0.177	0.677	3.75	0.062 .
400ms-500ms														
Subject	22.141	0.001 **	1.032	0.331	0.199	0.664	0.092	0.768	3.81	0.077 .	3.182	0.102	7.079	0.022 *
Items	16.127	0 ***	1.596	0.216	0.333	0.568	0.081	0.778	1.82	0.187	1.098	0.303	6.566	0.016 *
500ms-600ms														
Subject	33.933	0 ***	3.301	0.097 .	0.457	0.513	0.238	0.635	4.552	0.056 .	1.791	0.208	6.058	0.032 *
Items	12.908	0.001 **	1.796	0.19	0.336	0.566	0	0.985	1.55	0.223	0.519	0.477	5.255	0.029 *
600ms-700ms														
Subject	46.406	0 ***	4.732	0.052 .	0.4	0.54	0.017	0.898	4.856	0.05 .	2.987	0.112	3.347	0.095 .
Items	9.43	0.005 **	2.925	0.098 .	0.818	0.373	0.027	0.87	1.305	0.262	2.435	0.129	4.519	0.042 *
700ms-800ms														
Subject	14.66	0.003 **	1.191	0.299	0.628	0.445	1.595	0.233	0.474	0.505	4.011	0.07 .	0.011	0.918
Items	9.342	0.005 **	1.314	0.261	0.258	0.615	1.712	0.201	0.125	0.726	4.235	0.048 *	0.007	0.934
800ms-900ms														
Subject	6.13	0.031 *	1.232	0.291	0.564	0.468	2.032	0.182	0.711	0.417	2.346	0.154	0.13	0.725
Items	6.953	0.013 *	0.938	0.34	0.721	0.403	2.059	0.162	0.184	0.671	2.622	0.116	0.123	0.729
900ms-1000ms														
Subject	5.958	0.033 *	0.952	0.35	0.001	0.979	4.22	0.065 .	0.049	0.828	1.344	0.271	0.3	0.595
Items	2.712	0.11	1.097	0.303	0.018	0.893	5.517	0.026 *	0.13	0.721	1.551	0.223	0.159	0.693

Table 21: ANOVA analyses of the effects of our conditions and their interactions on our region of interest from 200ms to 1000ms post noun onset, as well as on each of eight individual 100ms windows within that region.

4.3 Overall Discussion

In this chapter we presented the results of two studies designed to further investigate the processing and representation of idiomatic expressions. In these experiments we continued the exploration of the relationship between idioms and their literal components begun in Chapter 3, specifically examining whether individuals consider the idiomatic interpretation of semantically related strings such as *kick the pail* when under literal or idiomatic contextual bias. Additionally, we further examined the effect of contextual bias on the processing of idiomatic expressions, hoping to confirm our interpretation of and expand upon the results of Experiment 2.

With regards to the effects of Contextual Bias, we predicted that results would confirm our hypothesis regarding the interpretation of Experiment 2 in Chapter 2. Thus given an idiomatic bias we expected to see consideration of both literal and idiomatic interpretations. However, given a literal bias we expected strong consideration of the literal interpretation only, with no consideration or only weak consideration of the idiomatic interpretation. We were also interested in whether participants would consider the idiomatic interpretation at all during literally biased trials. Recall that in Experiment 2, we hypothesized that literal contextual bias could either restrict interpretation to the literal meaning only, or could strongly inhibit the idiomatic interpretation.

The results of experiment 3 confirm our predictions regarding the pattern of consideration under various contexts. We see consideration of both literal and idiomatic meaning in

idiomatically biased contexts, while in literally biased contexts we see a strong general effect of bias with participants strongly considering only the literal interpretation in these trials. With regards to the question of whether idiomatic consideration is inhibited or prevented under literal bias, our results are mixed. We saw some evidence for the prevention view, with no evidence of idiomatic consideration at all, in our read-silently task. In the read-aloud task, however we saw evidence for the inhibition view, with weak consideration of the idiomatic interpretation. We interpret these findings as supporting the inhibition view over the prevention view.

The reason for this is that, as observed in this experiment as well as in Experiment 3, we interpret the differences between our read-aloud and read-silently task as primarily due to deeper semantic activation in the read-aloud task as compared to the read-silently task. Thus we would predict a stronger semantic activation of the idiomatic representation (and the literal representation) in the read-aloud task. The logic here is that the action of hearing *kick the bucket/pail* and just having read the word *death* (or *foot*) aloud, results in more activation in the lexical system than the act of reading the visual word silently. Thus when participants are biased to interpret the sentence literally, the lexical system utilizes this contextual cue and inhibits activation of the idiomatic interpretation, however the system is more capable of inhibiting activation of the idiomatic meaning *more completely* in the read-silently task due to the overall lower level of activation in that task.

We also predicted to find strong early consideration of the idiomatic interpretation in strings such as *kick the pail*, hence replicating our earlier *double-take* finding from Experiment 3, and generally predicted that both idioms and semantically-related strings should induce similar patterns of behavior with respect to consideration of idiomatic meaning. Generally our results support these predictions. We see evidence of strong early idiomatic consideration in our *kick the bucket* condition, and strong evidence for activation of both the idiomatic and literal interpretations of these strings under idiomatically biased contexts. Even under literally biased contexts, at least in the read-aloud task, we see evidence for weak activation of the idiomatic interpretations for both the idioms and semantically-related strings, with activation being delayed for semantically-related strings like *kick the pail* as compared to unaltered idioms such as *kick the bucket.*

We were also interested in the priority of literal interpretation, and the fact that these data give us a rich time-course of consideration allows us to make some suggestions regarding the process of interpreting idiomatic stings. Generally our results are compatible with a direct access view, in that we see early consideration of idiomatic meaning in our idiom biased trials, and early consideration of literal meaning in our literally biased trials. Thus one possible explanation would be that given the strong effect of contextual bias, participants are simply directly accessing the appropriate interpretation. However such an explanation would leave a great deal of the effects exhibited in this experiment unexplained. It would fail to explain why, for example, participants *do* seem to activate

the literal meaning in Idiomatically Biased trials eventually. Taken with the unbiased results of the experiment in the previous chapter this explanation becomes even more problematic. Recall that in those results we also saw competition between the two interpretations, and indeed in one condition saw an early preference for the Literal Interpretation.

Models such as the Configuration Hypothesis, and the Hybrid Representation Hypothesis are generally taken in the literature to be in support of a *literal priority* view, but closer examination reveals that this is actually an unfair evaluation. Tabossi & Zardon (1993), for example admit that contextual biases may indeed provide sufficient evidence to allow something like a direct access of idiomatic meaning in the Configuration Hypothesis. A more accurate characterization of these theories is that they suppose *literal obligation*. In the case of the Configuration Hypothesis this obligation may only exist in the absence of contextual bias (Tabossi & Zardon, 1993), essentially requiring the parser to process the literal structure of the expression enough to access the *idiom key*. Interpreting this idiom key essentially as sufficient stochastic evidence that the string in question is idiomatic, it may be the case that sufficient contextual evidence is sufficient to license the idiomatic expression even before any literal processing occurs. What the configuration hypothesis does seem to predict is that once the idiom is accessed, there is no further need for literal processing. Thus we might predict that upon discovering the relevant idiom key, the parser might simply stop processing the expression literally.

For hybrid models, the prediction is somewhat different. Whereas the primary mechanism for idiomatic access in the configuration hypothesis is sufficient evidence, which may include contextual information, access to the idiomatic representation in hybrid models requires activation of the literal component lemmas. Thus hybrid models lean more firmly into the *literal priority* camp than the configuration hypothesis. However, while to the super-lemma representation of an idiomatic expression does require activation of the literal component lemmas, there is no requirement that the literal meaning of the string of lemmas be computed *before* idiomatic access is available. Additionally, the hybrid theory makes no claims that context plays no role in idiomatic access. Indeed the structure of the model argues that the same sort of contextual effects that play a role in other domains of lexical access would apply similarly in the case of idioms.

Looking at our results, the more elegant explanation comes from adoption of the hybrid theory, though they are broadly consistent with the configuration hypothesis as well. Both models are compatible with the result that Literal Bias would inhibit access to the idiomatic interpretation, and that idiomatic bias would boost access to the idiomatic interpretation. Likewise both models are compatible with the finding that in Idiomatically Biased trials, the parser has an early preference for the idiomatic interpretation and later also considers the literal interpretation. However, hybrid models can explain this behavior by suggesting that the idiomatic bias boosts the already rapid access profile of the super-lemma representation, and that obligatory and slower literal meaning

computation proceeds normally and eventually catches up resulting in later competition. For the configuration hypothesis, this processing profile is somewhat mysterious, as the explanation would have to claim that the idiomatic bias suffices as an idiom key, permitting access to the idiomatic representation, and then sometime later literal meaning is accessed anyway despite this.

Chapter 5: General Discussion

The goal of this dissertation was to explore the representation and processing of idiomatic expressions. The primary questions we were interested in were (i) how idioms are structurally represented, (ii) how idioms are related to other elements in the lexicon and (iii) how individuals process expressions ambiguous between a literal and idiomatic interpretation. To address these questions we presented the results of three sets of experiments.

The first set (Exps 1 & 2) used a self-paced reading paradigm to examine how individuals process expressions ambiguous between a literal and idiomatic sense when context biases their interpretations toward one of these interpretations, and the processing costs of incorrect expectations. Our second experiment (Exp 3a & 3b) used eye-tracking to examine the structural representation of idioms by asking whether individuals utilize syntactic information on-line when determining the interpretation of ambiguous literal/idiom strings, and also examined whether idiomatic expressions are lexically related to their component words (e.g. *kick* and *bucket*) and through them to their related words (e.g. *pail*). Our final experiment (Exp 4a & 4b), continued the investigation of the relationship between idioms and their component words begun in experiment 3 by asking whether contextual factors influenced looking behavior for semantically associated strings (e.g. *kick the pail*) and provided further information regarding or observations of the effects of context on idiom processing from experiment 2.

Our first question regards the structural representation of idiomatic expressions. As we outlined in Chapter 1 and elsewhere, current thinking about the representation of idiomatic expressions is widely divergent with respect to the structural representations of idioms. Words-with-spaces views (Bobrow & Bell, 1973; Swinney & Cutler, 1979) propose that idiomatic expressions are represented as word-like chunks without detailed syntactic specifications. Decompositional views (Gibbs et al, 1989; Glucksberg, 1993) propose that idioms *do* have structure, but that the interpretation of that structure relies upon specialized components to interpret meaning and hence does not follow typical compositional processing. Hybrid views (Cutting & Bock, 1997; Sprenger et al, 2006) propose that idioms are represented as phrasal units, with detailed but potentially frozen specification of their syntactic configuration.

Generally, the results presented in this dissertation support the hybrid view of idiom representation. We found evidence in experiment 3 that individuals utilize syntactic information on-line to constrain the interpretation of ambiguous idiomatic expressions, and furthermore that this effect is both strong and early. While this behavior could be explained in a words-with-spaces view by claiming that incongruent syntactic context biases individuals to either adopt a literal mode of processing or restrict lexical consideration to the literal interpretation, such an explanation would require a separate mechanism by which the words-with-spaces representations are negatively associated with incompatible representations. While this is possible, the Hybrid view provides a more parsimonious explanation by simply stating that idiomatic representations *are* sets

of possible grammatical functions over their component parts. Our findings are thus most strongly support a structural representation of idiomatic expressions, and provide evidence in favor of the Hybrid view.

We also asked whether idiomatic expressions are related to their component parts, as predicted by the Hybrid view. Recall that under hybrid models, activation of the structural representation of the idiom during comprehension is mediated by the individual literal components of the idiom. We hypothesized that if this was the case, known lexical relationships (such as the semantic relationship between *bucket* and *pail*) should still induce consideration of the idiomatic interpretation despite the fact that strings such as *kick the pail* are unambiguously literal. Our results in experiments 3 & 4 support this conclusion, as we observe activation of the idiomatic interpretation for both unaltered idioms (e.g. *kick the bucket*) and semantically associated strings (e.g. *kick the pail*). Some evidence for this has been shown in production (Sprenger et al, 2007), however our results demonstrate that these effects area also present during comprehension.

Taken together with our findings regarding the structural representation makes Words-with-spaces views of idiomatic representation difficult to maintain. While again, it is theoretically possible to provide a relationship between an idiom such as *kick the bucket*, represented as a big word, and the words which seem to compose it (e.g. *kick* and *bucket*), the process by which this would occur is unclear. Furthermore these relationships would be arbitrary. Thus *kick_the_bucket* could just as easily be related to any other word in the

lexicon. The hybrid view is again more parsimonious, as the proposal is that idiomatic

structural representations are grammatical functions over their component pieces. Thus

the relationship between *kick* and the super-lemma representation of *kick the bucket* is not

arbitrary.

Finally, we also examined the effects of contextual cues on idiomatic processing and the

costs of recovering from incorrect expectations regarding the interpretation of an

ambiguous string. Again current thinking differs in the role of literal processing in the

interpretation of idioms with some models proposing no role of literal processes (Gibbs,

1980; Bobrow & Bell, 1973; Swinney & Cutler, 1979) and others proposing a priority

role for the literal over the idiomatic (Cacciari & Tabossi, 1988; Sprenger et al, 2007).

Our results support the latter view, namely that literal processing appears to be obligatory

when processing idiomatic expressions. In Experiment 2 we found evidence that

individuals are capable of recovering literal meaning faster than idiomatic meaning when

recovering from incorrect expectations induced by contextual bias. We suggested that this

recovery profile suggests that individuals obligatorily consider the literal meaning, and

hence have something to fall back on if their expectations turn out to be incorrect.

In experiment 4 we found similar effects, with evidence for literal processing even when

participants were biased to interpret the relevant string idiomatically. Critically, we did

not find a similar effect under literal bias. When biased literally, participants strongly

favored the literal interpretation, and we observed only weak and later consideration of

the idiomatic meaning. This result further strengthens our interpretation of the results of Experiment 2, and also provided us further details regarding the effects of contextual bias on idiom interpretation. We interpreted the weak activation of the idiomatic interpretation under literal bias as evidence that while contextual bias may serve to significantly inhibit consideration of the idiomatic interpretation; it does not appear to completely prohibit consideration. Taken as a whole, this suggests that the role of context in the interpretation of idioms is one in which the two interpretations compete in parallel with context acting to boost or inhibit consideration of one interpretation.

We also generally found that contextual bias seems to affect the idiomatic interpretation more than the literal interpretation. Recall that our data showed little consideration of the idiomatic interpretation under literal bias, suggesting that bias is strongly inhibiting idiomatic consideration. However we found competition between the idiomatic and literal interpretations under idiomatic bias, suggesting that idiomatic bias does not seem to inhibit literal processing as strongly. We suggest two possible interpretations of these data.

First, these results may simply be a result of the way the parser is structured. Under this interpretation, the parser would obligatorily consider the literal interpretation of any string which it receives as input, and if necessary retrieve the idiomatic interpretation if there is one only later. In this sense the parser may operate as proposed by the Configuration Hypothesis (Cacciari & Tabossi, 1988). The competition between the

literal and idiomatic interpretations under idiomatic bias would then be due to the general

operation of the parser in considering the literal interpretation obligatorily. This approach

has a more difficult time explaining the weak activation of the idiomatic interpretation

under literal bias, however. This could be built into the system by simply saying that the

parser does consider the idiomatic interpretation if it is potentially available, and perhaps

our sentential context was not strong enough to completely prohibit idiomatic

consideration.

Another possibility is that this pattern of behavior is due to the inherent structure of the

lexical representations. Assuming our view of the lexicon as a network of symbolic nodes

we can easily imagine that the activation pattern associated with the literal interpretation

of a string such as *kick the bucket* is necessarily more widely distributed than for the

idiomatic interpretation. Formally this is because the literal representation must recruit

symbolic nodes regarding the individual components, their grammatical properties and

relations and their individual conceptual representations. The idiomatic representation,

however requires only the activation of the component lemmas (*kick* and *bucket*) the

'preformatted' grammatical operations over them (in this case, a verb phrase), and the

single conceptual symbol associated with the idiomatic interpretation.

As a result of this difference in network distribution, we can imagine that idiomatic

representations should be easier/faster to activate as compared to literal representations,

as less nodes require less activation/activation may spread more rapidly. However,

inhibition of idiomatic representations will also be easier/faster for the same reason. Thus we can understand the asymmetry in the effects of contextual bias on the consideration of the idiomatic and literal interpretations as a result of an asymmetry in the distribution of their lexical representations. Our current experiments do not allow us to fully decide between these two views, and thus further exploration of these two possibilities will remain objects for further study.

In summary, this dissertation explored the representation and processing of idiomatic expressions and provided evidence that idioms (i) are represented as structural units, sensitive to syntactic information, (ii) bear principled lexical relationships with their literal component lemmas, and that during processing of these expressions we observed (iii) a priority for obligatory literal processing and (iv) an asymmetry between the effects of contextual bias on literal and idiomatic interpretations. This work also provides evidence in support of the Hybrid Representation Hypothesis as a valid hypothesis for exploring the comprehension of idiomatic expressions.

References

Aitchison, J. (1987). *Words in the mind: An introduction to the mental lexicon*. London: Blackwell.

Allopenna, P., Magnuson, J., & Tanenhaus, M. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*(4), 419-39.

Altmann, G. T. M. (2004). Language-mediated eye movements in the absence of a visual world: the blank screen paradigm. *Cognition, 93*(2), B79-B87.

Altmann, G. T. M., & Kamide, Y. (2009). Discourse-mediation of the mapping between language and the visual world: Eye movements and mental representation. *Cognition, 111*(1), 55-71.

Altmann, G. T. M., Garnham, A., & Henstra, J. A. (1994). Effects of syntax in human sentence parsing: Evidence against a structure-based proposal mechanism. *Journal of Experimental Psychology: Learning, Memory & Cognition, 20*(1), 209-16.

Altmann, G.T.M. & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition, 30*, 191-238.

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language, 62*(1), 67-82.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, Special Issue: Emerging Data Analysis, 59*(4), 390-412.

Barr, D. J. (2008). Analyzing "visual world" eyetracking data using multilevel logistic regression. *Journal of Memory and Language, Special Issue: Emerging Data Analysis, 59*(4), 457-74.

Bierwisch, M. and Schreuder, R. (1992) From lexical concepts to lexical items. *Cognition 42*, 23–60

Bobrow, S., & Bell, S. (1973). On catching on to idiomatic expressions. *Memory and Cognition, 1*(3), 343-46.

Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology 18*(3), 355-87.

Bock, K., & Levelt, W. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics*. San Diego, CA US: Academic Press. 945-84.

Bod, R. (1998). *Beyond grammar: An experience-based theory of language*. Center for the Study of Language and Information.

Boers, F. (2000). Metaphor awareness and vocabulary retention. *Applied Linguistics, 21*(4), 553-71.

Botterell, A. and Stainton, R.J. (2005). Quotation: Compositionality and Innocence without Demonstration. *Crítica, Revista Hispanoamericana de Filosofía 37*(110), 3-33.

Burgess, C., Tanenhaus, M.K., & Seidenberg, M.S. (1989). Nonword interference and lexical ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory & Cognition, 15*, 620-32.

Cacciari C. & Glucksberg, S. (1991). Understanding idiomatic expressions: The contribution of word meanings. In G.B. Simpson (ed.) *Understanding word and sentence*. Amsterdam: North-Holland. 217-40.

Cacciari, Christina, & Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language, 27*(6), 668-83.

Cacciari, Christina, Padovani, R., & Corradini, P. (2007). Exploring the relationship between individuals' speed of processing and their comprehension of spoken idioms. *European Journal of Cognitive Psychology, 19*(3), 417-45.

Caillies, S., & Declercq, C. (2010). Kill the Song-Steal the Show: What Does Distinguish Predicative Metaphors From Decomposable Idioms. *Journal of Psycholinguistic Research*, available http://dx.dio.org/10.1007/s10936-010-9.

Caillies, S., & Le Sourn-Bissaoui, S. (2008). Children's understanding of idioms and theory of mind development. *Developmental Science, 11*(5), 703-11.

Chomksy, N. (1965). *Aspects of the Theory of Syntax*. Cambridge: MIT Press.

Chomsky, N. (1970). *Remarks on Nominalization*. In R. Jacobs and P. Rosenbaum (eds), *Readings in English Transformational Grammar*. Waltham, Massachusetts: Blaisdell Publishing. 184-221.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407-28.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning & Verbal Behavior, 8*(2), 240-7.

Cronk, B. C., Lima, S. D., & Schweigert, W. A. (1993). Idioms in sentences: Effects of frequency, literalness, and familiarity. *Journal of Psycholinguistic Research, 22*(1), 59-82.

Cutting, J.C. & Bock, K. (1997). That's the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends. *Memory & Cognition, 25*(1), 57-71.

Dahan, D. & Tanenhaus, M.K. (2005). Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken-word recognition. *Psychonomic Bulletin & Review, 12*(3) 435-59.

Davies, M. (2011). Google Books (American English) Corpus (155 billion words, 1810-2009). Available online at http:googlebooks.byu.edu/.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production *Psychological Review, 93*(3), 283-321.

Dell, G.S. & O'Seaghdha, P.G. (1991). Mediated and Convergent Lexical Priming in Language Production: A comment on Levelt et al. (1991). *Psychological Review, 98*(4), 604-14.

Elman, J. L. (2009). On the Meaning of Words and Dinosaur Bones: Lexical Knowledge Without a Lexicon. *Cognitive Science, 33*(4), 547–82.

Estill, R.B. & Kemper, S. (1982). Interpreting Idioms. *Journal of Psycholinguistic Research, 11*(6), 559-68.

Ferguson, H. J., & Breheny, R. (2011). Eye movements reveal the time-course of anticipating behavior based on complex, conflicting desires. *Cognition, 119*(2) 179-96.

Ferreira, F. & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25*, 348-68.

Fillmore, C., Kay, P. & O'Connor, M. (1988). Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language, 64*, 501-38.

Fodor, J.A. (2001). Language, Thought and Compositionality. *Mind & Language, 16*(1), 1-15.

Fodor, J.A. (1987). *Psychosemantics*. Cambridge, MA: MIT Press.

Fodor, J., and Pylyshyn, Z. (1988). Connectionism and Cognitive Architecture: a Critical Analysis. *Cognition, 28*: 3–71.

Forster, K. I. (1976). Accessing the Mental Lexicon. In,R.J. Wales & E.C.T. Walker (eds.), *New Approaches to Language Mechanisms*. Amsterdam: North Holland.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(4), 680-98.

Foss, D.J. (1970). Some effects of ambiguity upon sentence comprehension. *Journal of Verbal Learning and Verbal Behavior, 9*, 699-706.

Fraser, B. (1970). Idioms within a Transformational Grammar. *Foundations of Language 6*, 22-42

Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies*. Bloomington, IN: Indiana University Linguistics Club.

Frazier, L., & Rayner, K. (1987). Resolution of syntactic category ambiguities: Eye movements in parsing lexically ambiguous sentences. *Journal of Memory and Language, 26*(5), 505-26.

Frege, G. (1884). *The Foundations of Arithmetic*. Evanston, IL: Northwestern University Press, 1980.

Fromkin, V.A. (1973). *Speech errors as Linguistic Evidence*. The Hague, Netherlands: Mouton.

Garnsey, S., Pearlmutter, N., Myers, E., Lotocky, M. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language, 37*, 58-93.

Garrett, M. (1980). Levels of processing in speech production. In B.Butterworth (ed.), *Language Production: Vol 1. Speech and Talk*. London: Academic Press. 177-220.

Gibbs, R.W., & Gonzales, G.P. (1985). Syntactic frozenness in processing and remembering idioms. *Cognition, 20*(3), 243-259.

Gibbs, R.W., & Nayak, N. (1989). Psycholinguistic studies on the syntactic behavior of idioms. *Cognitive Psychology, 21*, 100-38.

Gibbs, R.W. (1980). Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition, 8*(2), 149-156.

Gibbs, R.W. (1985). On the process of understanding idioms. *Journal of Psycholinguistic Research, 14*(5), 465-472.

Gibbs, R.W. (1993). Why Idioms are not Dead Metaphors. In C. Cacciari & P. Tabossi (eds.) *Idioms: Processing, Structure and Interpretation*. Hillsdale, NJ: Lawrence Erlbaum Associates. 57-78

Gibbs, R.W., Bogdanovich, J., Sykes, J., & Barr, D. (1997). Metaphor in idiom comprehension. *Journal of Memory and Language, 37*(2), 141-154.

Gibbs, R.W. (1984). Literal meaning and psychological theory. *Cognitive Science, 8*(3), 275-304.

Gibbs, R.W., Nayak, N., & Cutting, C. (1989). How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of Memory and Language, 28*(5), 576-593.

Giora, R. & Fein, O. (1999) On understanding familiar and less-familiar language. *Journal of Pragmatics, 31*, 1601-18.

Giora, R. (1999). On the priority of salient meanings: Studies of literal and figurative language. *Journal of Pragmatics, 31*(7), 919-929.

Glucksberg, S. (1993). Idiom meanings and allusional content. In C. Cacciari & P. Tabossi (eds.) *Idioms: Processing, Structure and Interpretation*. Hillsdale, NJ: Lawrence Erlbaum Associates. 3-26

Goldberg, A. (1995). Constructions: A construction grammar approach to argument structure. Chicago UP.

Goldberg, A. (2006). Constructions at work: The nature of generalization in language. Oxford UP.

Halle, Morris & Alec Marantz. (1993). Distributed Morphology and the Pieces of Inflection. In Hale, K. & S.J. Keyser (eds), *The View from Building 20*. Cambridge, MA: MIT Press. 111-176.

Higginbotham, James. (1986). Linguistic theory and Davidson's Program in Semantics. In LePore, E. (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Oxford: Basil Blackwell. 29-48.

Hsieh, S., & Hsu, C. (2010). Idiom Comprehension in Mandarin-Speaking Children. *Journal of Psycholinguistic Research, 39*(6), 505-22.

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition, 96*(1), B23-B32.

Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language, 57*(4), 460-482.

Jackendoff, R. (1995). The boundaries of the lexicon. In M. Everaert, E.-J. VanderLinden, & R. Schreuder (Eds.), *Idioms. Structural and Psychological Perspectives*. Hillsdale, NJ: Erlbaum. 133-65.

Jackendoff, R. (1990) *Semantic Structures*. Cambridge, MA: MIT Press.

Jackendoff, R. (1992). *Languages of the Mind*. Cambridge, MA: MIT Press.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434-46.

Jaeger, T.F. and Snider, N. (2008). Implicit learning and syntactic persistence: Surprisal and Cumulativity. *The 30th Annual Meeting of the Cognitive Science Society* (CogSci08), 827-812.

Jurafsky, D. (2003). Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. In R. Bod, J. Hay & S. Janned (eds.), *Probabilistic Linguistics*. Cambridge MA: MIT Press.

Kaiser, E., Runner, J., Sussman, R., & Tanenhaus, M. (2009). Structural and semantic constraints on the resolution of pronouns and reflexives. *Cognition, 112*, 55-80.

Katz, J. J. (1973). Compositionality, idiomaticity, and lexical substitution. In S. R. Anderson & P. Kiparsky (Eds.), *A festschrift for Morris Halle*. New York: Holt, Rinehart & Winston. 357-76.

Katz, J. J., & Postal, P. (1963). Semantic interpretation of idioms and sentences containing them (Quarterly Progress Report No. 70). Cambridge, MA: MIT Research Laboratory of Electronics. 275- 82

Kempen G, Huijbers P. The lexicalization process in sentence production and naming: Indirect election of words. *Cognition, 14*, 185–209.

Keysar, B., & Bly, B. (1999). Swimming against the current: Do idioms reflect conceptual structure. *Journal of Pragmatics, 31*, 1559-1578.

Kiparsky, P. (1982). Lexical Morphology and Phonology. In I.S. Yang (ed.), *Linguistics in the Morning Calm*. Seoul: Hanshin. 3-91.

Konopka, A. E., & Bock, K. (2009). Lexical or syntactic control of sentence formulation? Structural generalizations from idiom production. *Cognitive Psychology, 58*(1), 68-101.

Kuiper, K., van Egmond, M., Kempen, G., & Sprenger, S. (2007). Slipping on superlemmas: Multi-word lexical items in speech production. *The Mental Lexicon, 2*(3), 313-357.

Lakoff, G. & Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Levelt, W. (1999). Models of word production. *Trends in Cognitive Sciences, 3*(6), 223–32.

Levelt, W.J.M. (2001). Spoken word production: A theory of lexical access. *PNAS. Proceedings of the National Academy of Sciences, 98* (23), 13464-71.

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T. & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review, 98*, 122-42.

Levelt, W.J.M., Roelofs, A. & Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1-75.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition 106*(3), 1126-77.

Lupker, S.J. (1982). The Role of Phonetic and Orthographic Similarity in Picture-Word Interference. *Canadian Journal of Psychology, 36*, 349-67.

MacDonald, M. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language, 32*, 692-715.

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes, 9*, 157-201.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*, 676-703.

Masson, M. E. J., & Bodner, G. E. (2003). A retrospective view of masked priming: Toward a unified account of masked and long-term repetition priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: The state of the art*. New York: Psychology Press. 57-94.

Michel, J., Shen, Y.K., Aiden, A.P., Veres, A., Gray, M. K., The Google Books Team, Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A. & Aiden, E. L. (2011). Quantitative analysis of culture using millions of digitized books. *Science, 331*.

McClelland J. & D. Rummelhart. (1981). An interactive activation model of context effects in letter perception. Part 1: An account of basic findings. *Psychological Review, 88*, 60-94.

McGlone, M. S., Glucksberg, S., & Cacciari, C. (1994). Semantic productivity and idiom comprehension. *Discourse Processes, 17*, 167-190.

Meringer R. & Mayer, K. (1895). *Versprechen und Verlesen*. Goschenscher-Verlag (Reprinted 1978).

Meyer, A. (2005). *Eye-tracking investigations of lexical ambiguity*. Columbia, MO: University of Missouri.

Morton, J. (1970). A functional model for memory. In D.A. Norman (ed.), *Models of Human Memory*. New York: Academic Press.

Nayak, N.P., & Gibbs, R.W. (1990). Conceptual knowledge in the interpretation of idioms. *Journal of Experimental Psychology: General, 119*(3), 315-330.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://www.usf.edu/FreeAssociation/.

Nunberg, G. (1978). *The pragmatics of reference*. Bloomington, IN: Indiana University Linguistics Club.

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

Onifer, W. and Swinney, D. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency-of-meaning and contextual bias. *Memory and Cognition, 9*(3), 225-236.

Ortony, A., Schallert, D., Reynolds, R., & Antos, S. (1978). Interpreting metaphors and idioms: Some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior, 17*(4), 465–477.

Pagin, P, & Dag W. (2010). Compositionality II: Arguments and Problems. *Philosophy Compass, 5*(3), 265-82.

Pawley, A & Syder, F.H. (1983). Two puzzles for linguistic theory: nativelike selection and nativelike fluency, in L Richards and IL Schmidt (eds.), *Language and Communication*. London: Longman. 191-226.

Pelletier, F. J. (1994) The principle of semantic compositionality. *Topoi, 13*, 11–24.

Peterson, R., Burgess, C., Dell, G., & Eberhard, K. (2001). Dissociation between syntactic and semantic processing during idiom comprehension *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(5), 1223-37.

Peterson, R.R., & Savoy, P. (1998). Lexical Selection and Phonological Encoding During Language Production: Evidence for Cascaded Processing. *Journal of Experimental Psychology: Learning, Memory and Cognition 24*(3) 539-557

Pollard, C. & Ivan S. (1994). *Head-Driven Phrase Structure Grammar*. Chicago: Chicago University Press.

Pollio, H.R., Smith, M. K., & Pollio, M.R. (1990). Figurative language and cognitive psychology. *Language and Cognitive Processes, 5*, 141-167.

Popiel, S. J., & McRae, K. (1988). The figurative and literal senses of idioms, or all idioms are not used equally. *Journal of Psycholinguistic Research, 17*(6), 475-87.

Pustejovsky, James. (1995). *The Generative Lexicon*. Cambridge, MA: MIT Press.

Rayner, K., & Duffy, S. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory and Cognition, 14*, 191-201.

Rayner, K., Pacht, J. M., & Duffy, S. A. (1994). Effects of prior en- counter and discourse bias on the processing of lexically ambiguous words: Evidence from eye fixations. *Journal of Memory & Language, 33*, 527-544.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition, 64*, 249-84.

Rumelhart, D. E., & Norman, D. A. (1982). Simulating a skilled typist: A study of skilled cognitive motor performance. *Cognitive Science, 6*, 1-36

Sag, I., Baldwin, T., Bond, F., & Copestake, A. (2002). Multiword expressions: A pain in the neck for NLP. In *Proceedings of the 3rd International Conference on Intelligent Text Processing and Computational Linguistics* (CICLing-2002), 1-15.

Salverda, A., Dahan, D., & McQueen, J. M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition, 90*(1), 51-89.

Schriefers, H., Meyer, A.S. & Levelt, W.J.M. (1990). Exploring the Time Course of Lexical Access in Language Production: Picture-Word Interference Studies. *Journal of Memory and Language, 29*, 86-102.

Schweigert, W.A. (1986). The comprehension of familiar and less familiar idioms. *Journal of Psycholinguistic Research, 15*, 33-45.

Seidenberg, M. S. and McClelland, J. L. (1989). A Distributed, Developmental Model of Word Recognition and Naming. *Psychological Review, 96*, 523-68.

Selkirk, E. (1982). *The Syntax of Words*. Cambridge MA: MIT Press.

Smith, E. E., Shoben, E. J. & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review, 81*(3), 214-41.

Spivey, M.J. & Tanenhaus, M.K. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24*, 1521-43.

Spivey, M.J. Tanenhaus, M.K., Eberhard, K.M. & Sedivy, J.C. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology, 45*, 447-81.

Sprenger, SA, Levelt, W., & Kempen, G. (2006). Lexical access during the production of idiomatic phrases. *Journal of Memory and Language, 54*(2), 161-84.

Stroop, J.R. (1935). Studies of Interference in Serial Verbal Interactions. *Journal of Experimental Psychology, 18*, 643-62.

Swinney, D. (1979). Lexical access during sentence comprehension:(Re) consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*(6), 645-59.

Swinney, D., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior, 18*(5), 523-34.

Swinney, D., & Hakes, D. (1976). Effects of prior context upon lexical access during sentence comprehension. *Journal of Verbal Learning and Verbal Behavior, 15*(6), 681-9.

Tabossi, P, Wolf, K, & Koterle, S. (2009). Idiom syntax: Idiosyncratic or principled Journal of Memory and Language. *Journal of Memory and Language, 61*, 77-96.

Tabossi, P. & Zardon, F. (1993). The activation of Idiomatic Meaning in Spoken Language Comprehension. In C. Cacciari & P. Tabossi (eds.) *Idioms: Processing, Structure and Interpretation*. Hillsdale, NJ: Lawrence Erlbaum Associates. 145-62.

Tabossi, P., Colombo, L., & Job, R. (1987). Accessing lexical ambiguity: Effects of context and dominance. *Psychological Research, 49*, 161–7.

Tabossi, P., Fanari, R., & Wolf, K. (2008). Processing idiomatic expressions: effects of semantic compositionality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(2), 313-27.

Tabossi, P., Fanari, R., & Wolf, K. (2009). Why are idioms recognized fast *Memory & Cognition, 37*(4), 529-40.

Tanenhaus, M. K., Carlson, G., & Trueswell, J. C. (1989). The role of thematic structures in interpretation and parsing. *Language and Cognitive Processes. 4*(3-4), 211-34.

Tanenhaus, M.K. and Seidenberg, M.S. (1989). Context and lexical access: Implications of nonword interference for lexical ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 15(4), 620-32.

Tanenhaus, M.K., Leiman, J.M., and Seidenberg, M.S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior*, 18, 427-40.

Tanenhaus, M.K., Spivey-Knowlton, M.J., Eberhard, K.M. & Sedivy, J.C. (1995). Integration of Visual and Linguistic Information in Spoken Language Comprehension. *Science, 268*(5217), 1632-34.

Tanenhaus, M.K., and Seidenberg, M.S. (1981). Discourse context and sentence perception. *Discourse Processes, 4*, 197-220.

Titone, D, & Connine, C. (1999). On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics, 31*, 1655-74.

Titone, D.A. & Connie, C.M. (1994). The comprehension of idiomatic expressions: Effects of predictability and literality. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20*, 1126-38.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33*, 285 -318.

Trueswell, J.C. & Tanenhaus, M.K. (1991). Tense, temporal context and syntactic ambiguity resolution. *Language and Cognitive Processes, 6*, 339-50.

Trueswell, J.C. & Tanenhaus, M.K. (1994). Toward a lexicalist framework for constraint-based syntactic ambiguity resolution. In Clifton, Frazier & Rayner (Eds.). *Perspectives on Sentence Processing*, 155-179. Hillsdale, NJ: LEA Press.

Van Orden, G.C., Johnston, J.C. & Halle, B.L. (1988). Word identification in reading proceeds from spelling to sound to meaning. Journal of Experimental Psychology: *Learning, Memory & Cognition, 14*, 371-86.

Versace, R. & Nevers, B. (2003). Word frequency effect on repetition priming as a function of prime duration and delay between prime and target. *British Journal of Psychology, 94*, 389-408.

von Fintel, K. (1998). Quantifiers and if-clauses. *The Philosophical Quarterly, 48*(191), 209-214.

Wasow, T., Sag, I.A. & Nunberg, G. (1983). Idioms: an interim report. in S. Hattori & K. Inoue (eds.), *Proceedings of the XIIIth International Congress of Linguistics* (CIPL) Tokyo.

Weinreich, U. (1969). Problems in the analysis of idioms. In J. Puhvel (ed.), *Substance and structure of language*. Berkeley, Los Angeles CA: University of California Press. 23-81.

Yee, E. & Sedivy, J.C. (2006). Eye Movements to Pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 32*(1), 1-14.

Appendix I: Target Stimuli for Experiment 2

Target stimuli for experiment 2. Stimuli are listed in together in 16 quadruplets, one for each of our target stimuli. In each quadruplet the elements (a) and (b) are our Idiomatically Biased stimuli and (c) and (d) are our Literally Biased stimuli. Likewise elements (a) and (c) are our Idiomatically Resolving stimuli and (b) and (d) are our Literally Resolving stimuli.

(1a) The awkward geek, who was very thin, fit in at the society without any difficulties.

(1b) The awkward geek, who was very thin, fit in the hole without any difficulties.

(1c) The tiny contortionist, who was very thin, fit in the hole without any difficulties.

(1d) The tiny contortionist, who was very thin, fit in at the society without any difficulties.

(2a) The foolish entrepreneur, who liked living on the edge, rushed into the decision without a second thought.

(2b) The foolish entrepreneur, who liked living on the edge, rushed into the building without a second thought.

(2c) The brave fireman, who liked living on the edge, rushed into the building without a second thought.

(2d) The brave fireman, who liked living on the edge, rushed into the decision without a second thought.

(3a) The jealous boyfriend, who had stayed up very late, cut in halfway through the dance after getting a drink from the soda machine.

(3b) The jealous boyfriend, who had stayed up very late, cut in the cutting room after getting a drink from the soda machine.

(3c) The talented seamstress, who had stayed up very late, cut in the cutting room after getting a drink from the soda machine.

(3d) The talented seamstress, who had stayed up very late, cut in halfway through the dance after getting a drink from the soda machine.

(4a) The bored intellectual, who had won many contests, dove into the crossword almost immediately.

(4b) The bored intellectual, who had won many contests, dove into the pool almost immediately.

(4c) The muscular swimmer, who had won many contests, dove into the pool almost immediately.

(4d) The muscular swimmer, who had won many contests, dove into the crossword almost immediately.

(5a) The spoiled heir, who was very worried, came into a fortune late last Thursday.

(5b) The spoiled heir, who was very worried, came into a room late last Thursday.

(5c) The late employee, who was very worried, came into a room late last Thursday.

(5d) The late employee, who was very worried, came into a fortune late last Thursday.

(6a) The stubborn father, who was a strict vegetarian, stood by his decision despite other suggestions.

(6b) The stubborn father, who was a strict vegetarian, stood by the tree despite other suggestions.

(6c) The attractive model, who was a strict vegetarian, stood by the tree despite other suggestions.

(6d) The attractive model, who was a strict vegetarian, stood by his decision despite other suggestions.

(7a) The bored student, who hadn't slept properly the night before, drifted off during the lecture despite drinking a cup of coffee earlier in the day.

(7b) The bored student, who hadn't slept properly the night before, drifted off the road despite drinking a cup of coffee earlier in the day.

(7c) The tired driver, who hadn't slept properly the night before, drifted off the road despite drinking a cup of coffee earlier in the day.

(7d) The tired driver, who hadn't slept properly the night before, drifted off during the lecture despite drinking a cup of coffee earlier in the day.

(8a) The bored security guard, who was very strong for his size, flipped through the magazine to make the time go by faster on his lunch break.

(8b) The bored security guard, who was very strong for his size, flipped through the hoop to make the time go by faster on his lunch break.

(8c) The skilled acrobat, who was very strong for his size, flipped through the hoop to make the time go by faster on his lunch break.

(8d) The skilled acrobat, who was very strong for his size, flipped through the magazine to make the time go by faster on his lunch break.

(9a) The clever entrepreneur, who made a lot of money, jumped on the opportunity and was rewarded with a large sum.

(9b) The clever entrepreneur, who made a lot of money, jumped on the moving train and was rewarded with a large sum.

(9c) The daring stuntman, who made a lot of money, jumped on the moving train and was rewarded with a large sum.

(9d) The daring stuntman, who made a lot of money, jumped on the opportunity and was rewarded with a large sum.

(10a) The hungry waitress, who had been working all day, dug into a sandwich just after noon on Sunday.

(10b) The hungry waitress, who had been working all day, dug into a tomb just after noon on Sunday.

(10c) The daring archaeologist, who had been working all day, dug into a tomb just after noon on Sunday.

(10d) The daring archaeologist, who had been working all day, dug into a sandwich just after noon on Sunday.

(11a) The loveable waitress, who was saving up for a car, waited on a customer on a sunny Thursday afternoon.

(11b) The loveable waitress, who was saving up for a car, waited on the bench on a sunny Thursday afternoon.

(11c) The impatient commuter, who was saving up for a car, waited on the bench on a sunny Thursday afternoon.

(11d) The impatient commuter, who was saving up for a car, waited on a customer on a sunny Thursday afternoon.

(12a) The argumentative philosopher, who was wearing a brown jacket, backed down from the argument and admitted that he had made a mistake.

(12b) The argumentative philosopher, who was wearing a brown jacket, backed down the corridor and admitted that he had made a mistake.

(12c) The frightened explorer, who was wearing a brown jacket, backed down the corridor and admitted that he had made a mistake.

(12d) The frightened explorer, who was wearing a brown jacket, backed down from the argument and admitted that he had made a mistake.

(13a) The disappointed athlete, who wanted to be left alone, dwelt on the defeat for many years.

(13b) The disappointed athlete, who wanted to be left alone, dwelt on the mountain for many years.

(13c) The old hermit, who wanted to be left alone, dwelt on the mountain for many years.

(13d) The old hermit, who wanted to be left alone, dwelt on the defeat for many years.

(14a) The spoiled bride, whose husband was in the army, ran up the bill without paying much attention.

(14b) The spoiled bride, whose husband was in the army, ran up the stairs without paying much attention.

(14c) The worried mother, whose husband was in the army, ran up the stairs without paying much attention.

(14d) The worried mother, whose husband was in the army, ran up the bill without paying much attention.

(15a) The understanding professor, who wanted to enjoy himself over spring break, eased off the workload as he didn't want to end up exhausted.

(15b) The understanding professor, who wanted to enjoy himself over spring break, eased off the ledge as he didn't want to end up exhausted.

(15c) The cautious mountain climber, who wanted to enjoy himself over spring break, eased off the ledge as he didn't want to end up exhausted.

(15d) The cautious mountain climber, who wanted to enjoy himself over spring break, eased off the workload as he didn't want to end up exhausted.

(16a) The excellent student, who worked very hard, turned in the assignment and then left the room.

(16b) The excellent student, who worked very hard, turned in the doorway and then left the room.

(16c) The graceful ballerina, who worked very hard, turned in the doorway and then left the room.

(16d) The graceful ballerina, who worked very hard, turned in the assignment and then left the room.

Appendix II: Idioms and Semantic Associates for Exps 3 & 4

Unaltered Idioms	Altered Semantic Associates
kicked the bucket	kicked the pail
tightened his belt	tightened his buckle
smelled a rat	smelled a mouse
found her feet	found her toes
jumped the gun	jumped the rifle
know the ropes	know the cords
pulling his leg	pulling his arm
hit the hay	hit the straw
held her horses	held her ponies
pulling the strings	pulling the threads
spilling the beans	spilling the vegetables
hit the sack	hit the bag

Appendix III: Target Stimuli for Experiment 3

Target items for experiment 3. Stimuli are listed in together in 12 quadruplets, one for each of our target stimuli. In each quadruplet the elements (a) and (b) are our Syntactically Unavailable stimuli and (c) and (d) are our Syntactically Unavailable stimuli. Likewise elements (a) and (c) are our Lexically Available stimuli and (b) and (d) are our Lexically Unavailable stimuli.

(1a) It was surprising to see someone as skilled as John completely miss the ball when he kicked. The bucket full of orange slices was destroyed when he accidentally missed the ball.

(1b) It was surprising to see someone as skilled as John completely miss the ball when he kicked. The pail full of orange slices was destroyed when he accidentally missed the ball.

(1c) Mary kicked the bucket last Thursday evening.

(1d) Mary kicked the pail last Thursday evening.

(2a) Allan gave up trying to untie the knot and just began pulling. The strings were very strong, however, and eventually he had to use scissors.

(2b) Allan gave up trying to untie the knot and just began pulling. The threads were very strong, however, and eventually he had to use scissors.

(2c) Adam was pulling the strings for quite some time.

(2d) Adam was pulling the threads for quite some time.

(3a) The bowl was far too heavy for the young child and the food started spilling. The beans were in a trail on the floor leading from the kitchen to the table.

(3b) The bowl was far too heavy for the young child and the food started spilling. The vegetables were in a trail on the floor leading from the kitchen to the table.

(3c) Erika shouted at Sarah for spilling the beans in the conference room.

(3d) Erika shouted at Sarah for spilling the vegetables in the conference room.

(4a) It didn't matter how hard the kung-fu actor hit. The sack of fake blood just wouldn't burst!

(4b) It didn't matter how hard the kung-fu actor hit. The bag of fake blood just wouldn't burst!

(4c) Philip hit the sack as fast as possible.

(4d) Philip hit the bag as fast as possible.

(5a) When mountain climbing yesterday, John forgot to make sure his harness was tightened. His belt became hooked on a rock outcropping and nearly caused him to fall.

(5b) When mountain climbing yesterday, John forgot to make sure his harness was tightened. His buckle became hooked on a rock outcropping and nearly caused him to fall.

(5c) William tightened his belt sometime last week.

(5d) William tightened his buckle sometime last week.

(6a) Anne entered the cellar where she kept preserved food and noticed that the room smelled. A rat had chewed a hole in the plastic bag and the food inside was rotten.

(6b) Anne entered the cellar where she kept preserved food and noticed that the room smelled. A mouse had chewed a hole in the plastic bag and the food inside was rotten.

(6c) Jimmy smelled a rat in the office.

(6d) Jimmy smelled a mouse in the office.

(7a) After a day of searching the park Jane decided that her lost cell phone simply could not be found. Her feet were sore from walking around all day.

(7b) After a day of searching the park Jane decided that her lost cell phone simply could not be found. Her toes were sore from walking around all day.

(7c) Kathy found her feet after several hours.

(7d) Kathy found her toes after several hours.

(8a) To escape his pursuers the spy ran to the edge of the bridge and jumped. The gun he carried slipped out of his hand and was carried down stream.

(8b) To escape his pursuers the spy ran to the edge of the bridge and jumped. The rifle he carried slipped out of his hand and was carried down stream.

(8c) Valerie jumped the gun several times.

(8d) Valerie jumped the rifle several times.

(9a) How to tie a proper knot is something that every serious sailor should know. The ropes had become loose and the captain berated his crew for nearly half an hour.

(9b) How to tie a proper knot is something that every serious sailor should know. The cords had become loose and the captain berated his crew for nearly half an hour.

(9c) Most new recruits know the ropes by the end of the day.

(9d) Most new recruits know the cords by the end of the day.

(10a) A rescue worker grabbed Bernie's arm and began pulling. His leg was broken in three places, but thanks to the rescue team Bernie escaped before the car caught fire.

(10b) A rescue worker grabbed Bernie's arm and began pulling. His arm was broken in three places, but thanks to the rescue team Bernie escaped before the car caught fire.

(10c) Steven was angry at Kevin for pulling his leg this past Tuesday.

(10d) Steven was angry at Kevin for pulling his arm this past Tuesday.

(11a) The artillerist fired and scored a direct hit. The hay near the wagon immediately caught fire.

(11b) The artillerist fired and scored a direct hit. The straw near the wagon immediately caught fire.

(11c) Billy hit the hay as soon as he could.

(11d) Billy hit the straw as soon as he could.

(12a) Mary put the saddle on slowly and made sure that it held. Her horses were still untamed, so it was best to be extra careful.

(12b) Mary put the saddle on slowly and made sure that it held. Her ponies were still untamed, so it was best to be extra careful.

(12c) Harriet held her horses and remained calm.

(12d) Harriet held her ponies and remained calm.

Appendix IV: Target Stimuli for Experiment 4

Target items for experiment 4. Stimuli are listed in together in 12 quadruplets, one for each of our target stimuli. In each quadruplet the elements (a) and (b) are our Syntactically Unavailable stimuli and (c) and (d) are our Syntactically Unavailable stimuli. Likewise elements (a) and (c) are our Lexically Available stimuli and (b) and (d) are our Lexically Unavailable stimuli.

(1a) Swimming with sharks is a dangerous and unpredictable profession. As a result of the shark attack several oceanographers kicked the bucket last Thursday evening.

(1b) Swimming with sharks is a dangerous and unpredictable profession. As a result of the shark attack several oceanographers kicked the pail last Thursday evening.

(1c) John spent all day filling things with cement as a nasty prank. Several people broke their toes when they kicked the bucket last Thursday evening and may sue.

(1d) John spent all day filling things with cement as a nasty prank. Several people broke their toes when they kicked the pail last Thursday evening and may sue.

(2a) In the current financial climate it is wise to be frugal about money. A top CEO tightened his belt sometime last week as a result of these hard times.

(2b) In the current financial climate it is wise to be frugal about money. A top CEO tightened his buckle sometime last week as a result of these hard times.

(2c) Many people are losing weight with modern exercise programs. Bill tightened his belt sometime last week to keep his pants from falling down.

(2d) Many people are losing weight with modern exercise programs. Bill tightened his buckle sometime last week to keep his pants from falling down.

(3a) The department was in an uproar when top-secret information was leaked to the press. Jack was not surprised, he had smelled a rat in the office ever since they hired the new interns.

(3b) The department was in an uproar when top-secret information was leaked to the press. Jack was not surprised, he had smelled a mouse in the office ever since they hired the new interns.

(3c) Plumbers eventually get used to disgusting odors, but some are just too awful. Louis nearly gagged when he smelled a rat in the office which had crawled into a duct and died.

(3d) Plumbers eventually get used to disgusting odors, but some are just too awful. Louis nearly gagged when he smelled a mouse in the office which had crawled into a duct and died.

(4a) The hardest part about joining the pop group was learning to perform all of the dance moves. Luckily, Kate found her feet after several hours of practice.

(4b) The hardest part about joining the pop group was learning to perform all of the dance moves. Luckily, Kate found her toes after several hours of practice.

(4c) The gas explosion had completely destroyed the mannequin in the display window. The police found her feet after several hours of searching.

(4d) The gas explosion had completely destroyed the mannequin in the display window. The police found her toes after several hours of searching.

(5a) Randy read the paper and immediately regretted selling his stock shares last week. If he hadn't jumped the gun several times he could have made almost twice as much money.

(5b) Randy read the paper and immediately regretted selling his stock shares last week. If he hadn't jumped the rifle several times he could have made almost twice as much money.

(5c) After managing to get the artillery exhibit to loan them several items. The circus watched as the acrobat flipped through the hoop and then easily jumped the gun several times.

(5d) After managing to get the artillery exhibit to loan them several items. The circus watched as the acrobat flipped through the hoop and then easily jumped the rifle several times.

(6a) It takes time and experience to master any task. It was obvious to the more experienced programmers that James didn't know the ropes by the end of the day.

(6b) It takes time and experience to master any task. It was obvious to the more experienced programmers that James didn't know the cords by the end of the day.

(6c) The detective found fibers on the wrists of the victim and sent them to the crime lab. He wanted to know the ropes by the end of the day that were used to bind the victim

(6d) The detective found fibers on the wrists of the victim and sent them to the crime lab. He wanted to know the cords by the end of the day that were used to bind the victim

(7a) Joseph is the most gullible person in the office. On April Fool's Day he didn't realize that Fran was pulling his leg this past Tuesday when she told him about her pet elephant.

(7b) Joseph is the most gullible person in the office. On April Fool's Day he didn't realize that Fran was pulling his arm this past Tuesday when she told him about her pet elephant.

(7c) Attack dogs are trained to drag criminals to the ground. The burglar couldn't remain standing with a 200 pound dog pulling his leg this past Tuesday and was quickly arrested.

(7d) Attack dogs are trained to drag criminals to the ground. The burglar couldn't remain standing with a 200 pound dog pulling his arm this past Tuesday and was quickly arrested.

(8a) After a long day of work it is important to get enough rest. Derek hit the hay as soon as he could after dinner.

(8b) After a long day of work it is important to get enough rest. Derek hit the straw as soon as he could after dinner.

(8c) A good artillerist knows where to aim for maximum effect. Instead of aiming for the barn Zack aimed in front of it hit the hay as soon as he could which immediately caught fire.

(8d) A good artillerist knows where to aim for maximum effect. Instead of aiming for the barn Zack aimed in front of it hit the straw as soon as he could which immediately caught fire.

(9a) Zoe was far too eager to try the difficult ski slope. If she had held her horses and remained calm she might have avoided breaking her arm.

(9b) Zoe was far too eager to try the difficult ski slope. If she had held her ponies and remained calm she might have avoided breaking her arm.

(9c) The amateur riders learned that large animals can be dangerous if not kept calm. Felicia showed them how she held her horses and remained calm at all times to avoid frightening the animals.

(9d) The amateur riders learned that large animals can be dangerous if not kept calm. Felicia showed them how she held her ponies and remained calm at all times to avoid frightening the animals.

(10a) The company president knew surprisingly little about the controversial policy. The investigator suspected that someone else was pulling the strings for quite some time and vowed to find out who it was.

(10b) The company president knew surprisingly little about the controversial policy. The investigator suspected that someone else was pulling the threads for quite some time and vowed to find out who it was.

(10c) Sewing is always difficult when there are cats around. Every time Mandy thought she was making progress she would find that her cat had been pulling the strings for quite some time and unraveling the shirt.

(10d) Sewing is always difficult when there are cats around. Every time Mandy thought she was making progress she would find that her cat had been pulling the threads for quite some time and unraveling the shirt.

(11a) The interrogators were extremely intimidating. John started spilling the beans in the conference room almost as soon as they entered.

(11b) The interrogators were extremely intimidating. John started spilling the vegetables in the conference room almost as soon as they entered.

(11c) A good waiter has excellent balance. Joe walked carefully to avoid spilling the beans in the conference room as he brought them to the table.

(11d) A good waiter has excellent balance. Joe walked carefully to avoid spilling the vegetables in the conference room as he brought them to the table.

(12a) Dwight was completely exhausted when he got home from work today. After he got home, he hit the sack as fast as possible.

(12b) Dwight was completely exhausted when he got home from work today. After he got home, he hit the bag as fast as possible.

(12c) A talented boxer knows that speed is just as important as strength. At boxing practice yesterday Dale hit the sack as fast as possible.

(12d) A talented boxer knows that speed is just as important as strength. At boxing practice yesterday Dale hit the bag as fast as possible.