# Climbing on the Bandwagon of Idiomatic Variation: A Multi-Methodological Approach

by

Kristina Geeraert

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

Doctor of Philosophy

Department of Linguistics

University of Alberta

# Abstract

Idioms have traditionally been regarded as 'frozen' expressions, which are fixed in form. But recent corpus-based research has shown that idioms can occur with a range of variation (cf. Moon, 1998; Barlow, 2000; Duffley, 2013; Schröder, 2013), from lexical variation (e.g. *shake in one's boots/shoes*), to integrated concepts (e.g. *make rapid headway*) to partial or truncated forms (e.g. *the fat lady is warming up*). Few studies however have explored idiomatic variation from an experimental perspective (cf. Gibbs and Nayak, 1989; Gibbs et al., 1989a; McGlone et al., 1994). This dissertation attempts to fill that gap by investigating idiomatic variation using multiple methods. In one study, speakers were asked to rate the acceptability of several types of idiomatic variation to determine speaker preferences for particular variants. These same variants were also presented to participants in an eye-tracking study to determine if certain types of variants are easier to interpret and understand. Finally, speakers were asked to produce idiom variants in an elicitation task specifically designed to encourage creativity. These studies show that some variants, such as integrated concepts (e.g. *pull the political strings*), are produced quite frequently in the elicitation task and judged to be more acceptable in the ratings task, but show significantly longer fixations on the idiom as a whole due to the additional information. Other types of variants however, such as lexical variation (e.g. *tug the strings*), are less preferred and produced less often, but do not show longer reading times than the canonical form. A fourth study collected the transparency ratings of idioms in

their canonical form to determine whether the meaning as a whole influenced variation. Idioms in general are rated as more acceptable if they are also considered to be more transparent, but transparency was not found to be predictive of variation. The results from this study reveal that idioms have a much greater potential for variation than is often assumed. Idioms can be utilized with a considerable range of variation and yet are still interpretable with their idiomatic meanings. This study thus leads to a view of idioms as being not so much different from non-idiomatic or 'literal' language, even if idioms tend to convey semantically richer information.

# Preface

This dissertation is an original work by Kristina Geeraert. The research project, of which this dissertation is a part, received research ethics approval from the University of Alberta Research Ethics Board, Project Name "Investigating Blended Idiomatic Expressions in English", Study ID Pro00036936, on March 4, 2013.

# Acknowledgements

Thank you to everyone who has supported and encouraged me during the completion of this dissertation and throughout the PhD program.

Special recognition goes to my supervisors, Dr. John Newman and Dr. R. Harald Baayen, for all their mentorship, advice, guidance, and support. You were an amazing team and I feel honoured to have been your student. I would also like to thank Dr. Juhani Järvikivi, the third member of my supervisory committee, for his contribution, encouragement, and on-the-ground technical support. In addition, I would like to express gratitude towards my external examiners, Dr. Herbert Colston and Dr. Debra Titone, for participating in the examination and providing feedback. Acknowledgement also goes to Dr. Sally Rice for all her support and mentorship over the years, and for being an 'honorary supervisor'.

Thank you to everyone involved in the Centre for Comparative Psycholinguistics. I would particularly like to acknowledge Dr. Juhani Järvikivi, Director of the CCP, for allowing me to use the facility to collect my data; Lauren Rudat, for all the suggestions to improve the experimental stimuli and for helping me with any snags while running the experiments; and Vincent Porretta, for all his support and assistance both inside and outside the lab.

I would also like to thank the faculty and staff in the Linguistics Department, who were always there to answer any questions or support me in any way they could. Particular recognition goes to Elizabeth French and Lex Bos, who went far beyond administrative support to show great friendship. My fellow graduate students also deserve acknowledgement, for all the support, assistance, and memories. An abundance of gratitude goes specifically to Tamara Sorenson Duncan for her never-ending support and encouragement.

A huge thank you to all my friends who have supported me throughout this degree. In particular, Team Train Wreck (Tammy Borowiecki, Neil and Leah Chatten, Andrew Galloway, Annika Laale, Janeya Manley, and Pattra Winner) for all the laughter, shenanigans, and encouragement every week. And my AMAzing ladies: Chloe Albert, Leanne Stone Dransfeld, and Kalin Jensen, who provided a shoulder to cry on, strength to continue, and an unlimited supply of encouragement and support for which I will always be grateful.

I would also like to extend an enormous thank you to my family, who have been immensely supportive throughout this endeavor. Special recognition goes to my parents and brothers for their unwavering support and encouragement. Acknowledgement also goes to my aunt Frances, who has been a fabulous mentor throughout my education; and to my nana, for all her inspiration and support.

Finally, recognition must also be given to the Social Sciences and Humanities Research Council of Canada and to the Killiam Trust Foundation for their financial support in conducting this research and funding my education.

Thank you!

# Table of Contents

# List of Tables

# List of Figures

# CHAPTER 1

# Introduction

Research on idioms, such as *break the ice* or *kick the bucket*, has predominantly focused on the canonical form, but idioms can occur with a range of variation (e.g. *shatter the ice, no buckets were kicked*), employed for various creative, pragmatic, and discourse effects (cf. McGlone et al., 1994; Duffley, 2013). This dissertation further explores idiomatic variation from an experimental perspective using several methodologies. In what follows, I present an overview of the research on idioms and idiomatic variation before outlining the methods used in this dissertation and the respective goals of each study.

## 1.1 What is an Idiom?

There has been a substantial amount of debate over the last half-century about the notion of an idiom. An idiom was traditionally regarded as a multi-word unit, whose meaning could not be derived from the meanings of its parts (see Gibbs and Colston, 2012, p.50, for a discussion). This meant that the meaning of the expression was arbitrary and semantically opaque, that the form of the expression was frozen or structurally fixed, and that therefore the meanings must be stored separately in the Mental Lexicon, as if a large word, similar to the way word meanings are listed in a dictionary (cf. Sprenger, 2003; Gibbs, 2007; Langlotz, 2006). More recently, scholars have come to view idioms as multi-word units, which are prefabricated and have some degree of collocational restriction or fixedness (cf. Moon, 1998; Wray, 2002; Langlotz, 2006). This has led some scholars (e.g. Moon, 1998) to make a distinction between

'narrow' descriptions of idioms, in which idioms are seen as fixed and semantically opaque multi-word units, and 'broad' descriptions, in which idioms are viewed as multi-word units, regardless of whether they are semantically opaque or not. In fact, some approaches to language, such as Construction Grammar, account for idioms as one type of construction (cf. Goldberg, 2006). In this framework, all levels of form-meaning pairings are considered constructions, especially if some part of the form or function is not fully predictable and if these pairings occurs with sufficient frequency. Constructions then include morphemes (e.g. *-ing*), words (e.g. *kick*), idioms, including partially filled idioms (e.g. *kick the bucket, pull someone's leg*), and even constructional schemas (e.g. passives, ditransitives).

The description of an idiom is also blurred by the fact that idioms are often grouped together with 'fixed expressions' or 'formulaic language' (cf. Moon, 1998; Gibbs, 2007; Taylor, 2012), which can include other classification types, such as collocations (e.g. *foreseeable future*), proverbs (e.g. *a bird in the hand is worth two in the bush*), similes (e.g. *as cool as a cucumber*), and even routine formulae (e.g. *how do you do?*). Several attempts at classifying different types of idioms have also been made. One such example is Taylor (2012), who divides idioms into four types: semantic idioms (i.e. expressions whose meanings cannot be understood from the meaning of the component words, such as *kick the bucket* or *spill the beans*); syntactic idioms (i.e. idioms whose syntactic structure cannot be generated by the general syntactic rules of a language, as in *by and large* or *off with his head!*); lexical idioms (i.e. usage patterns associated with specific lexical items, such as the usage range of *fun*); and phrasal idioms (i.e. word combinations that have an internal structure of a phrase, but without a fully predictable meaning, like *of course* or *in fact*). While constructionist approaches regard all these types as idioms (or constructions), traditional accounts typically focused on semantic idioms, also commonly referred to as 'pure' or 'core' idioms (cf. Moon, 1998; Wulff, 2008). For the purposes of this study, I will also only be focusing on semantic idioms.

Idioms are often discussed along three dimensions. They are regarded as 'conventionalized' or 'institutionalized' formulae, specifically that they are recognized or accepted as multi-lexemic expressions of the language (Fillmore et al., 1988; Moon, 1998; Langlotz, 2006; Wulff, 2008). For Moon (1998, p.7), institutionalization is quantitative, although she acknowledges that some fixed expressions are not very frequent, while others are no longer institutionalized (e.g. *swim between two waters*

'be impartial'). Idioms are also discussed in terms of their lexico-grammatical fixedness, or their formal rigidity (cf. Moon, 1998; Wulff, 2008). Idioms often occur with preferred lexical items and syntactic restrictions for voice, aspect, and mood, often leading scholars to regard these restrictions as formal 'defectiveness' (Moon, 1998, p.7). Finally, the compositionality, analyzability, or isomorphism of idioms is almost always discussed (cf. Nunberg, 1978; Gibbs and Nayak, 1989; Moon, 1998; Langlotz, 2006; Gibbs, 2007; Geeraerts, 2009). These three terms vary slightly, depending on perspective. Compositionality, or rather non-compositionality, is when the word-by-word interpretation does not yield the unitary meaning of the string (Moon, 1998, p.8); this is essentially a bottom-up process. Analyzability on the other hand, is a more top-down analytical process, where the constituents are interpreted to have a distinct meaning which contributes to the meaning of the expression (Langlotz, 2006, p.27). Meanwhile, isomorphism is used for a neutral, non-directional interpretation where a one-to-one correspondence exists between the formal structure and the semantic structure (Geeraerts, 2009, p.89). Isomorphism (also called syntagmatic transparency) is often discussed along with motivation (or paradigmatic transparency), which is the speaker's ability to make sense of or 'motivate' the derived or figurative interpretation.

Finally, scholars tend to agree that these dimensions for understanding idioms are a matter of degree and should be seen as a cline or continuum (cf. Moon, 1998; Langlotz, 2006; Gibbs, 2007; Taylor, 2012). In fact, there has been much debate about the notion of 'idiomaticity', with pure idioms placed at the high end of the idiomaticity continuum (Wulff, 2008). Fillmore et al. (1988) suggest that idiomaticity is a repertory of semantic, morphosyntactic, and pragmatic clusters of information, whereas Nunberg et al. (1994) propose that idiomaticity is a semantic continuum, since syntactic structures and restrictions of a idiom can be motivated through semantics (e.g. abstract concepts tend to be talked about in terms of concrete things). Wulff (2008) however, proposes that the idiomaticity continuum is based on the limitation or restrictions of certain behavioural properties, specifically passivization (i.e. voice and tree-syntactic flexibility), mood and number of the verb, lexico-grammatical flexibility (i.e. the addition of any modifier and the number of adverbs), and lastly the compositionality and frequency of the idiom. She claims that speakers consider these criteria salient enough to assign the construction a position on the continuum. While the degree of idiomaticity certainly varies across

different types of idioms, what idiomaticity is exactly is still unclear.

## 1.2 Idiom Comprehension and Production

The majority of experimental research has focused on the comprehension of idioms, and in particular, how these expressions are understood in comparison with literal language. The processing of language has been thought to happen in a compositional manner, with the literal meaning of each word being understood in sequential order. However, idioms typically do not have components which contribute a literal meaning to the meaning of the whole. Therefore, idioms were considered to be the 'exception to the rule', and were thought to be understood differently. Numerous proposals then have been suggested in order to account for idioms.

Bobrow and Bell (1973) hypothesized that idioms belong to a separate "idiom list", where reference to this list would occur if literal processing failed. But Swinney and Cutler (1979) found that participants were faster at judging idioms (e.g. *break the ice*) as acceptable than they were at judging matched controls (e.g. *break the cup*). They took these results as evidence for the Lexical Representation Hypothesis, which suggests that idioms are stored and accessed as lexical items, and processed in parallel to literal language. Gibbs (1980, 1986) however, found that participants were significantly faster to respond to nonliteral targets, after a nonliteral prime, than to literal targets, after a literal prime. More importantly, participants were also faster at responding to a nonliteral target after a literal prime. These results led Gibbs to conclude that speakers process the idiomatic meaning immediately, even in literal contexts where an idiomatic meaning would not make sense, and therefore proposed the Direct Access Hypothesis.

Since idioms were assumed to be stored whole, studies were conducted to determine the degree of 'lexicalization' (i.e. the process whereby the phrase functions like a single word). Gibbs and Gonzales (1985) claimed that frozen idioms are more lexicalized than more flexible idioms, after finding that participants processed syntactically frozen idioms (e.g. *face the music*) faster than syntactically flexible idioms (e.g. *lay down the law*). Schraw et al. (1988) also found support for lexicalization. They compared familiarity and comprehension ratings for native and non-native speakers, the latter since they were more likely to have not lexicalized

these expressions already. They found that native speakers had high ratings for comprehension regardless of how familiar they were with the expressions, outperforming the non-natives, and therefore suggest that idioms must first be lexicalized in order to understand the figurative meaning. However, Swinney and Cutler (1979) found no support for their frozenness variable, while Schweigert (1986) found a significant effect of familiarity – sentences containing less familiar idioms required more time to read, in both literal and figurative contexts.

In order to determine how speakers know when to access the stored idiomatic meaning, Cacciari and Tabossi (1988) conducted a series of lexical decision experiments, where participants had to decide if a target word was a real word, after listening to a sentence containing an idiom (e.g. *After the excellent performance the tennis player was in seventh heaven*). The target could be semantically-related to the idiomatic meaning (e.g. *happy*), semantically-related to the literal meaning of the last word in the idiom (e.g. *saint*), or semantically anomalous (e.g. *umbrella*). Using highly predictable idioms, participants responded faster to the idiomatic targets when they were presented at the onset of the last word. Only the literal targets were responded to faster if less predictable idioms were used and placed in vague contexts. But if the presentation of the target was changed to 300ms after the onset of the last word, both idiomatic and literal targets were responded to faster than the controls. They conclude that the idiomatic meaning is not always processed immediately (cf. Direct Access Hypothesis), nor is there any indication that the idiomatic meaning is always processed in parallel with the literal meaning (cf. Lexical Representation Hypothesis). Thus, they propose the Configuration Hypothesis – idioms are associated with particular configurations, termed the "idiom key", and once this idiom key (i.e. sufficient input) has been reached then the idiomatic meaning is accessed.

Several follow-up studies have been conducted, further investigating the Configuration Hypothesis. Tabossi and Zardon (1993, 1995) manipulated the location of the presentation of the target. The target could appear on the word before the verb, on the verb, on the first content word after the verb, or on the second content word after the verb. They found that idioms differed in where the idiom key is located: some idioms are activated at the first content word after the verb (e.g. *mind* in *set your mind to rest*), while others are activated at the second content word after the verb (e.g. *head* in *hit the nail on the head*). Meanwhile, Fanari et al. (2010) investigated the role of length and context in activating the idiom key. They found

that the idiomatic meaning was more likely to be activated by the final word if the idiom was longer and if the idiom occurred in a biasing context.

But the literal meaning of the individual words are still accessible during idiom activation, as has been demonstrated through priming studies. Sprenger and colleagues (Sprenger, 2003; Sprenger et al., 2006) found that both identity primes (e.g. *hand*) and words semantically-related to one of the idiom's constituents (e.g. *foot*) lead to faster response times when producing the target idiom (e.g. *get out of hand*). In addition, using the idiom as a prime also leads to faster response times of semantically-related words. These results led the authors to propose the Superlemma Theory, where the idiom's lemma, or superlemma, exists in the lexicon and is connected to the lemmas of the idiom's constituents. However, they claim that the lexical items within this superlemma are fixed (Sprenger et al., 2006) – additional concepts cannot be inserted (e.g. *hit the icy road*) or the lexical items cannot be reversed (e.g. *a wolf in sheep's clothing* vs. *a sheep in wolf's clothing*), which is a major limitation of this theory.

Moreover, Sprenger et al. (2006) claim that syntactic constraints associated with the idiom are stored with this superlemma. Tabossi et al. (2009b) however explored the acceptability of different syntactic constructions for Italian idioms with native and non-native speakers, including general syntactic constraints of the language. They found that syntactic constructions in appropriate pragmatic contexts were reliably preferred over minimal contexts, even with less preferred syntactic constructions, such as passives. For example, bare nouns are not allowed in subject position in Italian as a general rule, and reliably higher ratings were found for idioms which had a full noun in the passive than idioms that had a bare noun. They conclude that idiom-specific syntax is not encoded in the lexicon with the idiom, but rather that idioms behave according to the general syntactic rules of the language.

Another approach to the processing of idioms has been to explore whether some idioms are processed differently than others. The compositionality of idioms has been thought to influence the way in which idioms are understood. Nunberg (1978) proposed a classification system to account for different types of idioms. 'Normally Decomposable Idioms' are idioms whose constituent parts contribute directly to the meaning of the whole, such as *pop the question*. In this idiom, *the question* refers to a specific question (i.e. a marriage proposal) while *pop* refers to the spontaneous act of uttering that proposal. 'Abnormally Decomposable Idioms', like *carry a torch*, are

idioms whose constituent parts contribute indirectly to the meaning of the whole, through a figurative relationship. *Torch* is a conventional metaphor of warm feelings or love for someone, and *carry* is the verb used to refer to feeling in that way. Lastly, 'Nondecomposable Idioms' are idioms whose constituents do not contribute to the meaning of the whole, as in *kick the bucket*. There is no relationship between *kick* or *bucket* and the meaning 'to die'.

Using this classification system, Gibbs et al. (1989b) investigated the role of decomposability in idiom processing. They found that decomposable idioms (both normally and abnormally decomposable) were processed faster than nondecomposable idioms. They therefore proposed the Idiom Decomposition Hypothesis, where literal language processing does not stop once one reaches the idiom. Instead, this literal language processing facilitates the processing of decomposable idioms, whose constituents contribute to the idiomatic meaning, but is not facilitative for nondecomposable idioms, since their meanings do not contribute to the meaning of the whole, causing these idioms to be processed slower. Titone and Connine (1999) investigated decomposable and nondecomposable idioms in an eye-tracking study, placing context before or after the idiomatic expression. They found that nondecomposable idioms were only processed slower if they followed the context, suggesting that decomposability only plays a role when integrating the idiom into the context. These results led the authors to propose the Hybrid Model, which incorporates aspects of the Configuration Hypothesis and the Idiom Decomposition Hypothesis. In the Hybrid Model, the idiomatic meaning is activated when a sufficient portion of the idiom is encountered, but literal processing does not stop once the idiom key is activated allowing this literal processing to facilitate the processing of decomposable idioms.

However, support for the semantic decomposability of idioms has not proved reliable. For example, Tabossi and colleagues (Tabossi et al., 2008, 2009a) have failed to find support for semantic decomposability in several studies. Tabossi et al. (2008) conducted a replication study of Gibbs et al. (1989b), but did not find an effect of decomposability. Meanwhile, Tabossi et al. (2009a) used a semantic judgement task, where participants had to respond to a meaningful sequence, in order to determine if speakers process different types of expressions differently. They found that all three types of expressions (i.e. decomposable idioms, nondecomposable idioms, and cliches) were recognized faster than matched controls, but that there was no difference between the different types of expressions. They claim their findings are further

support for the Configuration Hypothesis and disprove the Idiom Decomposability Hypothesis.

Numerous studies have been conducted on idioms, the majority of which have focused on the comprehension of idioms, each proposing a different theory to account for the varied results. These studies all assume that idioms are stored and accessed whole, and are therefore understood differently from literal language, which is processed compositionally. These studies have also assumed that idioms are simply reproduced whole with little or no variation. But as the studies in the next section demonstrate, idioms can occur with a range of variation.

## 1.3 Idiomatic Variation

### 1.3.1 Corpus Studies

Corpora searches have been an ideal method for exploring idiomatic variation, and in particular, examining how speakers utilize idioms in naturalistic language. A considerable amount of research on idioms and idiomatic variation has been explored through corpus studies, a selection of which will be discussed in some detail here. These studies serve to illustrate the extent with which speakers can modify idioms.

The largest corpus-based study is Moon (1998) who conducted a descriptive analysis of fixed expressions, using the Oxford Hector Pilot Corpus. For her, fixed expressions and idioms (FEIs) are holistic units of two or more words, which can consist of collocations, routine formulae, saying and proverbs, similies, as well as idioms. Among her results, she outlined a range of variation observable with idiomatic expressions. Idioms can occur with lexical variation for verbs (e.g. *fit/fill the bill, throw/toss in the towel, blow up/explode in one's face*), nouns (e.g. *a piece/slice of the action, run rings/circles around someone, miss the boat/bus*), adjectives (e.g. *a bad/rotten apple, scream blue/bloody murder, on a short/tight leash*), and even particles (e.g. *by/in leaps and bounds, rap someone on/over the knuckles, in/into/out of touch*). Variations can include truncations (e.g. *a bird in the hand* [*is worth two in the bush*]), reversals (e.g. *you can't have/eat your cake and eat/have it too*), dialectal variation (e.g. *flog a dead horse* [BrE] vs. *beat a dead horse* [AmE]), and even homophonous or erroneous variations (e.g. *dull as ditchwater/dishwater, to all intents and purposes/intensive purposes*).

Moon (1998, p.122) assumes that FEIs have a fixed, or canonical form, and that variation is derived or deviant from that fixed form. In order to account for variation, such as that shown in Example (1.1) from Moon (1998, p.162), she proposes that similar variants cluster into 'idiom schemas', where they share a similar metaphor in common, as well as common lexical items, but have no fixed structure or lexis. She suggests that compositional FEIs have clearer schemas, whereas non-compositional FEIs have lexicalizations which appear frozen. Variations, truncations, and allusions are therefore understood because they activate the relevant schema. Taylor (2012, p.77) further elaborates on Moon's (1998) example in (1.1), adding other variants like *scare/frighten/terrify the (living) daylights/hell/wits/stuffing* [*out of someone*], and puts forth a more sophisticated account of these idiom schemas, where variants are listed with their relative frequencies, such that *scare* is the most frequent while *terrify* is the least probable.

(1.1) a. scare the life out of SOMEONE
b. scare the shit out of SOMEONE
c. scare SOMEONE shitless
d. scare the pants off SOMEONE
e. frighten the life out of SOMEONE
f. be frightened out of ONE'S mind
g. be scared out of ONE'S wits

Wulff (2008) took a different approach to idiomatic variation, investigating the distributional behaviour of variants, while further exploring the notion of an 'idiomatic continuum'. She collected concordance lines from the BNC for 39 constructions (i.e. 33 V-NP idioms which returned at least 90 hits of the idiomatic meaning and 6 additional V-NP collocational phrases like *write a letter* chosen randomly). She then coded all concordance lines for select syntactic and modifying variations (e.g. tree-syntactic flexibility, such as declarative active/passive or imperative active/passive; lexico-grammatical modifications, as in attributive adjectives or nouns within the NP, relative clauses, number of adverbials; and morphological flexibility like tense, aspect, mood, number and person of the verb). She found that morphological flexibility of the verb slot tends to be most flexible (i.e. tense, aspect, mood, as well as the number and person of the verb), whereas lexico-grammatical flexibility

of the noun slot is least flexible (i.e. modifying attributive noun, or post-modifying relative clause or prepositional phrase).

Schröder (2013) also investigated idioms in the BNC, as well as COCA, and unlike Wulff (2008), found that adjectival/adverbial modification was the most common type of variation, which included attributive nouns. She examined the syntactic flexibility of nine idioms, grouped according to Horn's (2003) classification: fixed idioms (i.e. *kick the bucket, bite the dust, grasp the nettle*), mobile idioms (i.e. *spill the beans, break the ice, bury the hatchet*), and metaphors (i.e. *pull strings, keep tabs on, make headway*). According to this classification, fixed idioms do not allow any syntactic alternations, mobile idioms allow some variation, such as passivization and raising constructions, and metaphors can undergo all syntactic modifications. However, this straight-forward division is not what she found; the selected idioms behaved quite differently than this classification proposed, with great variability within each group. Fixed idioms occurred with adjectival/adverbial modifiers (e.g. *bit the desert dust*) and passives (e.g. *the nettle was grasped*), metaphors predominantly occurred with adjectival/adverbial modification (e.g. *pull political strings, make some headway*), while mobile idioms did not solely occur with the predicted modifications like passives (e.g. *the ice was broken*), but also with other types of alternations like adjectival/adverbial modification (e.g. *spill the many-flavoured beans*).

The following studies are less descriptive accounts of idiomatic variation (compared with the previous studies), and instead use idiom variants to illustrate or advance a particular theory or framework. The first study is Barlow (2000), who examined two idioms (i.e. *it isn't over until the fat lady sings* and *make hay while the sun shines*) in the Corpus of Spoken Professional American English, to demonstrate conceptual blending with idioms. These expressions can be integrated in their full form into a larger context, incorporating a particular subject(s), while linking the action of the idiom with a specific action in the context (e.g. *Long-shot Oscar nominees often try to make hay while the sun shines, lining up as many projects as possible between the announcement of their nomination and probable disappointment on Oscar night*). But conceptual blending can also produce variants of the idiom, integrating additional contextual information into the expression. For example, information about the current situation can be integrated into the subject position, as well as into the temporal phrase (e.g. *The Chiefs aren't the only NFL team that hasn't made hay while their quarterback shined*), or a partial form of the idiom can

be integrated into the larger context, alluding to the process of taking advantage of favourable conditions, and in this particular case accruing political advantage through an unfortunate situation (e.g. *their national leaders angling to make hay out of the tragedy*).

Langlotz (2006, p.175) utilizes corpus data to demonstrate and explain his cognitive-linguistic model of idiom representation, which states that idioms are "complex linguistic routines which are mentally represented as activation sets". In this model, these activation-sets are used as a "symbolic standard to encode a context-specific target-conceptualization" (Langlotz, 2006, p.185). Idiomatic variation, as outlined in this model, is an effect of creativity, where an idiom is manipulated in order to encode a target conceptualization in a communicatively functional and motivated way. He outlines five strategies for creative idiomatic variation. Constructional adaptation is the inflectional and syntactic modifications to the base form (e.g. *with many of its 15 districts 'walking a tightrope'*). Three types of variation manipulate the different layers in an activation set: literal-scene manipulation (i.e. altering the literal scene to encode in the target conceptualization, such as *The chancellor had a narrow tightrope to walk*), topic indication (i.e. modification to the figurative level only, *walk a financial tightrope*), and topic-related literal-scene manipulation (i.e. modification in both the literal and figurative scenes, such as this comment in regards to the making of the movie Jaws: *Bruce, a shark, found it a part he could really sink his three rows of teeth into*). His last variational strategy is ambiguation and punning, which appears to be specific to opaque, non-isomorphic idioms (i.e. nondecomposable idioms), as he describes it as "creative evocation of multiple referentiality" (e.g. *were home and well almost dry*; Langlotz, 2006, p.214).

Duffley (2013) further explores Langlotz's (2006) approach to creative idiomatic variation, looking specifically at two nondecomposable idioms (i.e. *kick the bucket* and *shoot the breeze*) on the internet. He found examples of idiomatic variation with these expressions, reflecting Langlotz's four main strategies: constructional adaptation (e.g. *Most of their buckets have been kicked, Most of the breeze was being shot by one sloppy looking 'veteran' driver*), literal-scene manipulation (e.g. *just shoot the afternoon breeze, reluctant to kick his brimming bucket of life*), topic indication (e.g. *ready to kick its digital bucket, Shoot the cosmic breeze*), and topic-related literal-scene manipulation (e.g. *All we know with certainty is that Titian died in 1576... He may have been over, or under, 90 years old when he kicked the paint*

*can*). Finally, he adds an additional strategy to which speakers can use to employ creativity: poetic or extravagant alternation, which is based on the originality of the speaker as introducing variation for the sake of variety (e.g. *my phone kicked the pail last week, shoot some air with some chums*). These examples illustrate that idiomatic variation can even occur with idioms regarded as opaque and non-compositional. He therefore suggests that as long as the literal scene is analyzable, creative idiomatic variation can still occur. The studies discussed in this section, and the illustrative examples obtained from corpora, have demonstrated the impressive extent with which idioms can be modified.

### 1.3.2 Experimental Studies

Few studies have been conducted on the comprehension or production of idiomatic variation. Gibbs and colleagues (Gibbs et al., 1989a; Gibbs and Nayak, 1989) investigated the comprehension of lexical and syntactic variation using a similarity ratings task. Participants were asked to rate the similarity in meaning between the idiom variant and its literal paraphrase. In Gibbs et al. (1989a), a synonym was used instead of the noun (e.g. *kick the pail*), verb (e.g. *punt the bucket*), or both (e.g. *punt the pail*). In Gibbs and Nayak (1989), the idiom was placed in six different syntactic constructions (e.g. passives, progressives, nominalizations). Both studies found that decomposable idioms were rated as more similar to their literal paraphrases than nondecomposable idioms. They suggested that the forms of decomposable idioms can be manipulated, while retaining the idiomatic meaning, because the constituents contribute meaning to the whole, whereas the idiomatic meaning is lost for nondecomposable idioms since the constituents do not contribute meaning. However, as Duffley (2013) has illustrated, nondecomposable idioms can be modified while retaining their idiomatic meaning. Furthermore, Gibbs and colleagues used semantic similarity ratings as a measure of comprehension, which may not have accurately reflected the comprehension of these variants. Participants may have been rating their approval or preference of the variant and not necessarily how difficult it was to understand the variant. In addition, semantic similarity has been shown to be largely predicted by the same local contexts as observed in corpora (Miller and Charles, 1991), suggesting that the measure collected simply reflected how interchangeable the variant is with its paraphrase.

Replication studies have once again shown inconsistent results in regards to

semantic decomposability. Tabossi et al. (2008) conducted a replication study of Gibbs and Nayak (1989) and found that abnormally decomposable idioms received lower semantic similarity ratings than both normally decomposable and nondecomposable idioms, but that the latter two were not significantly different. Meanwhile, Tabossi et al. (2008) replicated the classification task for grouping idioms into the decomposability categories, following the procedure outlined in Gibbs and Nayak (1989). They calculated the significance of the proportion of agreement between the participant judges using a binomial test and found that agreement on whether an idiom was decomposable or nondecomposable was not significantly different from chance. This non-significant result was also true of the agreement for normally and abnormally decomposable idioms. Titone and Connine (1994b) also found a similar result from the decomposability classification task. In their study, participants reliably classified idioms into one of the three categories 40% of the time, using a 75% agreement rate, in two separate classification tasks.

McGlone et al. (1994) explored the semantic productivity of idiom variation through reading and rating studies. The ratings task asked participants to rate the familiarity of idioms and the comprehensibility of their variants. They found that participants had little difficulty understanding the variants, but did find a correlation between familiarity and comprehensibility – more familiar idioms had variations which were easier to understand. In the self-paced reading study, they measured the reaction time for the participants to read the final sentence in a story, which contained idioms, variants, or literal paraphrases. They found that participants were significantly faster at reading the canonical form of the idiom, but that the variants were read as fast as the literal paraphrases. They suggest that canonical forms of idioms are accessed whole, but that variants are processed like literal language and are therefore processed slower. While the results of this study are certainly interesting, they did not control for the type of variation included. They had instances of lexical variation (e.g. *shatter the ice* for *break the ice*), quantifications (e.g. *not spill a single bean* for *spill the beans*), and even hyperboles (e.g. *it's raining the whole kennel* for *it's raining cats and dogs*). Based on their findings, it is unknown whether certain types of variants are easier than others to understand and interpret.

Studies investigating the production of idiomatic variation are also quite rare. Cutting and Bock (1997) explored the production of idiom blends. They presented two idioms together on a screen, labelled with either 'A' or 'B', and then asked par-

ticipants to produce one of the idioms after a delay. They were attempting to induce errors, or 'slips of the tongue'. More blends were produced if the syntax was similar (syntactic categories in the blends were maintained, such that nouns were replaced by other nouns) and if the semantics were similar (this was true of similar figurative meanings shared with another idiom, as well as similar literal meanings shared with a literal phrase). They did not find an effect of decomposability. These results demonstrate that both the syntax and the semantics (i.e. literal and figurative) are available when accessing and producing idioms.

Meanwhile, Konopka and Bock (2009) attempted to induce structural persistence by using a priming technique to determine whether syntactic criteria specific to idioms are stored in the lexicon or whether they are part of the general syntactic mechanism. They tested phrasal verbs, which can have an idiomatic meaning (e.g. *pull off a robbery*) or a literal one (e.g. *pull off a sweater*). Using rapid-serial-visual-presentation, they presented a sentence to the participants, then asked the participants to do a distracting task. After the distracting task, participants were asked to recall and produce the sentence. They found that idiomatic phrasal verbs were just as likely to be primed as non-idiomatic phrasal verbs (i.e. produced with the same phrasal verb order as the previous prime: verb-particle-object or verb-object-particle). They conclude that the syntax associated with idioms uses general syntactic mechanisms for sentence generation and is therefore separate from the lexicon.

Corpus-based studies have shown that idioms, even nondecomposable idioms, can occur with a range of variation. Idioms are not simply 'fixed expressions', but can be utilized in very novel and creative ways. Experimental research on idiomatic variation however has been quite limited. The comprehension of idiomatic variation either used a ratings task (i.e. judgements) to measure comprehensibility, or considered all variants as a homogeneous set without controlling for the different types of variation. Moreover, production studies of variation attempted to induce slips of the tongue, implying deviations from the canonical form are simply 'errors' and not instances of creative manipulation. This dissertation then attempts to further understand idiomatic variation from an experimental perspective, teasing apart speaker judgements from comprehension, controlling for different types of variants, and encouraging the creativity with which speakers can use idioms.

## 1.4 Current Study

Idiomatic variation has received considerable attention in corpus-based studies. These studies have shown that idioms can occur with a range of variation, including variant types which are rarely acknowledged in the psycholinguistic literature, such as partial forms (or truncations). In fact, some idioms are used to a great extent with variation, such as *make hay* [*while the sun shines*] or *make X headway* (Barlow, 2000; Schröder, 2013). Therefore, in order to understand the role of idioms and how they are organized in language, variation from an experimental perspective needs to be considered further. This dissertation begins to fill that gap, by exploring idiomatic variation using a multi-methodological approach. Four studies have been conducted which investigate idioms and idiomatic variation, each attempting to gain insight into how we use, evaluate, and understand idiomatic language.

The first study explores ratings of transparency for idiomatic expressions. The majority of research on idioms includes a variable for the compositionality of the expression, such as semantic decomposability, which indicates the contribution of each individual word to the meaning of the whole. However, the semantic decomposability classification has proved unreliable. Meanwhile, most scholars recognize that idioms can also be regarded as transparent or opaque, but this measure has yet to be explored in detail. Perhaps the individual contribution of each word is not as pivotal as the role that the words in combination play. This study then is two-fold: to determine which factors influence the perceived transparency of an idiom, as well as to collect a measure of transparency for inclusion in the analyses of the following three studies specific to idiomatic variation.

The second study explores speakers' acceptability of several types of idiomatic variation. Previously conducted rating studies aimed to elicit measures of comprehensibility (Gibbs and Nayak, 1989; Gibbs et al., 1989a). This study however, simply intends to elicit speakers' acceptability or preferences of different types of variation. Four types of variation were included in this experiment: lexical variation (e.g. *pull your limb*), integrated concepts (e.g. *pull your gullible leg*), partial or truncated forms (e.g. *pull it*), and formal idiom blends (e.g. *pull your goat*), along with the canonical form (e.g. *pull your leg*) and a literal meaning of the idiom. These variants were placed in contextually appropriate sentences, with the context identical for the canonical form and the variants, so that only the form of the idiom varied between

these conditions. A different context was provided for the literal meaning of the idiom, which included the idiom in its canonical form. This study then is three-fold: to determine which factors influence speakers' acceptability of variation, to ascertain which variant types are more acceptable, and to provide a measure of acceptability for inclusion in the analyses of the following study.

The third study explores idiomatic variation through an eye-tracking experiment. Using the same stimuli as in the previous experiment, these variants were presented to participants in an eye-tracking study to measure participants' fixation durations, and number of fixations, to determine which variants are more difficult to understand. Previous research found that variants take longer to comprehend than the canonical form, but did not control for the type of variant included (cf. McGlone et al., 1994). This study expands on this previous research by investigating several types of variants, which will be compared to the canonical form, and to each other, to determine if all variant types are more difficult to understand than the canonical form, as well as to determine whether all types of variants are equally as difficult.

Finally, the fourth study investigates the spontaneous and conscious production of idiomatic variation through an elicitation task. Participants were provided with short newspaper snippets and idioms and asked to create a headline for each snippet using the provided idiom. The results are analyzed quantitatively to determine which variables are influential in employing variation. Seven select idioms are also analyzed more qualitatively to illustrate the range of variation and the specific variational patterns which can be produced with a particular idiom. In addition, the headlines for these seven idioms will be compared to concordance lines extracted from corpora, to determine whether the variations elicited in this study converge with findings from corpora, or whether they reveal new uses and variants.

The final discussion will bring together the main results from the four experiments included in this dissertation, attempting to draw generalizations about each variant, while integrating the results into the larger discussion on idioms in the literature. I will then present a more recent approach to language, Implicit Grammar, and discuss how idioms and idiomatic variation might be better explained within this framework.

# CHAPTER 2

# Transparency of Idiomatic Expressions

## 2.1 Introduction

Decomposability, or the compositionality of the expression, is often discussed in the literature on idioms, specifically the extent to which the individual constituents of the idiom contribute to the overall meaning of the whole. One classification system that has been proposed for the different types of idioms and the degree to which the constituents contribute meaning is that of Nunberg (1978). In this classification scheme, 'Normally Decomposable Idioms' have constituents which directly contribute their literal meaning to overall meaning of the whole. The most often cited example of a normally decomposable idiom is *pop the question*; that is, *the question* is a specific question, a marriage proposal, while *pop*, 'suddenly ask', refers to the verb used in uttering that proposal. 'Abnormally Decomposable Idioms' have constituents which contribute meaning to the whole through an indirect or figurative relationship, such as *carry a torch*. The *torch* is a conventional metaphor used to convey warm feelings of love for someone, while *carry* is the verb used to describe the act of feeling in that way. Finally, 'Nondecomposable Idioms' have constituents which do not contribute to the meaning of the whole, as in *kick the bucket* 'to die'. The individual words in this idiom do not literally or metaphorically contribute meaning to the expression; there is no relationship between *kick* or *bucket* and 'die'.

Some support has been found for the role that decomposability plays in idiom processing, building upon Nunberg's (1978) classification. Gibbs and colleagues (Gibbs and Nayak, 1989; Gibbs et al., 1989a) found that decomposable idioms were preferred with syntactic and lexical variation. They attributed this finding to the

fact that decomposable idioms have constituents which independently contribute meaning to the whole; therefore, the meaning is still retrievable even when the form is manipulated. This differs from nondecomposable idioms, which cannot be manipulated in form and still maintain their idiomatic meaning. Moreover, Gibbs et al. (1989b) found that decomposable idioms are processed faster than nondecomposable idioms, which led them to suggest that these idioms have a processing advantage since they can be processed compositionally as well as retrieved directly from the mental lexicon.

Other studies however have found contradictory results in regard to the decomposability of idioms. Tabossi et al. (2008) investigated syntactic variation of idioms based on their decomposability, following Gibbs and Nayak (1989). They found that abnormally decomposable idioms received reliably lower semantic similarity ratings than both normally decomposable and nondecomposable idioms, but that the latter two were not significantly different. This finding suggests that normally decomposable idioms are not seen as more syntactically flexible than nondecomposable idioms. They also conducted a study investigating participants' reliability classifying idioms into the three categories: normally decomposable, abnormally decomposable, and nondecomposable. They utilized the same procedure as Gibbs and colleagues (Gibbs and Nayak, 1989; Gibbs et al., 1989a,b) – participants first grouped the idioms as either decomposable or nondecomposable, then further divided the idioms in the decomposable category into normally or abnormally decomposable. Using a binomial test, they calculated the proportion of agreement between the participant judges and found that classifying idioms as decomposable vs. nondecomposable, as well as normally vs. abnormally decomposable, was not significantly different from chance. Titone and Connine (1994b) also conducted a classification task following this same procedure used by Gibbs and colleagues. They also found that participants are unreliable at grouping idioms into the three decomposability categories. Participants were more reliable distinguishing decomposable idioms from nondecomposable idioms, than distinguishing between normally and abnormally decomposable idioms, but in two separate classification tasks only 40% of idioms were reliably classified into these three categories using a 75% agreement rate. These findings reveal that classification into these categories is unreliable across participants.

In fact, Gibbs et al. (1989a) had to remove the abnormally decomposable idioms from the analysis in order to achieve significance between normally decompos-

able idioms and nondecomposable idioms. Meanwhile, Gibbs et al. (1989b) indicated that they included idioms in a particular decomposability category when participants agreed on their placement "more than 50%"; in other words, at chance. Contradictory results are also observed in idiom processing. Gibbs and Gonzales (1985) found that frozen idioms (i.e. nondecomposable idioms) were responded to faster than more flexible, or decomposable, ones. While Tabossi et al. (2008) found no effect of decomposability, or flexibilty, in their experiment; the only significant effect they found was that idioms are processed faster than literal controls. Thus, categorizing the compositionality of idioms has not proved a reliable measure, and the effects of this categorization have not shown consistent results.

Some scholars have instead attempted to account for an idiom's 'transparency', or the degree to which the expression is considered related (i.e. a motivated relationship) to its figurative meaning (cf. Skoufaki, 2009; Keysar and Bly, 1999). Keysar and Bly (1995) found that knowledge and use of the idiom's meaning appears to influence measures of transparency. In that study, they taught half the participants the original meaning of unfamiliar idioms like *the goose hangs high*, 'things look good', and the other half of the participants the conceptual opposite meaning, 'things look bad'. They then asked the participants to rate the transparency of the expression, with either the original or the opposite meaning, from the perspective of "an uninformed individual". Participants were more likely to rate the expression as transparent if it was the meaning they had learned, regardless of whether it was the original or the conceptual opposite. Their second experiment followed the same procedure, but in addition they asked participants to use some of these newly learned idioms in sentences before proceeding to the testing phase. They again found that participants rated the learned meanings as more transparent but that the transparency increased with increased use of the idiom. Meanwhile, Nippold and Taylor (2002) observed that children and adolescents do not significantly differ in how transparent they rate idioms, despite differing in familiarity and comprehensibility. Given Keysar and Bly's (1995) findings, this result might be due to their knowledge of the idiom.

These findings demonstrate the important role that knowledge of the idiomatic meaning can have for how transparent speakers regard the idiom. But speakers can still have intuitions about the meaning of idioms not previously encountered. In the pre-test in Keysar and Bly (1995), they asked native speakers to choose the meaning of the idiom: its original meaning or its conceptual opposite. The mean for selecting

the correct meaning was 51%. However, some idioms did show a preference for one meaning or the other. *Have someone dead to rights* received a strong preference (82%) for its original meaning 'unquestionably guilty' (vs. the conceptual opposite 'unquestionably innocent'), while *lay out in lavender* received a strong preference (85%) for the conceptual opposite meaning 'to flatter' (vs. the original meaning 'to chastise harshly'). This variation shows that native speakers can still have strong intuitions about the meaning of individual expressions.

This study attempts to determine which factors contribute to the transparency of idiomatic expressions, while obtaining a mean measure of transparency for each idiom to be included in the analyses of the following chapters. A much larger sample of idioms (180) is included in this study compared with previous studies, which should include both familiar and unfamiliar idioms since speakers familiarity with idioms presumably varies. This larger sample of idioms should also coincide with a range of possible measures of transparency. This study asks participants to rate the transparency along a continuum, instead of grouping the idioms into discrete categories. Speakers intuitions about idioms undoubtedly vary, from the relationship of the component parts to the clarity of the expression as a whole; therefore, they will be using a continuous scale to more accurately account for this variable distribution. The idioms included in this study are a subset of multi-word units, situated towards the idiomatic end of the 'idiomaticity continuum', which ranges from more collocational in nature to more abstract in meaning and non-compositional (cf. Fillmore et al., 1988; Pawley, 1985; Wulff, 2009; Taylor, 2012); another reason to ensure transparency is reflected along a continuum.

## 2.2 Methodology

### 2.2.1 Materials

One hundred and eighty idioms were selected from the Oxford Dictionary of English Idioms (Ayto, 2009) and the Collins COBUILD Idioms Dictionary (Sinclair, 2011), along with their definitions and example sentences. These idioms were selected because of their inclusion in other experiments (cf. the following chapters). Of these idioms, 173 were used in another experiment (with 17 overlapping in all experiments) and an additional seven idioms were selected as "fillers" (4 opaque and 3 transparent)

for an even 180. Six additional idioms were used in the practice: 2 opaque (i.e. *chew the fat, elbow grease*), 2 transparent (i.e. *lend a hand, as flat as a pancake*), and 2 semi-transparent (i.e. *a foot in the door, call it quits*).

### 2.2.2 Procedure

Participants were asked to rate the transparency of idioms (i.e. how obvious is the meaning of each idiom). They were first provided with a definition of transparency, adapted from the Oxford English Dictionary (2015), see (2.1). They were also provided with an example of a transparent expression, as well as an opaque expression, shown in (2.2). They were then provided with a paraphrased meaning of transparency – How obvious is the meaning of the expression? To what extent does the meaning of the individual words contribute to the meaning of the whole? – before starting the practice.

(2.1) Definition of Transparency:
Evident, obvious, clear; easily recognized or understood
*Linguistics*: obvious in structure or meaning

(2.2) Examples:
transparent – *write a letter* 'write a letter'
not transparent – *don't have a cow* 'calm down, relax'

The participants were presented with each idiom, in random order, along with its definition and an example sentence on the same screen. The text was presented in a black, bold, 22-point Courier New font, centered on a white background. After each idiom, they were asked if they knew the expression (i.e. yes or no). Using the mouse, they clicked on the appropriate box to respond. They were then asked to rate the transparency of the expression. Also using the mouse, they could click anywhere on the provided scale, which was labelled with 'transparent' on the extreme right and 'not transparent' on the extreme left. This scale was a Visual Analogue Scale (VAS), which is a continuous graphical rating scale. The advantage of this continuous scale is that fine gradations are able to be measured, unlike an ordinal

scale of discrete categories which tends to involve compromises in decision-making (Hayes and Patterson, 1921; Freyd, 1923; Funke and Reips, 2012). Each pixel length of the computer screen corresponds to a possible value, and in this case, pixels were converted into a numerical value ranging between 0 – 100. The responses from VAS are output in a form that is ideal for determining differences in distributions and can be subjected to a greater number of statistical procedures (Aitken, 1969; Funke and Reips, 2012). Finally, VAS has been shown to be a reliable, valid, and responsive self-report measure (Gift, 1989; de Boer et al., 2004). Participants were reminded to use the whole scale before the experiment began. They were also given an opportunity to take a short break halfway through the experiment.

After the experiment, the participants were presented with a few additional questions. They were first asked three questions pertaining to their idiom usage: (1) How often do you use these expressions?; (2) How often do others around you use these expressions?; and (3) Do you like using these expressions? They responded to these questions by using the same VAS scale that they had used to rate transparency, this time labelled with a 'thumbs-up' picture on the right and a 'thumbs-down' picture on the left. They were next asked to rate the acceptability of seven prescriptively 'incorrect' sentences, shown below, also using the same VAS scale.

Language Questions (LQs):

1. The only option the school board has is to lay off a large amount of people.
2. Slot machines are thought to be more addicting than table games.
3. The document had to be signed by both Susan and I.
4. While cleaning the kitchen, Sally looked up and saw a spider on the roof.
5. I thought it could've went either way.
6. She could care less what he had to say about it.
7. You have to balance your life, irregardless of what anybody thinks.

### 2.2.3 Participants

Twenty native speakers of Canadian English participated in this experiment. All participants were first-year undergraduate linguistics students from the University of

Alberta. Half of the participants had previously participated in another experiment on idiomatic variation (i.e. one of the experiments discussed in a later chapter), which appears not to be a significant predictor. All participants were reimbursed for their time with course credit.

## 2.3 Results

### 2.3.1 Variables

The data collected from this experiment were analyzed using linear mixed-effects regression, using the `lme4` package (Bates et al., 2014) in R (R Core Team, 2014). Two response variables were analyzed: rating responses and reaction times. Both are continuous variables – the rating responses range from 0 – 100, while the reaction times are log transformed measures in milliseconds.

Eight significant predictor variables are included in the two models; half are subject-related responses to additional questions and half are idiom-related measures. `KnowIdiom` is a factor indicating the participant's knowledge of the idiom (i.e. 'yes' or 'no'). `LikeUsingIdioms` is a scaled measure of the participant's overall enjoyment of using idioms – a higher positive rating indicates that a participant enjoys using idioms, whereas a higher negative rating indicates a dispreference for their use.

Two variables, `LQ5` and `LQ7`, which pertain to the participant's 'flexibility' with language (i.e. they are the participant's responses to Language Question 5 and Language Question 7 respectively), are also included in the models. These questions were asked to obtain an approximate measure of the participant's 'permissiveness' with language to determine whether a more 'relaxed' attitude towards grammar influences one's judgement of idiomatic transparency. If speakers are more forgiving when it comes to language usage, then they might also be more flexible at gleaning meaning from forms typically considered non-literal or non-compositional. These two variables were selected because they produced the largest (`LQ5`) and smallest (`LQ7`) loadings in a Principal Components Analysis (PCA). PCA "reduces the number of dimensions required for locating the approximate position of the data points" (Baayen, 2008, p.120). This data set includes 180 idioms (i.e. data points) and 7 language acceptability questions (i.e. dimensions). The data points remain fixed

at their location, while new dimensions, or Principal Components (PCs), are determined. These PCs more succintly explain the data, while accounting for the same amount of variance as the original variables. However, the original variables (i.e. `LQ5` and `LQ7`) produced a lower AIC value[1] (independently and together) than using PC1 and/or PC2 from the PCA. Interestingly, `LQ5` and `LQ7` target different aspects of the grammar. `LQ5` (i.e. *could've went*) obtains an estimation of the participant's flexibility with grammatical forms, in this case participial forms in the irregular verb paradigm, whereas `LQ7` (i.e. *irregardless*) indicates the participant's flexibility with blended structures (i.e. *regardless* and *irrespective*), which result in reduced compositionality.

`logFrequencyIdiom` is the log frequency of the idiom in its canonical or citation form, extracted from the Corpus of Contemporary American English (COCA; Davies, 2008). `Length` indicates the number of words in the idiom, which ranged from two to eight words. `LSA.Score.Definition` is a Latent Semantic Analysis measure of similarity between the words of the idiom and its definition (i.e. the same definition provided to the participants during the experiment). The LSA scores were obtained from a pairwise comparison of two texts (i.e. an expression and a definition).[2] This analysis compares the local contexts of the two texts to obtain a value of similarity (Landauer et al., 1998). If the exact words of the expression have little, if anything, to do with the overall meaning, then the LSA score will be small (e.g. *highway robbery* 'charged a lot of money for something that should cost a lot less' = 0.06); however, if the exact words used share meaning or contribute to the overall idiomatic meaning more intrinsically, then the LSA score will be much larger (e.g. *as flat as a pancake* 'completely flat' = 0.94). Idioms with higher LSA scores should be rated as more transparent, as the words in the expression contribute to the idiom's meaning. Finally, `TrialScaled` is a scaled trial order indicating the order in which the idioms were presented to the participants. Since the stimuli was presented in random order, this order is different for each participant.

[1] AIC, or Akaike's Information Criterion, measures how well the predictor variables improve the predictiveness of the model (cf. Akaike, 1974, 1985). This measure is a method of model comparison, where a log likelihood ratio test is used to measure the difference of the entropies of two models (i.e. the negative of the entropy, or negentropy, is often regarded as the amount of information), given the number of parameters in the model. Thus, the model with the smallest AIC value is the preferred model.

[2] The LSA scores were obtained through the English Lexicon Project (Balota et al., 2007), available at http://lsa.colorado.edu/.

### 2.3.2 Rating Responses

The regression model for the rating responses includes seven significant predictor variables, four of which occur in a significant interaction with `KnowIdiom`. The estimates, standard error, and t-values for these predictors are shown in Appendix A.1, while the interactions are plotted in Figure 2.1.

The first interaction, presented in the top-left panel of Figure 2.1, is between the frequency of the idiom (`logFrequencyIdiom`) and whether the participant knows the expression (`KnowIdiom`). When speakers do not know a particular idiom, they are not affected by the frequency of the idiom, as expected (t = 0.75). Frequency is predictive however when a speaker is familiar with an expression. The more frequent the idiom, the higher speakers rate its transparency (t = 5.5).

`KnowIdiom` also interacts with `LikeUsingIdioms`, shown in the top-right panel of Figure 2.1. The greatest effect is observed for speakers who do not know a particular expression. If participants dislike using idioms, then they tend to rate the transparency of the unknown idiom extremely low (i.e. the idiom is not transparent at all), but the more speakers enjoy using these kinds of expressions, the higher they rate the transparency of unknown idioms. Thus, the more enjoyment one receives from using figurative language, the more interpretable and transparent idioms seem when first encountered. Sentiments towards idiomatic expressions contribute less if speakers are already familiar with the expression (i.e. `LikeUsingIdioms` is less predictive when the expression is known). `LikeUsingIdioms` also enters into an interaction with `Length`. The more speakers enjoy using idioms and the longer the idiom, the higher rated the transparency of the idiom (t = 2.22). Longer idioms are considered more transparent, especially if speakers generally enjoy using idioms.

`KnowIdiom` occurs in significant interactions with both `LQ5` and `LQ7`, shown in the bottom two panels of Figure 2.1. If participants know the idiom, then the degree to which their flexibility with language (i.e. their acceptability of prescriptively incorrect sentences) is minimally predictive of idiomatic transparency. However, if the participant does not know the idiom, then the participant's flexibility with language is more predictive. Unknown idioms were rated as less transparent the more `LQ5` (i.e. *I thought it could've went either way.*) was deemed acceptable. The reverse pattern emerged for `LQ7` – unknown idioms were rated as more transparent the more `LQ7` (e.g. *You have to balance your life, irregardless of what anybody thinks.*)

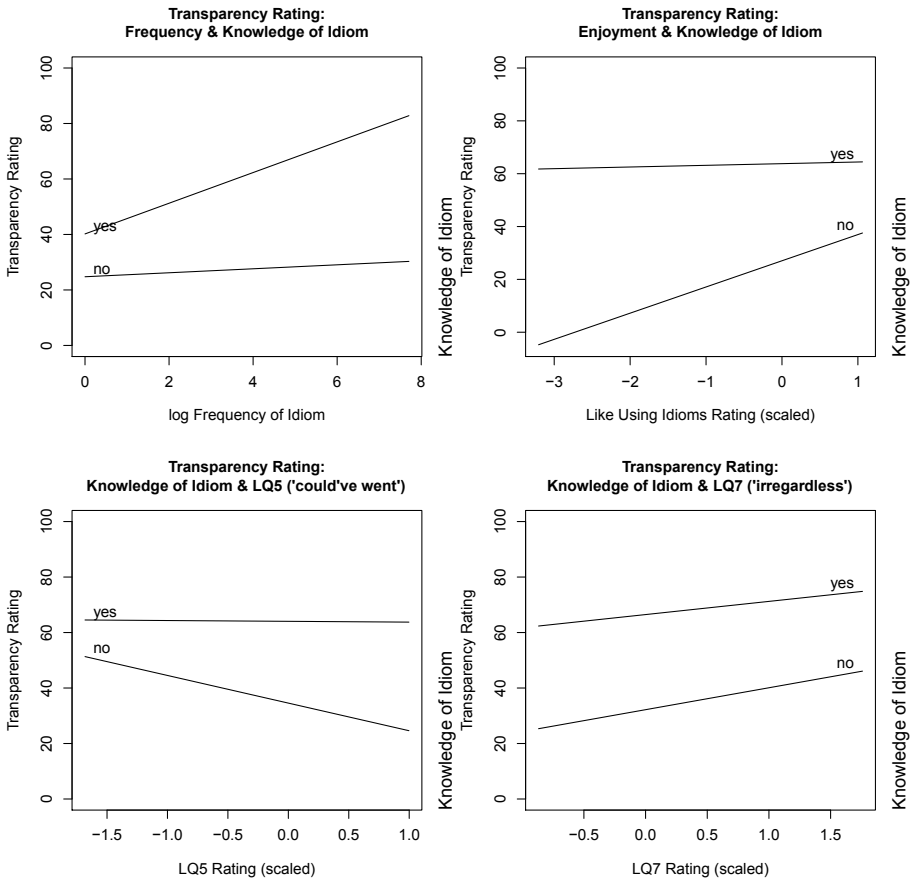


Figure 2.1: Interactions in the Mixed-Effect Linear Regression Model for the Transparency Rating Responses of English Idioms

was found acceptable. The interpretation of these findings may reside in the type of 'errors' found in each sentence. `LQ5` contains a grammatical error: the use of the past tense as a past participle (i.e. *could've went*). Being flexible with morphological tense and aspect marking in the verb phrase may not assist speakers with non-compositionality; instead this finding suggests that any flexibility these speakers have with language resides only in the grammar and not in semantic creativity hidden within a non-compositional structure. `LQ7` however, contains a morphological word blend (i.e. *irregardless*, from *regardless* and *irrespective*). Multiple units go into creating the wordform *irregardless*, resulting in a non-compositional structure. Thus, the more accepting speakers are of this structure, then the more accepting they are of the non-compositional structure of idioms, and in turn the more transparent they consider idioms to be, even ones they are unfamiliar with. Flexibility with language is less predictive when the speaker knows the expression because an interpretation and connections to the component parts have already been established and therefore other strategies are not needed to interpret the expression and determine its clarity.

This model also has one significant main effect: `LSA.Score.Definition`. The higher the LSA score (i.e. the more contextually similar an idiom is with its definition), the more transparent participants rated the idiom, as expected. This result was consistent for both those who know the idiom and those who do not; hence the variable is not significant in an interaction with `KnowIdiom`.

The random effects structure of the model shows by-Item random slopes with a correlation parameter for `KnowIdiom`, showing variability with which idioms the participants were familiar. The model also has by-Subject random slopes for `TrialScaled` and by-Subject random slopes for `logFrequencyIdiom`, indicating that participants showed variability in their ratings for when in the experiment the idiom was presented, as well as for its frequency. This latter result likely suggests that participants differed in terms of their experience and familiarity with certain idioms from how frequently they were represented in the corpus. This random-effects structure nicely illustrates the degree to which idiom use can vary among speakers.

### 2.3.3 Reaction Times

Four variables were significant in a model predicting reaction time, shown in Appendix A.2; two of which occur in significant interactions with `KnowIdiom`. The first

interaction, shown in the left panel of Figure 2.2, is between the frequency of the idiom and the participant's knowledge of the idiom. This interaction shows a similar pattern to what was observed in the previous section for the rating responses – frequency is only predictive if the idiom is already known. If the expression has a higher frequency and it is familiar, then the participants are faster in their responses.

The second interaction, shown in the right panel of Figure 2.2, is between `KnowIdiom` and `LikeUsingIdioms`. The participants who do not know the idiom are more affected by their overall enjoyment of using idioms in general, than those who know the expression. Participants who like using idioms tend to be slower in their responses, especially if they do not know the expression. This finding suggests that participants who like using idioms are trying to be more careful or thoughtful while rating the transparency, compared with those who do not enjoy using idioms. This is especially true if the participant does not know the expression; they are even more likely to spend extra time understanding the idiom and determining its transparency.

`TrialScaled` is the last significant fixed-effects predictor. The further the participant gets into the experiment, the faster their response times. However, participants differed in how much quicker they became at rating the transparency, as evident by the by-Subject random slopes for `TrialScaled`.

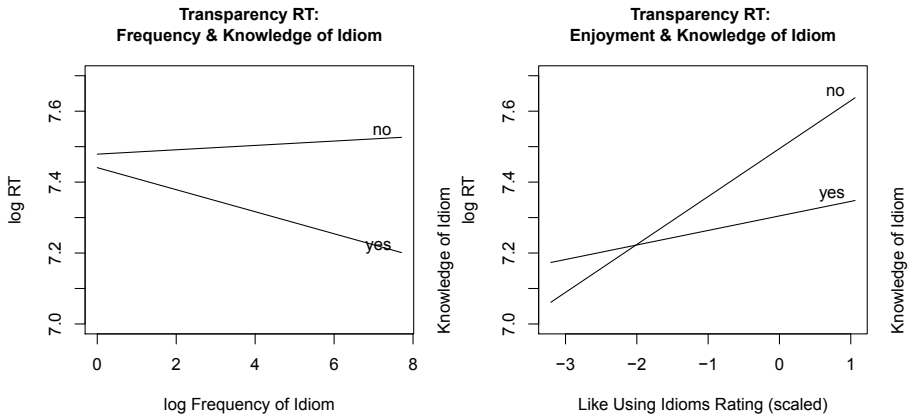


Figure 2.2: Interactions in the Mixed-Effect Linear Regression Model for the Reaction Times to Rate the Transparency of English Idioms

## 2.4 Discussion

The results from this study show a range of variables, both idiom-related and participant-related, that are relevant for predicting ratings of idiomatic transparency. Previous literature has focused mostly on idiom-related variables. For example, the length of the idiom has not been investigated in great detail; but Fanari et al. (2010) have shown that idiomatic meaning associated with longer idioms is recognized before one has reached the end of the idiom. This same result is not true of shorter idioms. Interestingly, speakers tend to rate longer idioms as being more transparent. Perhaps longer idioms convey more information, making their idiomatic meaning clearer and more assessible, thus facilitating the recognition of the idiom.

Contextual similarity (i.e. LSA scores) is also a predictive measure of how transparent speakers find an idiom to be. The more similarity exists between the two texts (i.e. the more similar the local contexts between an idiom and its definition), the more transparent the idiom. This measure indicates the extent to which the idiom is interchangeable with its definition, but more importantly, indicates the degree to which the individual words contribute to the overall idiomatic meaning. While this specific measure does not typically appear in idiom analyses, it more accurately accounts for the compositionality of the expressions than the traditional decomposability classification (cf. Nunberg, 1978; Gibbs and Nayak, 1989), objectively based on text analysis and not subjective categorization. As expected, idioms are regarded as more transparent when they are more compositional.

Frequency of the idiom is also shown to be significant, as anticipated due to the frequency effects observed with a variety of other multi-word expressions (cf. Arnon and Snider, 2010; Shaoul and Westbury, 2011; Tremblay et al., 2011). Previous studies on idioms do not typically include a measure of frequency, but often collapse frequency with familiarity – the subjective frequency with which the speaker encounters the expression (cf. Titone and Connine, 1994b; Tabossi et al., 2009a). This study has tried to tease apart the frequency of the expression from one's knowledge of the expression and found that idioms with higher frequency tend to be rated faster and tend to be rated with a higher value of transparency, but only if the speaker knows the expression. These results demonstrate that we retain information about the frequency of these expressions once we are exposed to them (cf. Shaoul and Westbury, 2011; Taylor, 2012).

However, the variability for frequency observed in the random-effects structure suggests that the speaker's individual familiarity and experience with each expression may still be important in understanding the transparency of idioms. This frequency measure, extracted from a corpus, may not accurately represent the frequency with which each speaker is aware of each idiom. This could be due to external factors, such as dialectal usage of these expressions – the corpus represents American English while the participants were speakers of Canadian English, but it could also be due to idiosyncratic usage and familiarity. Perhaps frequency as well as familiarity need to both be included in future studies on idioms: frequency representing each idiom's estimated frequency in the language and familiarity representing the speaker's personal experience with each idiom. Popiel and McRae (1988) included both measures in their study (albeit both were subjective ratings by their participants) and found that idioms received higher ratings for frequency than for familiarity. Familiarity has been shown to be a significant predictor in idiom comprehension (cf. Schweigert, 1986; Schraw et al., 1988; McGlone et al., 1994), but future studies may need to complement this variable with a more neutral measure of frequency, or overall usage in the language community.

Participant-related measures are also shown to be predictive in understanding idiomatic transparency – measures not typically discussed in the idiom literature. Whether the speaker enjoys using idioms seems to be an important predictor, for both rating responses and reaction times. Speakers who enjoy using idioms tend to consider them more transparent, especially if the idiom is unfamiliar. They also tend to take longer rating the transparency, possibly to reflect more on the idiom and its meaning. The speaker's flexibility with language also appears to be an important predictor, particularly when the expressions are unknown. Speakers who are more flexible with grammatical modifications (e.g. *could've went*) tend to consider unknown idioms less transparent, whereas speakers who are more flexible with morphological blending (e.g. *irregardless*), and therefore greater non-compositionality, consider unknown idioms more transparent. These variables provide some information about the individual differences evident in speaker's judgements and demonstrate the need for a more in-depth individual differences study on idioms and idiomatic transparency (cf. Bates et al., 1995; Kuperman and Van Dyke, 2011; Dabrowska, 2014). Idioms have traditionally been regarded as a homogeneous class of non-compositional units and this study nicely illustrates the different degrees of transparency between id-

ioms, as well as the ways in which speakers' personality traits might influence their opinions of an idiom's transparency.

The most important predictor of idiomatic transparency appears to be knowledge of the idiom. This variable occurs in several interactions and even shows variability within the random-effects structure of the rating response model. Speakers who know an expression rate it as being more transparent and are typically faster in their responses. These findings are in line with those of Keysar and Bly (1995), who found that speakers rated learned meanings of unfamiliar idioms as more transparent, regardless of whether the meaning was the original meaning of the idiom or its conceptual opposite. Once we learn the meaning of an idiom, it becomes more transparent because we are able to make sense of the meaning; we are able to create a story about it and why it means what it does (Keysar and Bly, 1999). Idioms are created from non-conventional metaphors and metonymies which are used to express abstract concepts by expanding the literal meanings of the words (Langlotz, 2006). We interpret and understand idioms, predominantly from their usage in context, with these associated metaphors. For example, *spill the beans* means 'to reveal a secret' and we understand this expression by connecting the *beans* to 'secrets' through metaphors such as IDEAS ARE PHYSICAL ENTITIES, and when we *spill* these entities, they disperse everywhere and are difficult to retrieve. The metaphors used to create this expression are accessible and we employ them in using and understanding it. However, if this expression was created with the meaning 'to keep a secret', we would use different strategies for interpreting this meaning. The *beans* might now be connected to 'distracting information' through the same metaphor IDEAS ARE PHYSICAL ENTITIES, but this time interpreted as *spilling* distracting and useless information in order to divert the listener's attention and 'keep a secret'.

This ability to create meaningful interpretations is observed when original metaphors or metonymies associated with idioms are no longer accessible. For instance, the idiom *kick the bucket* originated as a euphemistic way to describe the slaughtering of hogs, which would literally *kick the buquet* in their last reflex of life. *Buquet* was a loanword from French (cognate with modern French *boucher* and modern English *butcher*) for the wooden framework used in slaughtering the animal; now realized as *bucket* – a corruption of *buquet* – in the expression (Urdang, 1988; Gibbs, 1994). Most people however, are not familiar with this etymology, nor would it make much sense even if they were, as this instrument, and lexical item, no longer

exist. But speakers are able to create a new story around its meaning and establish new connections for why this idiom means what it does, which is why speakers today often folk-etymologize the meaning of this idiom as involving a hanging, where the condemned person about to die is standing on a bucket with a rope around his neck when the bucket is kicked out from underneath (Gibbs, 1994, p.276). This ability to create meaningful connections for known idioms, with possible scenarios and metaphorical links, is what allows speakers to regard known idioms as more transparent than unknown ones.

Studies investigating the learning of idioms reveal some of the strategies used in interpreting idioms and creating meaningful connections for them. Cacciari (1993) looked at adults' ability to make sense of unfamiliar idioms. Among the strategies employed, she found speakers would compare the unknown idioms to similar known idioms, interpret one of the words in the expression literally and then use the other words to expand on the semantics to obtain the meaning, and even instances where speakers would visualize an image of what the idiom could mean. These strategies were also observed in Skoufaki (2009), who investigated the transparency of unfamiliar idioms with second language learners. In that study, learners specified possible interpretations for unknown idioms, and although fewer interpretation types were produced for more transparent idioms, multiple interpretations were provided. This reinforces the idea that idioms could have another possible meaning and that as long as a story can be created (or re-created) for their meaning, idioms will make sense and become regarded as more transparent (Keysar and Bly, 1999).

Idioms can also come to be judged as less transparent over time, as culture and society changes, generations have different experiences, and the metaphorical references and connections become less obvious (cf. Bowdle and Gentner, 2005). This is observed with *kick the bucket* where agricultural advancement resulted in a lost instrument, lexical item, and metonymic link to the preserved idiom. Reduced transparency is also observed in the expression *three sheets to the wind* 'very drunk' (originally *three sheets in the wind*). Today, this idiom seems quite obscure; in fact, its mean transparency rating is 13.6 (0 = 'not transparent'; 100 = 'transparent'). However, to a seafarer back in the 1800's this expression would make perfect sense. The *sheets* refer to "the ropes attached to the corner of a ship's sail, used for controlling the extent and direction of the sail" (Ayto, 2009). If all three of these sheets are loose and *in the wind*, then the sails will be flapping around, and the ship will

be out of control "lurching about like a drunken sailor".[3]

Once an idiom has been encountered, connections are created between the associated metaphors or metonymies, as well as any folk etymologies or visual scenarios that may be created to further explain or rationalize the expression, especially if the original metaphors are no longer immediately accessible. Connections are also formed between the literal meanings of the words and their abstract figurative extensions. Over time, these figurative extensions can be reinterpreted to form additional, polysemous meanings of the individual word, used in isolation outside of the original idiomatic context, such as *spill* meaning 'reveal' in contexts other than *spill the beans* (Geeraerts, 1995). Additional aspects of the idiom's usage may also be incorporated into these connections, such as how or when the expression should be used. For example, *kick the bucket* is largely used to refer to an immediate and instantaneous death, likely due to its origins and preserved in the verb *kick* (Hamblin and Gibbs, 1999). This idiom is also reserved for distant and unfamiliar persons, or inanimate objects that have stopped working, and never about the passing of a loved one, also due to the idiom's origins euphemising the slaughtering of hogs, until it was minimally extended into other contexts. These additional associations are why idioms never equal their literal paraphrases; *kick the bucket* does not simply mean 'to die' (cf. Gibbs, 2007). Thus, learning an idiom is about creating a complex array of connections, which provide detailed information and context, making the expression seem semantically clearer and more interpretable to speakers, as well as more memorable, especially considering their relatively low frequency in the language. This complexity of intertwined meaning and associations is what makes categorizing these idioms into the traditional decomposability categories so difficult – the task forces speakers to sever the connections they have established and focus on the individual elements rather than the package.

Research on language has traditionally assumed that each individual word contributes meaning in a linear order. This contributed meaning was also thought to be the word's literal meaning. But this assumption then meant that idioms had to be processed differently, as most idioms do not have individual components which contribute a literal meaning. Therefore, it was assumed that idioms were stored whole, as a unit or a 'large word', and were accessed directly from the mental lexicon. This

[3]Retrieved from http://www.phrases.org.uk/meanings/three-sheets-to-the-wind.html on May 10, 2015.

lead to research trying to determine how idioms are activated and integrated into literal language processing (cf. Swinney and Cutler, 1979; Gibbs, 1986; Cacciari and Tabossi, 1988) and whether different types of idioms are processed differently, such as decomposable idioms are processed like literal language whereas nondecomposable idioms are accessed whole (cf. Gibbs et al., 1989b; Titone and Connine, 1999; Tabossi et al., 2008). However, recent research suggests that this traditional view and approach to language may not be entirely correct (cf. Elman, 2004, 2011; Baayen et al., 2011; Baayen and Ramscar, 2015). In these more recent approaches, words themselves do not possess meaning, but are instead cues to meaning, modulated by experience and context. Words are a signal intended to reduce the listener's uncertainty about the world, feelings of the speaker, etc. (Ramscar and Baayen, 2013). Under this view, idioms would not need to be regarded any differently, but instead would be a sequence of words which are cues to meaning. Usage and experience with these idioms (i.e. word sequences) allow speakers to create complex connections and associations with their metaphors, metonymies, pragmatic nuances, etc., which are available the next time these words are utilized in context. Thus, individual words of an idiom do not independently contribute meaning, but are instead utilized together to convey a wealth of information.

# CHAPTER 3

# Acceptability Ratings of Idiomatic Variation

## 3.1 Introduction

Researchers investigating idioms have a long history of only focusing on the canonical form of the expression, predominantly through experiments aimed at determining how we process these non-literal forms of language. For decades, researchers assumed idioms were frozen lexical units (i.e. fixed in form), either stored in a separate 'idiom list' (cf. Bobrow and Bell, 1973) or stored whole as a lexical unit in the mental lexicon (cf. Swinney and Cutler, 1979; Schraw et al., 1988; Sprenger, 2003). More recently however, researchers have begun to explore variation within idiomatic expressions, largely through corpus-based studies (cf. Moon, 1998; Barlow, 2000; Schröder, 2013). These studies have shown that idioms are not fixed lexical units, frozen in form, but are actually utilized in a variety of ways and with varying degrees of productivity. For example, Moon (1998) has shown that idioms can occur with lexical variation in nouns (e.g. *a skeleton in the closet/cupboard*), verbs (e.g. *say/kiss goodbye to something*), and even particles (e.g. *rap someone on/over the knuckles*). Variations can also include truncations (e.g. *don't count your chickens* [*before they're hatched*]), reversals (e.g. *can't eat/have your cake and have/eat it too*), with homophonous words (e.g. *dull as ditchwater/dishwater*), and even insertions (e.g. *we're a little late getting our Christmas act together*).

Schröder (2013) adds to Moon's extensive corpus study by exploring nine idioms in-depth for their syntactic variation. She found that 'metaphorical' idioms,

such as *make headway*, undergo a large degree of syntactic variation (close to 50%), mostly adverbial modification like *make rapid headway*, whereas 'fixed idioms' like *kick the bucket* are used with syntactic variation less than 5%. However, she does not include instances of creative wordplay, such as *Arthur kicked the detonator of the bomb, and, consequently, the bucket*, as tokens of variation in her analysis (Schröder, 2013, p.68). Meanwhile, Duffley (2013) specifically explored two fixed (i.e. nondecomposable) idioms, *kick the bucket* and *shoot the breeze*, on the internet and found a surprising amount of variation for these idioms – idioms which have traditionally been regarded as unalterable – from passives (e.g. *no buckets were kicked*), causatives (e.g. *before we kicked his bucket for him*), lexical variation (e.g. *my phone kicked the pail last week*), and even inserted concepts (e.g. *was ready to kick its digital bucket*). These examples illustrate that idioms are not nearly as fixed or frozen as has been previously assumed. In fact, even those idioms claimed to be the most fixed or frozen (i.e. *kick the bucket*) show a remarkable amount of variation.

Although scholars have begun to explore idiomatic variation in corpora, few studies have examined variation from an experimental perspective. Gibbs and colleagues (Gibbs and Nayak, 1989; Gibbs et al., 1989a) investigated syntactic and lexical variation with idioms in separate similarity rating studies, both attempting to investigate the comprehensibility of such variation. Both studies explored the effects of variation with 'Normally Decomposable' (e.g. *lay down the law*), 'Abnormally Decomposable' (e.g. *meet your maker*), and 'Nondecomposable' (e.g. *kick the bucket*) idioms. Decomposability is the degree to which the constituent parts of the idiom contribute to the figurative meaning of the whole. Normally decomposable idioms have constituents which directly contribute meaning to the expression; for example, *the law* is the set of rules or guidelines being implemented and *lay down* is the verb used to describe the act of enforcing those rules. Meanwhile, the constituents in abnormally decomposable idioms contribute meaning indirectly or through a metaphorical relationship – *Maker*, in *meet your maker*, is a conventionalized metaphor in Western society referring to God. Nondecomposable idioms however do not have constituents which contribute to the meaning of the expression; neither *kick* nor *bucket* contribute to the meaning 'to die'. In these studies, variants were presented to the participants, along with literal paraphrases, and participants were asked to rate the similarity in meaning between the two. The results show that normally decomposable idioms are preferred with lexical and syntactic alternations

(i.e. they were rated as more similar in meaning).

Tabossi et al. (2008) conducted a replication study of Gibbs and Nayak (1989) where they investigated syntactic variation of decomposable and nondecomposable idioms in Italian. They followed the same procedure, asking participants to rate the similarity in meaning between the idiomatic variant and its literal paraphrase. Unlike Gibbs and Nayak however, they found that both normally decomposable idioms and nondecomposable idioms were preferred with syntactic variation more than abnormally decomposable idioms. Meanwhile, Titone and Connine (1994b) conducted a replication study where they asked participants to group idioms into 'decomposable' or 'nondecomposable' categories, and then to group the decomposable idioms into 'normally decomposable' and 'abnormally decomposable' idioms. They found that participants were unreliable at categorizing idioms – only 40% of 171 idioms were reliably classified into the three categories using a 75% agreement rate among participants in two separate classification tasks. In fact, there are even discrepancies in the literature to which category an idiom belongs; for example Gibbs et al. (1989a) list *button your lips* as normally decomposable, while Libben and Titone (2008) list it as abnormally decomposable. Discrepancies even occur within studies by the same researcher: Gibbs et al. (1989a) list *cook one's goose* as nondecomposable, but Gibbs and Nayak (1989) list it as abnormally decomposable. These findings and observations show little support for the decomposability measure, instead they demonstrate that people differ with how they regard the relationship between an idiom's component parts and the meaning of the whole.

McGlone et al. (1994) also investigated variation in idioms, focusing less on synonymous lexical variation and more on semantic productivity within idiom variants. In their first experiment, they asked participants to paraphrase idiom variants, rate the comprehensibility of the variant, and then to rate their familiarity with each expression. They found that participants rated the variants as more comprehensible when they were more familiar with the expression.

These studies have all attempted to explore the effects that variation has on the comprehension of idioms. However, they implemented a ratings task to access participants' understanding of the variants. What they may have observed instead was not necessarily the degree of difficulty in understanding these variants, but the participants preference for, or approval of, the variants. Additionally, these ratings were predominantly a measure of similarity. Studies have shown that semantic

similarity is largely predicted by the same local contexts as observed in corpora (Miller and Charles, 1991); thus, these studies are more likely measuring how interchangeable the variant is with its literal paraphrase. However, idioms convey more information than their literal paraphrases (Gibbs, 1994, 2007). For example, when someone says *kick the bucket*, we know more information than just 'someone died' – we know that it was an immediate or punctual death, that the person was not close to us personally or that it was an inanimate object which ceases to work, that it is an informal situation, etc. This information is not conveyed in the literal paraphrase 'die'. By asking participants to rate the semantic similarity, the researcher is asking for the degree to which these items can be used interchangeably. It seems reasonable then that nondecomposable idioms received lower similarity ratings – they are less likely to be used in the same contexts.

These studies also have designs which make it difficult to interpret some of the results. Gibbs et al. (1989a) presented the variants and their literal paraphrases in isolation. Speakers may have rated variants differently had they occurred in context – an important criterion considering all language occurs in context. Both Gibbs and Nayak (1989) and Tabossi et al. (2008) presented their stimuli in context, but did not control for the contexts between the different conditions. Their findings may therefore be confounded. The ratings might not reflect the preference for each idiom with the particular syntactic variation, but rather the context in which it occurred. Meanwhile, McGlone et al. (1994) did not control the type of variation utilized in the stimuli; for example, they included lexical variation (e.g. *He was the sort of person who always bit off less than he could chew*), quantification of a noun (e.g. *He didn't spill a single bean*), and reversals (e.g. *He had gone from riches to rags*). While this study reveals that people prefer variation when they are familiar with the expression, it does not show whether people prefer certain types of variation.

This study differs from the previous studies in two important ways. First, I attempt to explore people's preferences for idiomatic variation through an acceptability ratings task. I explicitly ask people to rate how 'acceptable' they find each variant. This measure allows people's criteria in evaluating the variants to vary, as it specifically aims to capture their preferences, whatever they may be. Second, several types of variation are included in the same study to determine if people prefer one type of variation (e.g. lexical variation) over another (e.g. partial form of the idiom).

## 3.2 Methodology

### 3.2.1 Materials

Sixty idioms were extracted from the Oxford Dictionary of English Idioms (Ayto, 2009) and the Collins COBUILD Idioms Dictionary (Sinclair, 2011). These idioms varied in length and syntactic structure: 20 three-word idioms consisting of a verb and a noun phrase (i.e. V-NP; e.g. *rock the boat*); 20 four-word idioms consisting of a verb and a prepositional phrase (i.e. V-PP; e.g. *jump on the bandwagon*); and 20 five- or six-word idioms (10 each) consisting of a verb, noun phrase, and a prepositional phrase (i.e. V-NP-PP; e.g. *hear something through the grapevine*). Two contexts were created for each idiom: one literal and one figurative (e.g. *I used to be a socialite, and I would hear things through the grapevine* = figurative; and *I used to pretend I could talk to plants, and I would hear things through the grapevine* = literal). Both contexts had identical final clauses, with the idiom in sentence-final position.

These idioms were manipulated for four types of variation within the figurative context, in addition to the canonical form. First, lexical variation, where one of the lexical items in the expression was altered to a synonymous or near-synonymous word (e.g. *rock the canoe* or *discover something through the grapevine*). An online thesaurus was often utilized for synonymous words[1]. Second, partial form of the idiom, where only a portion of the expression was presented, usually a key word or words (e.g. *rock things* or *use the grapevine*). In order for the sentence to still be grammatically correct, pronouns or lexically-vague words were used to replace the missing elements of the expression, such as *it, them, things* for nouns, or *have, be, do, use* for verbs. Third, integrated concept, where an additional concept was integrated into the idiom (e.g. *rock the initiative boat* or *hear something through the judgemental grapevine*). These additional concepts expanded or emphasized the figurative context in which the idiom occurred. Finally, a formal idiom blend, where two idioms were blended together (e.g. *rock the applecart* – blending *rock the boat* with *upset the applecart*, or *get wind through the grapevine* – blending *hear something through the grapevine* with *get wind of something*). Each experimental idiom (i.e. one of the 60 idioms selected) was paired with a non-experimental idiom for use in the idiom blend variant. These "blending" idioms were chosen for their intuitive

[1] http://www.thesaurus.com/

Table 3.1: Four types of blends used in the Idiom Blend condition

| **Type of Blend** | **Example** | **Source Idioms** | **Total** |
|---|---|---|---|
| sSYN, sSEM | *rock the applecart* | *rock the boat*<br>*upset the applecart* | 15 |
| sSYN, dSEM | *shoot your tongue* | *shoot the breeze*<br>*hold your tongue* | 15 |
| dSYN, sSEM | *pass the mustard* | *cut the mustard*<br>*pass muster* | 15 |
| dSYN, dSEM | *face the wringer* | *face the music*<br>*put through the wringer* | 15 |

plausibility, but controlled for their syntax and semantics with the experimental idioms (cf. Cutting and Bock, 1997). Four types of blends were created: same syntax, similar semantics (sSYN, sSEM); same syntax, different semantics (sSYN, dSEM); different syntax, similar semantics (dSYN, sSEM); and different syntax, different semantics (dSYN, dSEM), exemplified in Table 3.1. Five instances of each type of blend occurred with the three 'syntactic types' (i.e. V-NP, V-PP, or V-NP-PP), totalling 15 of each blend type.

Half of the idioms had the beginning portion of the expression altered, while the other half of the idioms had alternations made to the final portion of the expression. In total, there are six conditions: one in a literal context and five in a figurative context (i.e. one canonical form and four variants). The experiment utilized a Latin-square design, where every participant saw each idiom once in one of the six conditions. Therefore, six versions of the experiment were created, each one containing 10 idioms in each of the six conditions.

CONDITIONS:

1. **Literal Meaning** of the idiom in its canonical form
   (e.g. *While the guys were reshingling, they suddenly went through the roof.*)
2. **Canonical Form** of the idiom in a figurative context
   (e.g. *Although these were new stocks, they suddenly went through the roof.*)
3. **Lexical Variation** of the idiom in a figurative context
   (e.g. *Although these were new stocks, they suddenly went through the ceiling.*)
4. **Partial Form** of the idiom in a figurative context
   (e.g. *Although these were new stocks, they suddenly went through it.*)

5. **Integrated Concept** within the idiom in a figurative context
   (e.g. *Although these were new stocks, they suddenly went through the investment roof.*)
6. **Idiom Blend** of two idioms in a figurative context
   (e.g. *Although these were new stocks, they suddenly went through the charts.*)

Since the "blending idioms" only occurred in one condition (i.e. Idiom Blend), they were used as fillers in their canonical form in the other five versions of the experiment, occurring in either a figurative or a literal context. Each blending idiom was excluded as a control in the version of the experiment where it occurred in the idiom blend condition in order to avoid a bias in the materials. Therefore, in each version of the experiment, 10 of the blending idioms occurred in a formal blend in the idiom blend condition, while the remaining 50 appeared in their canonical form as fillers. Of these fillers, 20 occurred in a figurative context, while 30 occurred in a literal context. This was done to increase the number of literal contexts in the experiment so that they were not so underrepresented. In sum, each participant saw 110 items: 60 experimental idioms (10 in each of the six conditions) and 50 blending idioms as fillers (20 in a figurative context and 30 in a literal one).[2]

Lastly, six practice sentences were created using six "practice" idioms. These idioms all occurred in their canonical form. Three were in a figurative context and three in a literal context. These were the same for all participants.

### 3.2.2 Procedure

Using E-prime 2.0 standard edition software, each sentence (i.e. context and idiom clauses) was presented in random order at the top of the computer screen. The text was presented in a black, bold, 24-point Courier New font, centered on a white background. Below each sentence was a Visual Analogue Scale (VAS) for participants to rate the acceptability of each expression. VAS is a continuous graphical rating scale, which allows fine gradations to be measured (Hayes and Patterson, 1921; Freyd, 1923; Funke and Reips, 2012). Each pixel length on the computer screen is converted to a possible numeric value, in this case, ranging between 0–100. This scale affords participants the precision to rate their acceptance of idiomatic variation

[2]I thank Lauren Rudat for all her helpful comments and suggestions on improving the stimuli in this experiment.

along a continuum, without the compromises associated with discrete categories and the potential bias of verbal labels for those categories (Aitken, 1969; Gift, 1989). In addition, the output from VAS is already in a form ideal for statistical analysis and can be subjected to a greater number of statistical procedures (Aitken, 1969; Funke and Reips, 2012). Thus, a visual analogue scale was utilized in this study instead of the more traditional Likert scale.

Participants were explicitly told that they would be reading sentences which contained English expressions, but that some of the expressions had been modified in various ways. They were asked to rate the acceptability of the expression, as it occurred in the sentence, by clicking the mouse anywhere on the provided scale, which was labelled with 'acceptable' on the extreme right and 'unacceptable' on the extreme left. The mouse repositioned itself to the centre of the scale on each trial. Participants were reminded to use the whole scale before the experiment began. They were given an opportunity to take a short break halfway through the experiment.

After the participants had rated the acceptability of the variants, they were then asked if they knew each expression. Each idiom appeared, in its canonical form, in a black, bold, 22-point Courier New font, centered on a white background. Above the idiom was the question "Do you know this expression?" and below were two boxes, one labelled 'yes' and the other labelled 'no'. Using the mouse, participants clicked on the appropriate box to respond. The mouse repositioned itself to the center of the screen on each trial.

At the end of the experiment, participants were presented with a few additional questions. They were asked three questions pertaining to their idiom usage: (1) How often do you use these expressions?; (2) How often do others around you use these expressions?; and (3) Do you like using these expressions? They responded to these questions by using the same VAS scale that they had used to rate the acceptability of the expression, but this time labelled with a 'thumbs-up' image on the right and a 'thumbs-down' image on the left. They were then asked to rate the acceptability of seven prescriptively 'incorrect' sentences, shown below, also using the same scale. These sentences attempted to access the participant's flexibility with language and non-standard usage.

Language Questions (LQ):

1. The only option the school board has is to lay off a large amount of people.
2. Slot machines are thought to be more addicting than table games.
3. The document had to be signed by both Susan and I.
4. While cleaning the kitchen, Sally looked up and saw a spider on the roof.
5. I thought it could've went either way.
6. She could care less what he had to say about it.
7. You have to balance your life, irregardless of what anybody thinks.

### 3.2.3 Participants

Seventy-two undergraduate linguistics students from the University of Alberta participated in this experiment. Participants were native or near-native speakers of English, with 48 participants declaring English as their first language. There were 55 female and 17 male participants, ranging from 17–43 years of age. Six participants were left-handed. All participants were reimbursed for their time with course credit.

## 3.3 Results

### 3.3.1 Variables

The results were analyzed with mixed-effects linear regression using the `lme4` package (Bates et al., 2014) in `R` (R Core Team, 2014). Only the 60 experimental idioms were included in this analysis (i.e. the fillers, or blending idioms, were not included outside of the idiom blend condition). Two response variables were investigated: the acceptability rating responses and the reaction times of those responses. All models analyzing these response variables are summarized in Appendix B.

Ten predictor variables were included in the models. First, `Condition` is a factor indicating the condition in which the idiom occurred (e.g. canonical form, lexical variation, idiom blend). `Length` specifies the number of words within the

idiom's canonical form. `PairedIdiomSemantics` is a factor specifying whether the two idioms used in the formal idiom blend have similar or different semantics (e.g. *spill the beans* & *let the cat out of the bag* = similar; *shoot the breeze* & *hold your tongue* = different), while `PairedIdiomSyntax` is a factor indicating whether the two idioms in the formal idiom blend have similar or different syntactic structures (e.g. *spill the beans* & *let the cat out of the bag* = different; *shoot the breeze* & *hold your tongue* = similar). `TrialScaled` is the scaled order of presentation of the stimuli in the experiment. Since the stimuli was presented randomly, this order will be different for each participant.

`PC1.logFrequency` is the first principal component (PC) of a Principal Components Analysis (PCA) on the log co-occurrence frequencies of the idiom. PCA "reduces the number of dimensions required for locating the approximate position of the data points" (Baayen, 2008, p.120). The data, on which the PCA was carried out, includes 120 rows (i.e. idioms) and 29 columns (i.e. frequency measures); thus in this data, there were 120 points in a 29 dimensional space. The 29 frequency measures included frequency of the whole idiom, frequencies of the individual words, and all possible combinations of adjacent words (e.g. word1 and word2; word2 and word3; word1 and word2 and word3); see Appendix E for details. The data points remain fixed at their location, while new axes, or dimensions, are determined from the original 29, in order to remove superfluous dimensions. These new dimensions are called Principal Components (PCs) – they are uncorrelated, more succintly account for the data, and are ranked by the proportion of variance explained. In the models below, only the first Principal Component (PC1, henceforth `PC1.logFrequency`) was significant. The highest loadings on this PC were the co-occurrence frequencies associated with the longest idioms (e.g. the frequency of words 5 and 6 occurring together – *the grindstone* in the idiom *keep your nose to the grindstone* – or the frequency of words 4, 5, and 6, *to the grindstone*). This latent variable proved to be a more effective predictor than the log frequency of the idiom as a whole, evidenced by a lower AIC value[3] produced during model comparison (i.e. one model containing the log frequency of the whole idiom and one model containing `PC1.logFrequency`).

[3]AIC, or Akaike's Information Criterion, measures how well the predictor variables improve the predictiveness of the model (cf. Akaike, 1974, 1985). This measure is used as a form of model comparison, by measuring the difference of the entropies of the models (i.e. the negative of the entropy is often regarded as the amount of information), given the number of parameters in the model, by means of a log likelihood ratio test; thus the model with the smallest AIC value is the preferred model.

Several participant variables were also included in the models. `KnowIdiom` is a factor indicating the participant's knowledge of the idiom (i.e. yes or no). `IsL1English` is a factor specifying whether the participant declared English as their first language(s). `HowOftenUseIdioms` is a scaled rating between 0 and 100 for how often the participant uses idioms in general. `LQ3` is a scaled rating between 0 and 100, indicating the participant's acceptability of Language Question 3 (i.e. *The document had to be signed by both Susan and I.*). In order to assess the Language Questions (LQs), I ran a Principal Components Analysis on the seven LQs. LQ3 emerged as diagnostic, as it had the highest loadings on PC1 (i.e. `PC1.LQ`). However, the original variable, `LQ3`, produced a lower AIC value than `PC1.LQ` and is therefore included in the models.

Finally, `meanTransparencyRating` is a scaled average rating for the transparency (or clarity) of the idiom's meaning as a whole. Since speakers differ in how they interpret the decomposability of idioms, as evidenced by the low reliability of the decomposability classification task (cf. Titone and Connine, 1994b), I collected a measure of how clear or obvious people find the meaning of the expression 'as a whole'. This measure then, may provide some indication of how literal or figurative people consider the idiom to be, and is in line with other proposals of an idiomaticity continuum, ranging from more collocational at one end to more idiomatic at the other (cf. Fillmore et al., 1988; Wulff, 2009; Taylor, 2012). These ratings were collected in a separate experiment, specifically designed to elicit ratings of transparency. Those participants saw each idiom, along with a definition and an example sentence, and were asked to rate the transparency of the idiom using a VAS scale from 0 (not transparent) to 100 (transparent), see Chapter 2 for a further discussion. The average rating for each idiom was included as a separate predictor in the models below to determine whether transparency influences people's preferences of variation.

### 3.3.2 Acceptability Rating Responses

A model with all predictors revealed a significant three-way interaction between `KnowIdiom`, `Condition`, and `HowOftenUseIdioms` (model not shown). For ease of interpretation, I therefore created two separate models: one modelling the ratings of known idioms and one modelling the ratings of unknown or unfamiliar idioms. These two models are shown in Appendices B.1 and B.2.

First consider the results for those participants who knew the idiom. The interaction between `Condition` and `HowOftenUseIdioms` is shown in the top-left panel of Figure 3.1. The canonical form of the idiom is rated as more acceptable by participants who use idioms more often. This same pattern is observed when an additional concept is integrated into the idiom; the slope is less steep, but not significantly different from the canonical form (t = -1.47), and therefore shown in grey in Figure 3.1. Meanwhile, idioms in the three conditions: lexical variation, formal idiom blends, and the literal meaning show the reverse pattern, significantly lower acceptability ratings when participants use idioms more often. These variants are rated more similarly to the canonical form when the participant does not use idioms often. Finally, a partial form of the idiom has significantly lower acceptability ratings compared with the canonical form regardless of whether the participant uses idioms often or not.

Now consider the participants who reported that they did not know the idiom. As expected, these participants are not affected by the different conditions (i.e. this variable is not significant as a main effect or in any interaction). These participants are not familiar with the canonical form and therefore do not recognize how it has been altered. They are however, still influenced by the amount they use idioms in general, and tend to rate the idiom as less acceptable if they use idioms often. This result likely emerges because they do not know this particular idiom despite having greater experience with and exposure to a variety of different idioms, because they use idioms often. In other words, they rate it as less acceptable because they are not familiar with the expression, not because of its form.

Returning to the participants who know the idiom, several additional predictors are significant in the model. `Length` is significant in an interaction with `Condition`, shown in the top-right panel of Figure 3.1. Participants tend to rate idioms as less acceptable if they are longer. This pattern includes most variants. Integrated concepts, lexical variation, and formal idiom blends have slopes which are not significantly different from the canonical form and are therefore depicted in grey. The literal meaning and the partial form however, show a different pattern. These two conditions are rated as more acceptable if the idiom is longer. This is understandable for partial forms, which would include more words alluding to the idiom when the idiom itself is longer (e.g. *driving them up it* for *driving someone up the wall* vs. *lose it* for *lose your marbles*). The greater number of words provides additional

context and information to understand the idiom's meaning, allowing it to be judged more favorably. Longer idioms might also provide additional context and information to help speakers interpret the literal meaning of the idiom as well. This finding is not dependent on the transparency of the idiom, as `meanTransparencyRating` does not occur in a significant interaction with `Condition`, `Length`, or both (i.e. a three-way interaction). Apparently, longer utterances simply invite a more literal reading.

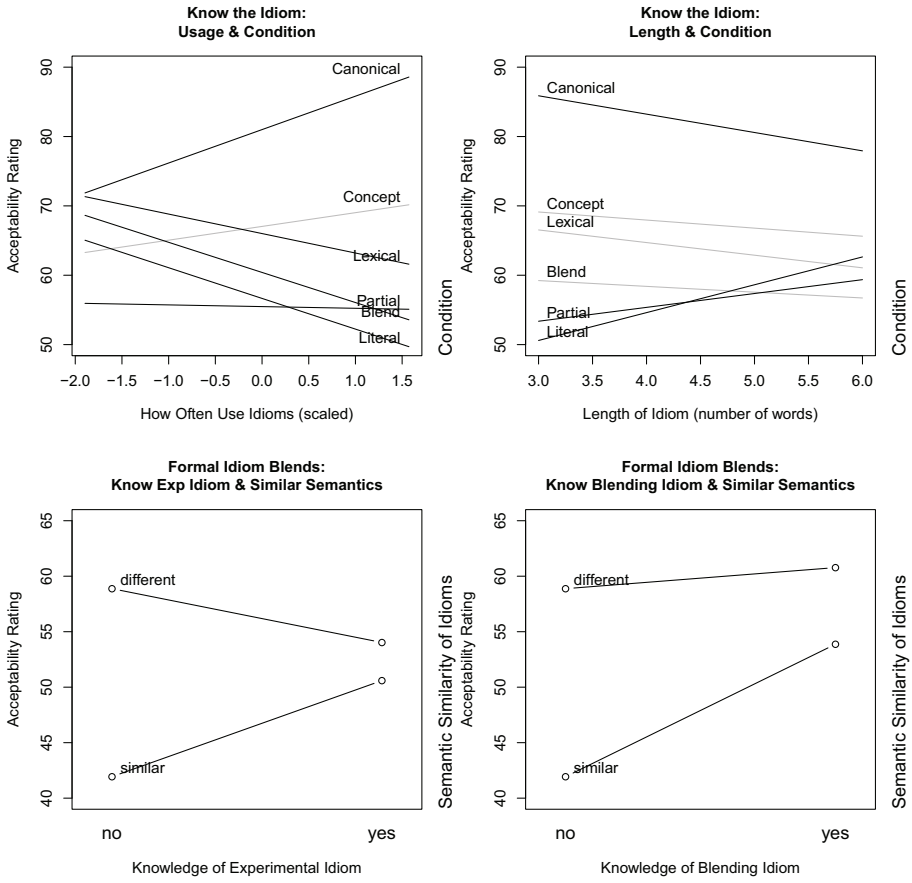


Figure 3.1: Interactions in the Mixed-Effects Linear Regression Model for the Acceptability Ratings of Idiomatic Variation. Lines in grey represent non-significant slopes.

The last two interactions observed for the acceptability rating responses focus specifically on the idiom blend condition. These interactions are shown in the bottom

two panels of Figure 3.1, while the models are presented in detail in Appendix B.3. Both interactions are between the participant's knowledge of the idiom (i.e. whether they know the idiom or not) and the semantics of the two idioms merged in the formal idiom blend (i.e. whether the semantics are similar or different). However, the interaction observed in the bottom-left panel of Figure 3.1 is with the knowledge of the experimental idiom, whereas the interaction in the bottom-right panel of Figure 3.1 is with the knowledge of the blending idiom. Both of these interactions indicate that formal idiom blends are rated as less acceptable when the semantics of the two idioms are similar, regardless of whether the speaker knows the idiom. Participants are significantly more likely to rate blends with similar semantics with a lower acceptability rating if one of the idioms is unknown. The blends which contain idioms that have different semantics tend to receive higher acceptability ratings, and although slight differences in the ratings emerge depending on whether one knows the idiom, these differences are not significant for experimental idioms ($t = -1.34$) or blending idioms ($t = 0.52$). A three-way interaction between these variables (i.e. knowledge of both idioms and the semantic similarity of the idioms) is not significant, which suggests that people only need to be unfamiliar with one of the idioms in order to evaluate the semantically similar idiom blends considerably less acceptable. This observation may be explained by the fact that these blends are more noticeable to those unfamiliar with one of the expressions, and since they stand out more, speakers find them less acceptable. Those who know both idioms are presumably able to access the meaning of the blend, as they are familiar with the idioms from which the parts belong, and therefore are not as surprised or confused by the blends. Additionally, these idioms may be judged more like errors, to those who know the expressions and especially to those who do not, than the blends that contain idioms which are semantically different, potentially because they vary along the paradigmatic axis (Algeo, 1977; Taylor, 2012), a sentiment also conveyed in previous research (cf. Fay, 1982; Cutting and Bock, 1997; Kuiper et al., 2007).

Two additional main effects are observed for all participants (i.e. those who know the idioms and those who do not). Higher acceptability ratings are given to idioms judged to be more transparent. This finding was not affected by the condition in which the idiom occurred – the higher the mean transparency rating of the idiom, irrespective of whether the idiom occurred in its canonical form or as a variant, the higher the acceptability rating. In other words, people find idioms more

acceptable if they are more clearly interpretable, regardless of their form. Second, participants who rated LQ3 (i.e. *The document had to be signed by Susan and I*) as more acceptable, also rated idioms, in both their canonical forms and their variants, as more acceptable. An interaction with `Condition` is not significant, so people are simply more accepting of idiomatic expressions if they are more flexible with language. This sentence contains an 'incorrect' form of first person singular (i.e. subject case, *I*, where object case, *me*, would prescriptively be required). This form, while incorrect according to the grammar of English, still allows one to access the correct information, namely the person being referenced; thus, a greater acceptance (or flexibility) of another form to convey the same meaning, the more acceptable idiomatic expressions are in general, even if they include variation.

A more complex random-effects structure is observed in the model for those who know the idiom, as expected. All models have by-Subject random slopes for `TrialScaled`, indicating that participants varied as to how much their responses differed as they proceeded through the experiment. However, the model for those who know the idiom also had by-Item random slopes for `Condition` with correlation parameters, suggesting that speakers rated the idioms differently depending on the condition in which they occurred; people prefer certain idioms with specific types of variation. Speakers acquire this preference and experience when they become familiar with an idiom, which is why this same finding is not observed for those who do not know the idiom.

### 3.3.3 Acceptability Reaction Times

I also analyzed the reaction times (RTs) for how quickly the participants responded to the acceptability ratings, shown in detail in Appendix B.4. Faster reaction times are associated with easier judgements of the idiom's acceptability. Only one interaction, between `KnowIdiom` and `Condition`, is significant in this model, illustrated in Figure 3.2. The RTs associated with each condition are similar for both those who know the idiom and those who do not. Significantly longer RTs are observed with integrated concepts, while significantly shorter RTs are observed with partial forms. These results may simply reflect the fact that the integrated concept condition has an additional word inserted into the idiom, whereas the partial form condition has a word omitted from the expression. This reaction time difference therefore likely

corresponds to length of the expression. The RTs for the literal meaning and the idiom blend conditions are also similar for both those who know the idiom and those who do not. Only two differences are observed. First, those participants who do not know the idiom tend to rate lexical variation faster than those who know the idiom, but not significantly faster (t = -1.26). Second, and more importantly, participants are significantly faster rating the canonical form of the expression if they know the idiom (t = -3.58). For those who do not know the expression, the RT to rate the canonical form does not differ significantly from variants of the same length: formal idiom blends, idioms with lexical variation, and idioms within a literal context. These results illustrate that the canonical form has an advantage if it is familiar, but that all other types of variation are rated as quickly as if one does not know the expression, as long as the length of the variant is taken into account.

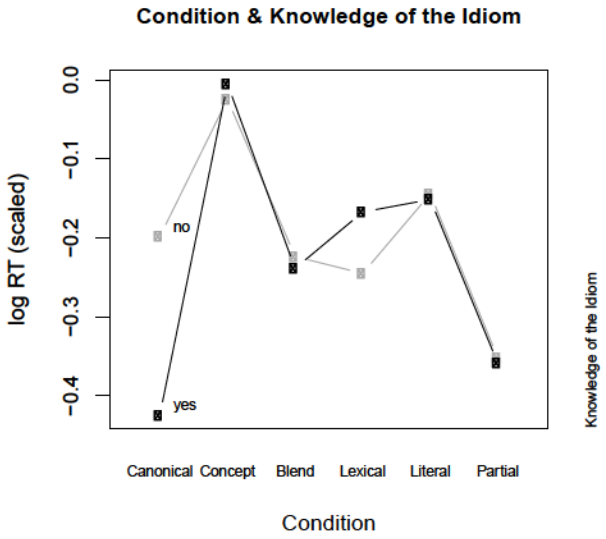


Figure 3.2: Interaction in Mixed-Effects Linear Regression Model for the Reaction Time to Rate the Acceptability of Idiomatic Variation.

Four additional main effects are observed in both models on RTs (i.e. RT to rate the acceptability for all idioms and the model specific to formal idiom blends shown in Appendix B.5). Participants have faster RTs the more they use idioms in general and the greater the idiom's co-occurrence frequencies. Participants who acquired English as a first language are significantly faster at responding than those

who acquired it as a second language, despite the fact that the latter acknowledge a native or near-native competency in English and show no difference in their acceptability ratings in the previous section. Finally, participants' RTs become faster the further they advance through the experiment. However, both models show by-Subject random slopes for `TrialScaled`, indicating that participants vary as to exactly how much faster they become.

## 3.4 Discussion

This study investigated speaker's preferences and acceptability of multiple types of idiomatic variation. Several findings have emerged about each type of variant. First, integrated concepts seem to be the most preferred type of variant, generally receiving the highest acceptability ratings, except from those who do not use idiomatic expressions often. This variant is similar to the adjective and adverb insertion variants discussed in Gibbs and Nayak (1989) and Tabossi et al. (2008). These two studies separated the type of concept by word class and found that inserted adverbs have higher similarity ratings to their literal paraphrases. However, these adverbs were positioned, at least for Gibbs and Nayak (1989), before the verb (i.e. before the idiom), as in *The boss will quickly lay down the law if anyone shows up late.* Their participants found these most similar in meaning to the literal paraphrase, but they probably preferred these variants more since they disrupted the idiom less. A similar finding is observed in this study; people prefer integrated concepts more presumably because they disrupt the idiom less – they always occur inside the idiom, but the idiom still occurs in its full form, with the canonical lexical items, and is only modified by adding additional information or extra specification to the figurative context. People prefer elabortation of the context, especially if they are familiar with the idiom and enjoy using idioms, over modification to the form and lexical items of the idiom.

If the form is altered in some way – a synonymous word, formally blended with another idiom, or used in a non-figurative way – then the variant tends to be judged as less acceptable, especially if one knows the expression and enjoys using idiomatic expressions in general. This is most likely due to the fact that the form varies from the canonical form. Literal uses of the idiom are rated more favourably when the expressions are longer, but are not found nearly as acceptable as when the canonical

form occurs in a non-literal context. This result may be because additional information or cues are available in the expression for it to make meaningful literal sense. Meanwhile, formal idiom blends are less preferred when the two idioms included in the variant are semantically similar. These blends are probably regarded as errors, and perhaps larger, more consequential errors by those who do not know one of the expressions.

Partial forms of idioms are predominantly rated as less acceptable. This variant receives low ratings from all participants, regardless of whether they enjoy using idioms or not, but did show a significant increase in acceptability when the canonical form of the expression is longer (cf. *hear something through the grapevine* vs. *cut the mustard*). When the idiom occurs in a truncated form, less information associated with the original expression is available, potentially making it more difficult to interpret, and therefore judged to be less acceptable. However, when more of the expression is available (i.e. there are more words in the variant because the expression itself is longer), participants rate the variant more favourably possibly because it is easier to interpret. The more cues that are available to interpret the meaning of the expression, the more preferred the variant.

The beginning of the expression (i.e. the verb) and the end of the expression (i.e. the noun) were both manipulated in this study. This variable however, was not significant in any of the models. This suggests that either the noun or the verb can be modified – one is not preferred more than the other – but that both are less preferred than the canonical form. Gibbs et al. (1989a) reported a similar finding, that either the noun or the verb could be lexically varied (i.e. there was no significant difference between the two), but that this change was considered less semantically similar to the literal paraphrase than the canonical form, and even less similar if both the noun and verb were changed. Thus, changing one element of an idiom comes with a cost, a dis-preference, but changing more than one element is considered even worse.

People who know the idiom not only show varying preferences between the different conditions, they also show preferences for particular idioms in specific conditions, as illustrated by the random effects structure of the rating responses for those participants who know the idiom. Learning the idiom and gaining experience with its use biases the speaker to prefer the idiom with certain variants. Thus, not only are some conditions preferred more than others, but these conditions also depend on which idiom they occur with. Those who know the idiom also show an advantage

when rating the canonical form, as expected. They are significantly faster rating the canonical form than other types of variants, which is not observed for those who do not know the idiom. Both groups show the same overall pattern – integrated concepts and partial forms of the idiom take significantly longer and shorter, respectively, to rate, compared with the other variants, reflecting the number of words in the expression. Only the canonical forms show an advantage, they are the forms people are most familiar with and tend to be preferred over a modified form.

Idioms also tend to be rated with higher acceptability ratings when they are considered to be more transparent. Different participants rated the transparency of each idiom in its canonical form. The mean rating for each idiom from that experiment was used as a predictor variable in this study, representing approximately how clear or interpretable speakers regard each idiom's meaning to be. When speakers consider the expression to be more transparent, they then find the idiom to be more acceptable, regardless of the form in which the idiom occurs. There is no significant interaction between the transparency of the idiom and the condition in which it occurs; therefore, there is no evidence or support for speakers preferring variation with idioms deemed more transparent. This finding contrasts with Gibbs and colleagues (Gibbs et al., 1989a; Gibbs and Nayak, 1989) who found support for lexical and syntactic variation with decomposable idioms, or idioms whose component parts contribute meaning to the meaning of the whole. However, Gibbs et al. (1989a) only found support for lexical variation if they removed abnormally decomposable idioms from the analysis, while Tabossi et al. (2008) did not find support that syntactic variation was preferred with decomposable idioms. Additionally, these studies asked participants to rate the semantic similarity of a variant with a literal paraphrase, whereas this study was explicitly asking for people's preferences. If we assume that nondecomposable idioms are also less transparent, then these findings suggest that people do not have a preference for variation with only transparent expressions, but instead find a difference in the similarity between an idiom and a paraphrase.

The fact that there is no significant interaction between transparency and condition suggests that speakers simply like idioms more when they are viewed as more transparent. Interestingly, one of the strongest predictors of transparency is knowledge of the idiom, see Chapter 2. When speakers are familiar with the expression and know its meaning, they tend to rate the expression as more transparent than if they do not know the expression. Other research has shown similar findings; for

example Keysar and Bly (1995) taught half their participants the original meaning of unfamiliar idioms, like *the goose hangs high* 'things look good', and the other half of the participants the conceptual opposite 'things look bad'. They then asked the same participants to rate the transparency of one of the possible meanings of the expression (i.e. either the one they learned or the opposite one), as if they were rating the meaning from the perspective of "an uninformed individual". They found that participants were more likely to rate the learned meaning, regardless of whether it was the original meaning or its conceptual opposite, as more transparent. Thus, speakers find idioms to be more transparent when they know their meaning, and in turn find idioms to be more acceptable, regardless of their form, when they are regarded as more transparent.

# CHAPTER 4

# Comprehension of Idiomatic Variation

## 4.1 Introduction

Research on idioms has predominantly focused on the canonical form (cf. Swinney and Cutler, 1979; Gibbs, 1980; Cacciari and Tabossi, 1988; Titone and Connine, 1999; Fanari et al., 2010). But recent research using corpus-based methods has demonstrated that idiomatic variation is more common than initially assumed (Moon, 1998; Schröder, 2013; Duffley, 2013). Idioms can occur with lexical variation (e.g. *throw/toss in the towel, a piece/slice of the action, hard/close/hot on the heels of someone*), truncations (e.g. [*he who pays the piper*] *calls the tune, make hay* [*while the sun shines*]), and even aspectual variation (e.g. *cross one's fingers/keep one's fingers crossed*). Idioms have also been shown to occur with adverbial and adjectival modification (e.g. *spill royal beans, pull innumerable strings, pulling political strings, make rapid headway*). These types of variation have even been shown to occur with idioms whose constituents do not contribute meaning to the whole (e.g. *when his parents kick their gold-plated bucket*, *shoot the warm April breeze over burgers*, and even *My 5-year old Dell kicked the can a few weeks ago*).

Few studies have explored the comprehension of idiomatic variation. Gibbs and colleagues (Gibbs et al., 1989a; Gibbs and Nayak, 1989) investigated the lexical and syntactic flexibility of idioms (i.e. the degree to which an idiom can alter its lexical items or syntactic structure and still maintain its idiomatic meaning). They were particularly interested in whether a difference in flexibility was observed between decomposable and nondecomposable idioms. Semantic decomposability has been described as the relationship between the idiom's component parts and the meaning

of the whole (Nunberg, 1978). In this classification scheme, 'Normally Decomposable Idioms' have constituents whose literal meaning directly contributes to the idiomatic meaning, such as *pop the question*. Here, *the question* refers to a specific question, a marriage proposal, while *pop* refers to the act of suddenly uttering that proposal. 'Abnormally Decomposable Idioms', as in *carry a torch*, have constituents which indirectly contribute meaning to the whole through a figurative relationship. *The torch* is a conventional metaphor used to convey warm feelings of love for someone, while *carry* is used to convey the act of feeling in that way. Lastly, 'Nondecomposable Idioms' like *kick the bucket* have constituents which do not contribute meaning to the whole. The individual words in this idiom do not literally or metaphorically contribute meaning to the expression; there is no relationship between *kick* or *bucket* and 'die'.

Gibbs and colleagues (Gibbs et al., 1989a; Gibbs and Nayak, 1989) found that decomposable idioms received higher similarity ratings to their literal paraphrases for both lexical and syntactic variation than nondecomposable idioms. However, these studies implemented a similarity ratings task to access the participant's understanding of the variants; thus, they may have observed the participant's preferences or acceptability of the variants, and not necessarily the degree of difficulty in understanding the variants. Furthermore, Tabossi et al. (2008) conducted a replication study of Gibbs and Nayak (1989) and found that abnormally decomposable idioms received reliably lower semantic similarity ratings than both normally decomposable and nondecomposable idioms, but that the latter two were not significantly different. In other words, normally decomposable idioms are not reliably rated as more flexible than nondecomposable idioms.

Other studies have also found unreliable results for the semantic decomposability classification. Following the same procedure described in Gibbs and Nayak (1989), Titone and Connine (1994b) found that only 40% of 171 idioms were reliably grouped into the three categories using a 75% agreement rate in two separate classification tasks. Meanwhile, Tabossi et al. (2008) calculated the proportion of agreement between participant judges using a binomial test and found that classification into the decomposable and nondecomposable categories, as well as further classification into normally and abnormally decomposable categories, was not significantly different from chance. The unreliability in classifying semantic decomposability certainly raises questions as to the reliability of this compositionality measure.

Meanwhile, McGlone et al. (1994) investigated the comprehension of semantic productivity. Variants were created in this study that produced a new idiomatic meaning based on the original (e.g. *shatter the ice*, from *break the ice*, meaning 'to break an uncomfortable or stiff social situation in one fell swoop'). Using a self-paced reading task, participants read sentences of stories, with the idiom or its variant presented in the final sentence, in either a specific or general context. They found that the original idioms were read faster than the variants, and that all idioms were read faster in the specific contexts. Variants then behaved like the original, but took extra time to read. In a second experiment, they included literal paraphrases of the idiom and of the variant and found that the original idiom was read faster than the variants and the literal paraphrases, but that the variants were read as fast as the literal paraphrases. They suggest that individual word meanings are activated when the idiomatic meaning is activated and accessed, even with opaque idioms like *kick the bucket*, allowing for semantic productivity to be utilized and understood.

While McGlone et al.'s (1994) study clearly shows that idioms can be processed when modified in context, they did not control for the different types of variants that they used. The variants included lexical variation playing on the literal meaning of a word (e.g. *shatter the ice*), quantification of a noun in the idiom (e.g. *not spill a single bean*), and even hyperboles (e.g. *it's raining the whole kennel*). It is therefore uncertain whether some types of variation are easier to interpret than others.

The current study further explores the processing of several types of variation, as well as the literal meaning of the idiom, through an eye-tracking study. Three research questions are explored in this study, summarized below. The first question plans to determine whether variants are still processed significantly different from the canonical form when the type of variation is controlled for. For example, is lexical variation more difficult to comprehend than the original idiom? What about partial forms of the expression? Second, I will compare the processing of these variants to the literal meaning of the idiom, as it is used in context, rather than a paraphrase of the literal meaning. By changing the context in the literal meaning condition, so that the words in the expression are processed literally, not idiomatically, how do variants compare to processing the literal meaning of the expression? Are they still read just as fast? Finally, by including several types of variation and controlling for them, a comparison can be made between the different types of variants. Are there processing differences between these variants?

Research Questions:

1. Are variants processed differently from the canonical form?
2. Are variants processed differently from the literal meaning of the idiom?
3. Are different types of variation processed differently from each other?

## 4.2 Methodology

### 4.2.1 Materials

Sixty idioms were extracted from the Oxford Dictionary of English Idioms (Ayto, 2009) and the Collins COBUILD Idioms Dictionary (Sinclair, 2011). These idioms varied in length and syntactic structure: 20 three-word idioms consisting of a verb and a noun phrase (i.e. V-NP, e.g. *rock the boat*); 20 four-word idioms consisting of a verb and a prepositional phrase (i.e. V-PP, e.g. *jump on the bandwagon*); and 20 five- and six-word idioms (10 each) consisting of a verb, noun phrase, and a prepositional phrase (i.e. V-NP-PP, e.g. *hear something through the grapevine*). Two contexts were created for each idiom: one literal and one figurative (e.g. *I used to pretend I could talk to plants, and I would hear things through the grapevine* = literal; and *I used to be a socialite, and I would hear things through the grapevine* = figurative). Both contexts had identical final clauses, with the idiom in sentence-final position.

These idioms were manipulated for four types of variation within the figurative context, in addition to the canonical form. First, lexical variation, where one of the lexical items within the expression was altered to a synonymous or near-synonymous word (e.g. *rock the canoe* or *discover something through the grapevine*). An online thesaurus was often utilized for synonymous words[1]. Second, partial form of the idiom, where only a portion of the idiom was presented, usually a key word or words (e.g. *rock things* or *use the grapevine*). In order for the sentence to still be grammatically correct, pronouns or lexically-vague words were used to replace the

[1] http://www.thesaurus.com/

Table 4.1: Four types of blends used in the Idiom Blend condition

| **Type of Blend** | **Example** | **Source Idioms** | **Total** |
|---|---|---|---|
| sSYN, sSEM | *rock the applecart* | *rock the boat*<br>*upset the applecart* | 15 |
| sSYN, dSEM | *shoot your tongue* | *shoot the breeze*<br>*hold your tongue* | 15 |
| dSYN, sSEM | *pass the mustard* | *cut the mustard*<br>*pass muster* | 15 |
| dSYN, dSEM | *face the wringer* | *face the music*<br>*put through the wringer* | 15 |

missing elements of the expressions, such as *it, them, things* for nouns, or *have, be, do, use* for verbs. Third, integrated concept, where an additional concept was integrated into the idiom (e.g. *rock the initiative boat* or *hear something through the judgemental grapevine*). These additional concepts expanded or emphasized the figurative contexts in which the idiom occurred. Finally, formal idiom blend, where two idioms were blended together (e.g. *rock the applecart* – blending *rock the boat* with *upset the applecart*, or *get wind through the grapevine* – blending *hear something through the grapevine* with *get wind of something*). Each experimental idiom (i.e. one of the 60 idioms selected) was paired with a non-experimental idiom for use in the idiom blend condition. These "blending" idioms were chosen for their intuitive plausibility, but controlled for their syntax and semantics with the experimental idioms (cf. Cutting and Bock, 1997). Four types of blends were created: same syntax, similar semantics (sSYN, sSEM); same syntax, different semantics (sSYN, dSEM); different syntax, similar semantics (dSYN, sSEM); and different syntax, different semantics (dSYN, dSEM), illustrated in Table 4.1. Five instances of each type of blend occurred with the three 'syntactic types' (i.e. V-NP, V-PP, or V-NP-PP), totalling 15 of each blend type.

Half of the idioms had the beginning portion of the expression altered, while the other half of the idioms had alternations made to the final portion of the expression. In total, there are six conditions: one in a literal context and five in a figurative context (i.e. one canonical form and four variants). The experiment utilized a Latin-square design, where every participant saw each idiom once in one of the six conditions. Therefore, six versions of the experiment were created, each one containing 10 idioms in each of the six conditions.

CONDITIONS:

1. **Literal Meaning** of the idiom in its canonical form
   (e.g. *While the guys were reshingling, they suddenly went through the roof.*)
2. **Canonical Form** of the idiom in a figurative context
   (e.g. *Although these were new stocks, they suddenly went through the roof.*)
3. **Lexical Variation** of the idiom in a figurative context
   (e.g. *Although these were new stocks, they suddenly went through the ceiling.*)
4. **Partial Form** of the idiom in a figurative context
   (e.g. *Although these were new stocks, they suddenly went through it.*)
5. **Integrated Concept** within the idiom in a figurative context
   (e.g. *Although these were new stocks, they suddenly went through the investment roof.*)
6. **Idiom Blend** of two idioms in a figurative context
   (e.g. *Although these were new stocks, they suddenly went through the charts.*)

Since the "blending idioms" only occurred in one condition (i.e. Idiom Blend), they were used as fillers in their canonical form in the other five versions of the experiment, occurring in either a figurative or a literal context. Each blending idiom was excluded as a control in the version of the experiment where it occurred in the idiom blend condition in order to avoid a bias in the materials. Therefore, in each version of the experiment, 10 of the blending idioms occurred in a formal blend in the idiom blend condition, while the remaining 50 appeared in their canonical form as fillers. Of these fillers, 20 occurred in a figurative context, while 30 occurred in a literal context. This was done to increase the number of literal contexts in the experiment so that they were not so underrepresented. In sum, each participant saw 110 items: 60 experimental idioms (10 in each of the six conditions) and 50 blending idioms as fillers (20 in a figurative context and 30 in a literal one).[2]

Finally, six practice sentences were created using six "practice" idioms. These idioms all occurred in their canonical form. Three were in a figurative context and three in a literal context. These were the same for all participants.

[2]I thank Lauren Rudat for all her helpful comments and suggestions on improving the stimuli in this experiment.

### 4.2.2 Procedure

This experiment used the Eye-Link 1000, desk-top mounted video-based eye-tracking device, manufactured by SR Research. The participants placed their chin on a chin rest during the experiment, for stability and to reduce any movements they might make during the experiment. The eye-tracker sampled the pupil location and size at a rate of 1000Hz, and was calibrated using a 9-point calibration grid. Calibration occurred at the beginning of the experiment, after the practise, and again after every 22 sentences, for a total of five blocks. Auto-thresholds for pupil and corneal reflection were used, along with the centroid pupil tracking mode. The resolution of the computer screen was set to 1920 x 1080 pixels.

The stimuli were presented in two parts. Participants first saw the "context clause" (i.e. the clause containing the context; e.g. *Although these were new stocks,*), followed by the "idiom clause" (i.e. the clause which contained the idiom; e.g. *they suddenly went through the roof.*) on a separate screen. Each trial began with a fixation cross presented for 1,000 msec on the left side of a light-grey screen, at 140 x 400 pixels, in a bold, black, Courier New 30-point font. This was followed by a blank, light-grey screen for 500 msec. The context clause was presented next, also on a light-grey background in a bold, black, Courier New 30-point font. Every context clause was displayed in full and fit on one line, beginning at 140 x 400 pixel location. To exit this screen, participants had to trigger an invisible boundary in the bottom right corner of the screen. The invisible boundary was a rectangle shape, 100 x 150 pixels in size, located at 1820 x 930. The screen remained for 2,000 msec before disappearing. A blank, light-grey screen was presented for 1,000 msec before the fixation cross preceding the idiom clause appeared. The sequence of screens for the idiom clause were identical to the context clause. The trial ended with the final blank screen lasting for 3,000 msec.

Ten percent of the stimuli were followed by a true/false comprehension question, which pertained to the immediately preceding sentence (i.e. both context and idiom clauses) and were presented randomly throughout the experiment. Participants pushed one of two buttons on a game controller to answer the comprehension questions, which were clearly labelled on the question screen. The experiment began with a practise session, which consisted of six practise sentences and three questions. These were the same for all participants, although their order varied.

All participants had normal or corrected-to-normal vision. The right eye of each participant was tracked in this experiment. Participants sat approximately 85cm from the computer screen, with the camera placed on the desk approximately 50cm in front of the participant, between them and the computer screen. The participants sat in a sound-proof booth, while the experimenter sat outside the booth, running the experiment. The experimenter and participant were able to communicate through a Morantz audio system (i.e. both the booth and the experimenter station outside the booth were equiped with a microphone and a speaker). The experimenter's microphone was turned off during each experiment block and turned on again during calibration. The participant's microphone was never adjusted, so that they could always be heard during the experiment if needed. The lights were kept on. The experiment was self-paced and took about 45 minutes to complete. Each participant was given an opportunity for a short break half-way through the experiment.

After the participants had participated in the eye-tracking portion of the experiment, they were then asked to indicate their knowledge of each expression in a separate task. Each idiom appeared, in its canonical form, in a black, bold, 22-point Courier New font, centered on a white background. Above the idiom was the question "Do you know this expression?"; below were two boxes, one labelled 'yes' and the other labelled 'no'. Using the mouse, participants clicked on the appropriate box to respond. The mouse repositioned itself to the center of the screen on each trial.

At the end of this second task, participants were presented with a few additional questions. They were asked three questions pertaining to their idiom usage: (1) How often do you use these expressions?; (2) How often do others around use these expressions?; and (3) Do you like using these expressions? Below each question was a Visual Analogue Scale (VAS), which is a continuous graphical rating scale (cf. Hayes and Patterson, 1921; Freyd, 1923; Funke and Reips, 2012). VAS scales allow fine gradations to be measured by converting the pixels on the computer screen to a possible numeric value, in this case, ranging between 0–100. Participants thus responded to these questions by clicking the mouse anywhere along the VAS scale. The scale was labelled with a 'thumbs-up' image on the right for a positive response and a 'thumbs-down' image on the left for a negative one. Lastly, participants were asked to rate the acceptability of seven prescriptively 'incorrect' sentences, shown below, using the same VAS scale. These sentences attempted to elicit a measure of

the participant's flexibility with language and non-standard usage.

Language Questions (LQ):

1. The only option the school board has is to lay off a large amount of people.
2. Slot machines are thought to be more addicting than table games.
3. The document had to be signed by both Susan and I.
4. While cleaning the kitchen, Sally looked up and saw a spider on the roof.
5. I thought it could've went either way.
6. She could care less what he had to say about it.
7. You have to balance your life, irregardless of what anybody thinks.

### 4.2.3 Participants

Sixty University of Alberta linguistics undergraduate students participated in this experiment. All were native speakers of English, predominantly monolingual speakers of Canadian English. This study only focused on native speakers of English for two reasons: firstly, idiomatic expressions are typically one of the last types of expressions people learn when learning a second language, and secondly, these expressions were altered in form. In other words, not only would they be relatively new and unfamiliar to non-native speakers, but they would also be more difficult given the range of variations used in this experiment. Since this study is focused on the comprehension of variation, only native speakers were included as they would be more consistent to compare. There were 43 female and 17 male participants, ranging from 17–29 years of age. Four participants were left-handed. All participants were reimbursed for their time with course credit.

## 4.3 Results

### 4.3.1 Variables

The results were analyzed using mixed-effects regression, using the `lme4` package (Bates et al., 2014) in `R` (R Core Team, 2014). Three response variables were inves-

tigated: Total Fixation Duration (i.e. the total amount of time spent fixating on the Area Of Interest, or AOI), First Fixation Duration (i.e. the first pass on the AOI), and Fixation Count (i.e. the number of times the participant fixated on the AOI). The first two response variables are numeric data, whereas the last response variable is count data. All variables were analyzed using linear regression, but the count data were assumed to have a poisson distribution. Two Areas Of Interest (AOIs) were analyzed: the idiom as a whole (i.e. the fixations on all words within the idiom were summed), and the altered word within the idiom (i.e. the synonymous word in lexical variation, the additional word in the integrated concept, the semantically vague 'replacement' word in partial forms, and the word from another idiom in formal idiom blend). Besides the idiom blend condition (where the 'blending idioms' are part of the variant), the analyses focus on the 60 experimental idioms. All models analyzing these response variables are summarized in Appendix C.

Ten predictor variables appeared significant in the models. `Condition` is a factor indicating the type of variation with which the idiom occurred (e.g. lexical variation, partial form), or whether the idiom occurred in its canonical form. `Length` specifies the number of words within the idiom's canonical form. `PortionIdiomAltered` is a factor specifying the portion of the idiom (i.e. beginning or ending) which was manipulated in the variant. The portion not manipulated was preserved from the original 'experimental' idiom. `TrialScaled` is the scaled order of presentation of the stimuli in the experiment. As the stimuli was presented randomly, this order will be different for each participant.

`MeanVariationRating` is a scaled mean measure of acceptability for the particular idiom with a specific type of variation. This measure was collected in a separate experiment (see Chapter 3), where participants were asked to rate the acceptability of the variants in the same contexts using a VAS scale ranging from 0 (not acceptable) to 100 (acceptable). These ratings were included to determine if participants' preferences of variants influence their ease of comprehension. Meanwhile, `meanTransparencyRating` is a scaled average measure of transparency for the idiom's meaning as a whole. I collected a measure for how clear or obvious the meaning of the expression is considered 'as a whole'. These ratings were collected in a separate experiment (see Chapter 2), which was specifically designed to elicit ratings of transparency. Participants in that study saw each idiom, along with its definition and an example sentence, and were asked to rate how obvious the meaning of the

expression was using a VAS scale from 0 (not transparent) to 100 (transparent). The average rating for each idiom was included as a predictor to determine whether the transparency of the idiom influences speakers processing of idiomatic variants.

The variables `LSA.Score.Paraphrase` and `LSA.Score.Definition` are Latent Semantic Analysis (LSA) measures of similarity, between the words in the idiom and its paraphrase (e.g. *spill the beans* 'reveal a secret') or its definition[3] (e.g. *spill the beans* 'reveal secret information, especially unintentionally or indiscreetly'), respectively. These LSA scores were obtained from a pairwise comparison of two texts (i.e. an expression and its paraphrase or an expression and its definition).[4] This analysis compares the local contexts of the two texts in order to obtain a value of similarity (Landauer et al., 1998). If the exact words in the expression have little, if anything, to do with the expression's overall meaning, then the LSA score will be small (e.g. *cut the mustard* – paraphrase: 'be acceptable' = 0.07; definition: 'come up to expectations; meet the required standard' = 0.29). But if the exact words used share meaning or contribute more intrinsically to the idiom's overall meaning, then the LSA score will be much larger (e.g. *stop something in its tracks* – paraphrase: 'stop something' = 0.87; definition: 'make something immediately stop continuing or developing' = 0.83). The LSA scores between the paraphrases and the definitions sometimes varied considerably (e.g. *hear something through the grapevine* had a score of 0.81 for its paraphrase 'hear gossip', but only 0.36 for its definition 'acquire information by rumour or unofficial communication'), which possibly contributes to the significance of either variable in the models. In some models only one of the LSA measures was significant, whereas in other models one measure produced a lower AIC value than the other.[5] The relevance for one or the other variable in the models may be explained by the similarity of the idiom and its meaning (i.e. definition) or the similarity of the idiom and its usage (i.e. paraphrase).

A Principal Components Analysis (PCA) was conducted on the co-occurrence frequency measures of the idiom, which included the frequency of the whole idiom,

[3]From the Oxford Dictionary of English Idioms (Ayto, 2009).

[4]The LSA scores were obtained from the English Lexicon Project (Balota et al., 2007), available at http://lsa.colorado.edu/.

[5]Akaike's Information Criterion (AIC) measures how well the predictor variables improve the predictiveness of the model (cf. Akaike, 1974, 1985). This measure is used as a form of model comparison by measuring the difference of the entropies of the models (i.e. the negative of the entropy, or negentropy, is often regarded as the amount of information), given the number of parameters, by means of a log likelihood ratio test. Therefore the model with the smallest AIC value is the preferred model.

frequencies of the individual words, and all possible combinations of adjacent words (e.g. word1 and word2; word2 and word3; word1 and word2 and word3); see Appendix E for details. This dataset therefore included 29 frequency measures (i.e. columns) for the 120 idioms (i.e. rows). In other words, the data contain 120 points in a 29-dimensional space. PCA "reduces the number of dimensions required for locating the approximate position of the data points" (Baayen, 2008, p.120), in order to remove superfluous dimensions. Thus the data points remain fixed at their location, while new dimensions, or Principal Components (PCs), are determined. These new dimensions are uncorrelated, are ranked by the proportion of variance they explain, and more succintly account for the data. In the models below, only the first PC (henceforth `PC1.logFrequency`) is significant. The highest loadings on this PC were the co-occurrence frequencies associated with the longest idioms (e.g. the frequency of words 5 and 6 occurring together – *the grindstone* in the idiom *keep your nose to the grindstone* – or the frequency of words 4, 5, and 6 – *to the grindstone*). Meanwhile, the smallest loading was for the frequency of the whole idiom (i.e. *keep your nose to the grindstone*). This latent variable proved to be a more effective predictor than the log frequency of the idiom as a whole, evidenced by a lower AIC value produced during model comparison (i.e. one model containing the log frequency of the whole idiom and one model containing PC1.logFrequency).

Several participant-related variables are also significant in the models, such as `KnowIdiom` and `KnowBlendingIdiom`, factors indicating each participant's knowledge of the experimental and blending idioms respectively (i.e. 'yes' or 'no'). `Gender` is a factor specifying whether the participant is male or female. Lastly, a second PCA was conducted on the rating responses for the seven Language Questions (LQs) above. Participants were asked to rate these prescriptively 'incorrect' sentences for their acceptability and their answers for all seven were included in the PCA. This analysis then had 60 data points (i.e. participants) in a 7-dimensional space (i.e. LQs). Once again, the number of dimensions was reduced to more succinctly account for the data using uncorrelated variables. In the models below, only PC2 (henceforth `PC2.LQ`) was significant, producing a lower AIC value than any of the LQs independently (i.e. comparing eight models, one containing `PC2.LQ` and seven containing one LQ from LQ1 through LQ7). The highest loadings on `PC2.LQ` are LQ6 (i.e. *could care less*) and LQ7 (i.e. *irregardless*) – two variants which reflect idiomatic variation and blending – while the smallest loadings are LQ4 (i.e. *roof* instead of *ceiling*)

and LQ1 (i.e. *amount* instead of *number*), which utilize a different word for a similar concept. This variable then, succinctly reflects the participant's flexibility with language usage.

### 4.3.2 Idiom as AOI

This section discusses the results for the idiom as an Area Of Interest (AOI). Fixation durations on each word in the idiom were summed to obtain the three response measures (i.e. Total Fixation Duration, First Fixation Duration, and Fixation Count) for the whole idiom. These three models are discussed together, as the same predictors show similar patterns in each model, and shown in detail in Appendix C.

Two significant interactions are observed with `Condition`, see Figure 4.1, for the Total Fixation Duration (TFD) model. The first interaction, in the left panel of Figure 4.1, is between `Condition` and `KnowIdiom`. The canonical form, and the majority of variants, show the same general pattern: shorter fixation durations are observed on known idioms. These variants (except integrated concepts) are therefore shown in grey in Figure 4.1, as they do not significantly differ from the canonical form. Known idioms in the First Fixation Duration (FFD) and Fixation Count (FC) models also show shorter and fewer fixations respectively than unknown idioms. Partial forms in the TFD model however show a different pattern. Fixation durations on this variant are relatively similar regardless of whether the participant is familiar with the expression or not; thus a facilitation effect for knowing the idiom is not observed with partial forms, as it is with the other variants. This particular variant is fixated upon less than the canonical form, likely due to the partial form being shorter in length (i.e. fewer number of words). This is in line with the longer fixations observed on integrated concepts, as a significant main effect. An additional word is integrated into the idiom for this variant, therefore making it longer in length and requiring additional fixations (i.e. larger summed fixation duration measure).

The second interaction in Figure 4.1, shown in the right panel, is between `Condition` and `Length`. The general pattern observed here is that longer idioms show longer summed fixation durations, as expected, due to the increased number of words in the idiom. Lexical variation, formal idiom blends, and literal meaning of the idiom are not significantly different from the canonical form (shown in grey). The other two variants show a pattern that is significantly different from the canonical

form. Idioms with integrated concepts show a slight inhibitory effect of length, where an additional concept is a little more difficult to integrate into shorter idioms (i.e. extra time is needed). Partial forms of the idiom show the reverse pattern. A partial form of a shorter idiom has even fewer words to fixate upon and therefore shows considerably shorter fixation times. In sum, durations on integrated concepts and partial forms are more comparable to the canonical form when the idiom is longer but these durations are more deviant when the idiom is shorter. This interaction between `Length` and `Condition` is the only interaction to be evident in all three response measures (TFD, FFD, and FC) for the idiom as AOI, with partial forms appearing significantly different from the canonical form in all three models.[6]

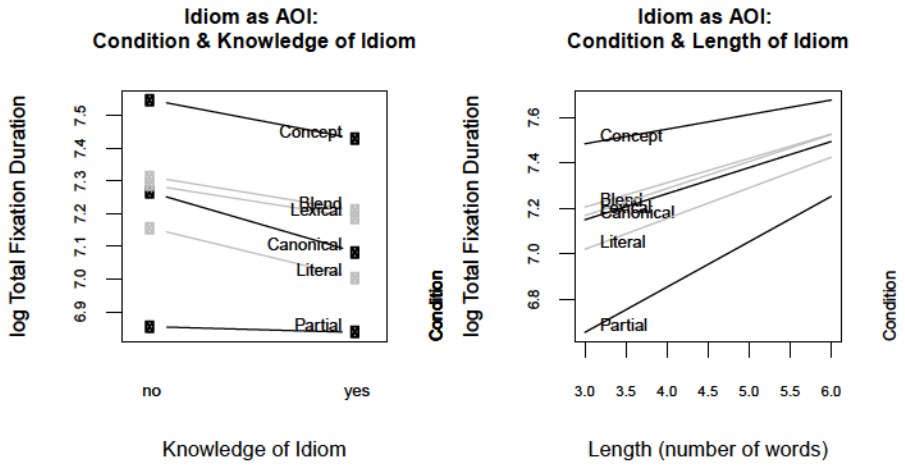


Figure 4.1: Interactions with Condition in the Mixed-Effects Linear Regression Model for the Summed Total Fixation Duration on the Whole Idiom as an Area of Interest. Lines in grey represent factors levels which are not significantly different from the canonical form.

Interestingly, the literal meaning of the idiom in these two plots shows shorter fixation durations than the canonical form. These fixations are not significantly shorter (t = -1.94), but certainly trending towards significance. The literal meaning

[6] `PC1.logFrequency` was also significant in the models on the idiom as AOI. However, this variable is strongly correlated with `Length` ($r$ = -0.9). This correlation is unsurprising given that `PC1.logFrequency` was created based on adjacent co-occurrence frequencies. Model comparison shows that `Length` is the more significant predictor in these models, producing a considerably lower AIC value, and therefore was retained in the models for idiom as an AOI at the expense of `PC1.logFrequency`.

is also associated with an increased number of comprehension questions answered incorrectly (X = 63.75; $p = 0$). A recursive partitioning analysis of the two variables `Condition` and `KnowIdiom`, using the `party` package (Hothorn et al., 2006) in R (R Core Team, 2014), reveals that this result is not due to the participant's knowledge, or lack of knowledge, of the idiom. Recursive partitioning utilizes conditional inference trees, which test all variables fed into the model for their predictability of the data, and then implements a binary split based on the most significant variable. The first partition was the literal level of `Condition` ($p < 0.001$). In other words, participants were more likely to get the comprehension questions associated with the literal meaning of the idiom wrong, regardless of whether they knew the expression or not. Two other possibilities might explain this outcome. First, participants may have needed additional time processing the idiom, a duration more comparable with the canonical form, in order to fully integrate the literal meaning of the idiom into the preceding context. Hence, they could have gotten the comprehension questions wrong because they read the idiom too quickly. Second, the 'literality' (i.e. the degree to which an idiomatic phrase has a plausible literal interpretation) of the idioms used in these questions was not controlled for (cf. Titone and Connine, 1994a,b). Some idioms, such as *have a card/ace up your sleeve*, have been shown to elicit higher ratings for literality than idioms like *foot the bill* (Titone and Connine, 1994b). Therefore, it is likely that the questions associated with idioms which have a lower literality rating brought the overall correctness down for this condition.

Two additional significant interactions are observed in the FFD model, but are realized as main effects in the TFD and FC models. First, `PortionIdiomAltered` occurs in an interaction with `Condition`. Fixations on the majority of idiom variants in the FFD show similar durations regardless of whether the beginning or ending of the idiom was altered. Partial forms however have considerably shorter fixations on the idiom if the beginning was altered (e.g. *use the grapevine*) and longer fixations if the ending was altered (e.g. *spilled it*). This interaction does not persist in the TFD model nor in the FC model, which show significantly longer fixations and more fixations on the idiom if the beginning was altered. These results indicate then that additional fixations on partial forms were required to process this variant, especially if the beginning was manipulated, resulting in a similar total number of fixations and fixation durations to the other variants.

The second significant interaction observed in the FFD model is between

PC2.LQ and KnowIdiom. Those who know the idiom have shorter fixations than those who do not know the idiom if they are more flexible with language (i.e. those who had higher acceptability ratings on the seven Language Questions). Once again, this interaction does not persist in the TFD and FC models. Instead only main effects are observed for PC2.LQ, where all participants fixate less on the idioms, both the canonical form and the idiom variants, if they are more flexible with language.

A significant main effect of Gender is observed for both fixation duration models (i.e. TFD and FFD) – males tend to fixate less long on the idiom than females. This predictor is not significant in the FC model, suggesting that they may fixate less long on the idiom, but they still show a similar number of fixations. A significant main effect is also observed for meanVariationRating, in all three models. Variants which received higher acceptability ratings are fixated on less long; preferred variants are easier to understand and interpret (or perhaps variants easier to interpret are preferred). Additionally, longer fixation durations appear on idioms which have higher LSA scores for the idiom's paraphrase (i.e. LSA.Score.Paraphrase), in all three models. This finding seems initially surprising, as previous analyses on the comprehension of idioms and specifically idiom variants suggest that idioms are easier to understand when the individual components contribute meaning to the whole (cf. decomposable idioms in Gibbs et al., 1989a). However, the LSA scores are a rating for how similar the local contexts are between the idiom itself and its paraphrase (i.e. how interchangable is the expression with its paraphrase). If the LSA score is high (i.e. the paraphrase is easily interchangable) then looking time increases as the contexts are not distinctive, or discriminative, for the idiom. But if the LSA score is low, then the idiom and its paraphrase are less interchangable, making the context more distinctive and the idiom more predictable.

Finally, a main effect of TrialScaled is also significant in both the TFD and FC models; participants fixate less long on the idiom the further into the experiment they get. But the degree to which each participant is affected by the order of presentation varies, as evidenced by the significant by-Subject random slopes for TrialScaled in all three models. The TFD model also has significant by-Item random slopes for Condition with correlation parameters. These slopes indicate that participants' fixation durations varied depending on which idiom occurred in which condition – participants found certain idioms easier or more difficult to understand depending on the condition in which they occurred.

### 4.3.3 Altered Word as AOI

This next section explores the altered or changed word within the idiom as an Area Of Interest (AOI). For variants in which the beginning portion of the idiom was altered, it may appear to the participant reading the text as though the ending was manipulated (e.g. as if the 'blending idiom' was the intended idiom, in *call the strings*, or part of an idiom was inserted into an otherwise non-idiomatic text, such as *use the grapevine*). Therefore, I analyzed the AOI of the altered word in two ways: the Manipulated Word (i.e. the word that I purposely altered in the idiom) and the Surprising Word (i.e. the word the participant may believe has been changed or the word that might signal an idiom is actually present to the participant). When the ending of the idiom was altered, these AOIs are identical, but for idioms which had the beginning altered, these two AOIs differ. The AOI for the surprising word would be on the final word, or in the case of V-NP-PP idioms, on the first content word after the alternation (e.g. on *someone* for the variant: *twiddle someone around your finger*). In addition, this section only focuses on the four idiom variants (i.e. lexical variation, partial forms, idiom blends, and integrated concepts) and how they compare to the canonical form. Since there is no altered word in the literal meaning condition, this variant has been excluded from the analyses in this section.

Both analyses (i.e. Manipulated Word as AOI and Surprising Word as AOI) show a significant interaction between `Condition` and `PortionIdiomAltered`, seen in the top panels of Figure 4.2. The overall pattern evident in this interaction on the manipulated word is that longer fixation durations occur when this word appears at the end of the idiom (i.e. the ending has been altered). This pattern is also true for the canonical form; significantly longer fixations are on the ending of the idiom ($t = 4.49$). Since the idiom occurs at the end of a sentence, these longer fixations on the canonical form (and variants) may reflect a sentence wrap-up effect (cf. Rayner et al., 2000; Hirotani et al., 2006). Nevertheless, the manipulated word for most variants do show significantly longer fixations than the canonical form. This is not true of lexical variation, which is the only variant that does not take significantly longer to process than the canonical form ($t = 1.54$). The Manipulated Word as AOI shows one variant, partial forms, which appears considerably different from the canonical form. Longer fixations are observed on the manipulated word in a partial form when the beginning has been altered, such as *use the grapevine*. But when the ending is altered (e.g. *spilled it*), fixations on the manipulated word are not significantly different

from the canonical form (t = -1.44). Since altering the verb does not always result in significantly longer fixations (cf. the non-significantly different lexical variant when the beginning is altered), this finding suggests that altering the verb to a semantically vague verb, in order to make the sentence grammatical, significantly inhibits processing. This interaction is observed in all models (TFD, FFD, and FC).

Fixation durations on the surprising word of the idiom variants (i.e. the final word when the beginning was altered) show a spillover effect. The top-right panel of Figure 4.2 shows the significant main effects observed for all variants which had the beginning of the idiom altered (e.g. *use the grapevine*). These words (i.e. *grapevine*) are the same words from the canonical form; however, the beginning portion of the idiom has been altered. They therefore take significantly longer to process, when following a manipulated word (e.g. *use*). The surprising word when the ending has been altered is the manipulated word (e.g. *it* in *spilled it*); these are the same in both plots. Interestingly, integrated concepts show longer fixations, albeit not significantly longer, if the ending of the idiom is manipulated.[7] Half of the longer idioms (i.e. V-NP-PP) had an integrated concept inserted earlier in the expression (e.g. *kept his overwhelmed nose to the grindstone*). It may be these instances which decrease the overall fixation durations for idioms which had the beginning altered, potentially suggesting that concepts integrated earlier into the expression are less surprising and therefore result in shorter fixations, which could be due to the greater predictability of the idiom (cf. Titone and Connine, 1994b; Titone and Libben, 2014). The largest spillover effect is with partial forms. These variants showed significantly longer fixations on the manipulated word when the beginning was altered (i.e. on the beginning of the idiom), as well as significantly longer fixations on the word following the manipulated word in partial forms. It appears that the semantically vague words used in these sentences make these partial forms more difficult to process and cause considerable spillover effects. It remains to be determined whether partial forms from more naturalistic language produce this same effect.

The middle panels of Figure 4.2 illustrate the interaction between knowledge of the idiom (i.e. `KnowIdiom`) and the participant's flexibility with language (i.e. `PC2.LQ`). Interestingly, when the manipulated word is the AOI, flexibility with language appears to only be facilitative for those who do not know the idiom, illustrated by the non-significant slope for those who know the expression (t = -1.29). But when

[7] The manipulated word and surprising word for integrated concepts are identical.

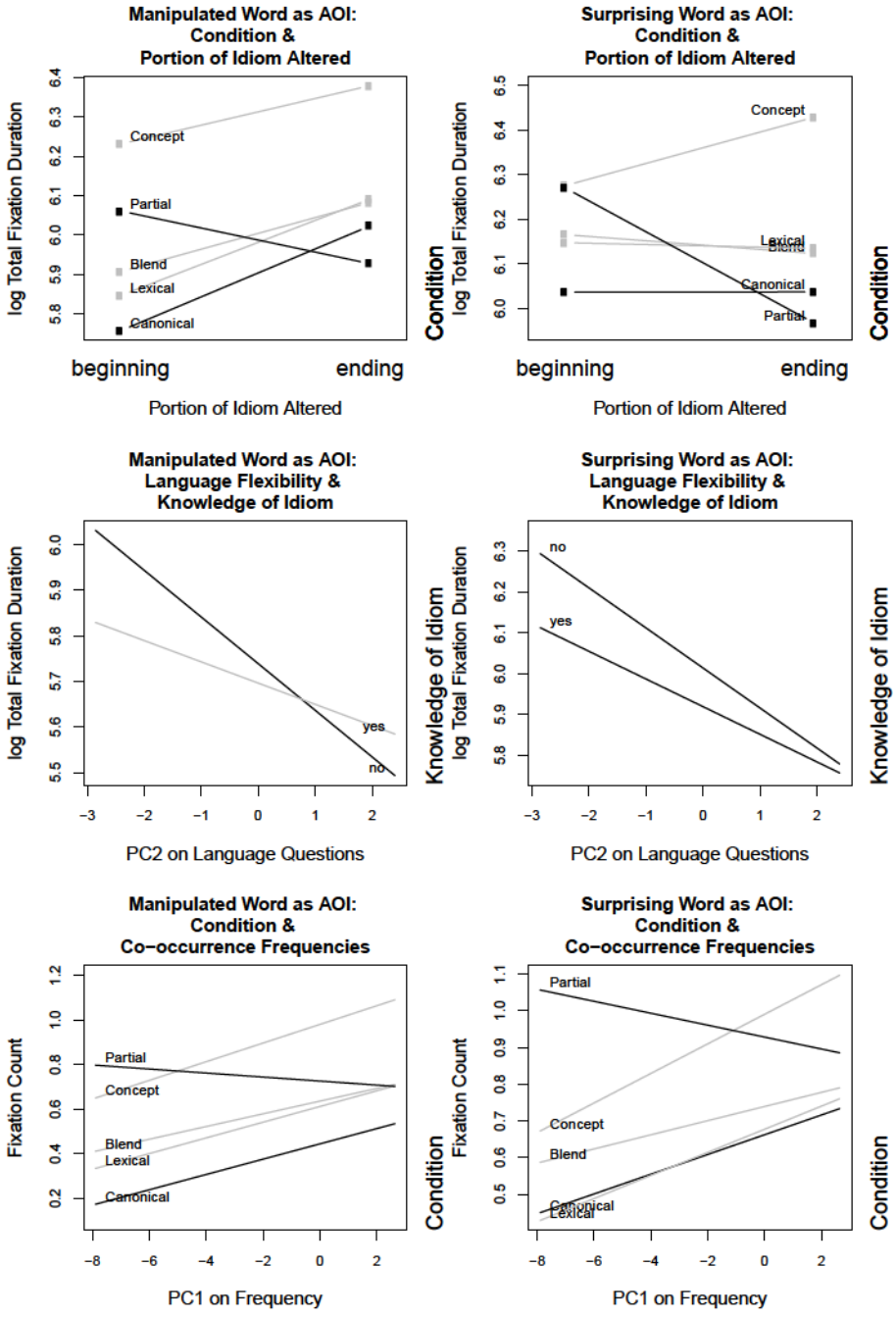


Figure 4.2: Interactions in the Mixed-Effects Regression Models for the Total Fixation Duration and Fixation Count on the Manipulated and Surprising Words of the Idiom as Areas of Interest. Lines in grey represent factors levels which are not significantly different from the canonical form in an interaction or slopes which are not significant in a single interaction.

the surprising word is the AOI, language flexibility becomes facilitative for both those who do not know the expression and those who do, resulting in an interaction that is no longer significant (t = 1.23). Thus, both knowledge of the idiom and a flexibility with language result in easier processing and interpretation of idiomatic variants.

The third interaction shown in the bottom panels of Figure 4.2 is between `Condition` and `PC1.logFrequency`. This interaction is only present in the FC models, where the co-occurrence frequencies of an expression appear to influence the number of times one fixates on a particular variant condition. The majority of conditions show the same general pattern as the canonical form, a greater number of fixations are observed on the altered word (both manipulated and surprising) for those idioms which contain higher co-occurrence frequencies. Partial forms however, show a different pattern. Fixations on the altered word in a partial form seem to show a greater number of fixations when the co-occurrence frequencies are smaller. This might suggest that idioms with smaller co-occurrence frequencies have more distinctive words in the idiom, but when one of these words is omitted it causes a greater number of fixations because it is more surprising to the reader. This interaction is not significant in the fixation duration models, suggesting that participants look a greater number of times at partial forms associated with smaller co-occurrence frequencies, but that the actual fixation time is comparable with all variants.

Looking specifically at formal idiom blends, two additional interactions are observed on the Manipulated Word, shown in Figure 4.3. In fact, both interactions are with `KnowBlendingIdiom`, the participant's knowledge of the second, or 'blending', idiom. The first interaction is with `PortionIdiomAltered`, seen in the left panel of Figure 4.3. When the ending of the idiom is altered, participant's fixation durations on the manipulated word (i.e. the portion of the blending idiom) do not significantly differ, regardless of whether they know the blending idiom or not. However, if the beginning portion of the idiom was altered, participants who do not know the blending idiom fixate significantly less than those who know this idiom (t = -2.62). Thus, participants who know the second idiom used in the blends, fixate for a comparable duration regardless of whether the blending idiom occurs at the beginning or ending of the blend. Participants who do not know the blending idiom notice the idiom variant when the blending idiom is incorporated into the final portion of the expression, showing fixations comparable to participants who know the blending idiom. However, they are not aware when the beginning has been manipulated and

therefore show shorter fixations on this manipulated word.

Given the significantly longer fixations on the surprising word (in the top-right panel of Figure 4.2), this suggests that both those who know the blending idiom and those who do not show significantly longer fixations on the surprising word in an idiom blend when the beginning portion of the experimental idiom has been altered. Simply put, participants are 'surprised' when they reach the final part of an idiom blend and need additional time to interpret the blend. This interaction between `KnowBlendingIdiom` and `PortionIdiomAltered` on the Manipulated Word is not present in the FFD (t = -1.09), suggesting that those participants who know the blending idiom, refixate on this manipulated word, resulting in a processing cost for those who know the blending idiom. This effect does not emerge when the experimental idiom comes first in the blend possibly due to the fact that the context better facilitates the processing of this portion of the idiom.

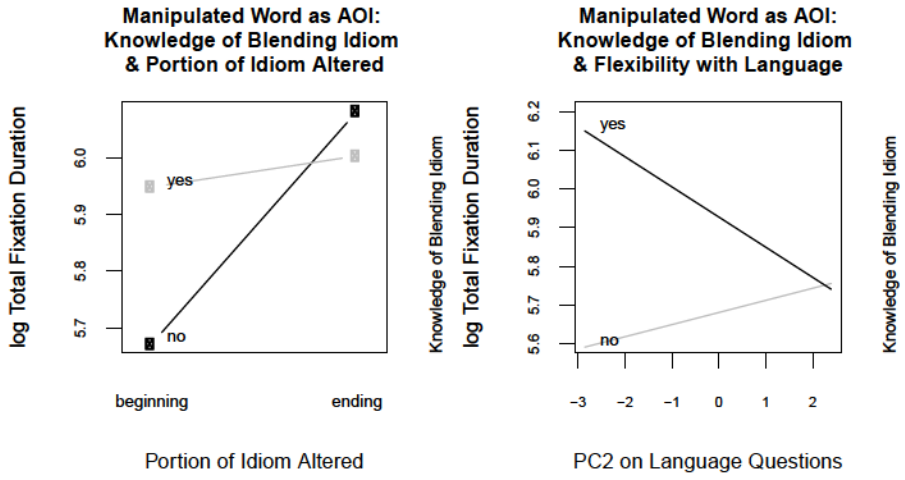


Figure 4.3: Interactions in the Mixed-Effects Linear Regression Model for the Total Fixation Duration on the Manipulated Word of the Idiom as an Area of Interest for Formal Idiom Blends. Lines in grey represent non-significant slopes.

The second significant interaction observed on formal idiom blends is between `PC2.LQ` and `KnowBlendingIdiom`, shown in the right panel of Figure 4.3. Participants who do not know the blending idiom are not significantly affected by their flexibility with language (t = 0.62); they fixate for a comparable duration on the manipulated

word no matter how they rated the acceptability of the seven LQs. Those participants however, who do know the blending idiom, show significantly longer fixation durations on the manipulated word if they are inflexible with language (t = 2.18). But if participants who know the blending idioms are flexible with variants and non-standard forms, then they show fixation times comparable to those who do not know the expression. While idiom blends can produce inhibitory effects, especially if one knows the second idiom involved in the blend, these effects can be minimized the more the participant is accepting of non-standard or 'incorrect' language usage.

A few additional main effects are also observed on the altered word. Participants have shorter fixation durations and fewer fixations on the altered word (both manipulated and surprising) when the variant is rated as more acceptable (i.e. `meanVariationRating`). This predictor is significant in both TFD and FC models, and is even significant in predicting fixation duration on the manipulated word specifically within formal idiom blends. In the FFD model for the Manipulated Word, `LSA.Score.Paraphrase` is significant, showing shorter fixations on the manipulated word when it occurs in an idiom which has a higher LSA score. In other words, modifications to an idiom are fixated on less if the idiom occurs in similar contexts as its literal paraphrase. However, this main effect of `LSA.Score.Paraphrase` does not persist in the TFD model, implying that this is a momentary processing advantage. Interestingly, the model specific to formal idiom blends for the Manipulated Word prefers a main effect with `LSA.Score.Definition`. Shorter fixations are observed on manipulated words within idiom blends if the experimental idioms occur in more similar contexts with their definitions. This perhaps suggests that idiom blends are easier to process if their form is more similar to their meaning. These effects with LSA scores are not evident on the Surprising Word, suggesting that similarity to the definition or paraphrase is not facilitative if one is surprised by a lexical item.

Finally, `TrialScaled` is a significant predictor in the TFD models for both the Manipulated and Surprising Word. The further the participant gets into the experiment, the shorter their fixation durations on the altered words. But participants vary as to how much faster they become, as evident from the by-Subject random slopes for `TrialScaled` in the TFD and FC models. The random effects structure in the TFD models on both the manipulated and surprising words show significant by-Item random slopes for `Condition` with correlation parameters. This indicates that participants vary in their fixation durations on altered words depending on how

the idiom was modified – specific idioms are preferred with certain types of variations and the fixation durations on these alternations reflect those preferences.

## 4.4 Discussion

### 4.4.1 Canonical Forms and Idiom Variants

The first research question asks whether the four variants included in this study are processed differently from the canonical form. The variant showing the greatest difference from the canonical form is the partial form of the idiom. This idiom variant is fixated on less than the canonical form, largely due to the omission of a word (or words) from the expression. The greatest difference in fixation duration is observed on partial forms of shorter idioms, showing significantly shorter fixations than canonical forms. These shorter fixations are not facilitated by the speaker's knowledge of the expression, unlike the canonical form and the other variants, but instead partial forms show no processing advantage when speakers know the idiom.

This lack of a facilitation effect is most likely due to the fact that speakers, both those who know the idiom and those who do not, need to fixate longer on the remaining portion of the expression and the altered word, especially if the beginning of the idiom has been modified. In order to make partial forms grammatical in the sentence when the beginning of the idiom was altered, semantically vague verbs (e.g. *be, do, have, use*) were used to connect the idiom to the sentence. However, this strategy had serious consequences for the results. Participants fixated significantly longer on these 'replacement' verbs and significant spillover effects were observed on the surprising word of the expression. A similar inhibitory effect was not observed if the ending of the expression was modified (e.g. *spilled it*). In fact, fixations were similar to the canonical form when the ending was altered. These results then (i.e. when the beginning of the idiom was altered) are likely due to the design of the experiment. Using such controlled stimuli made these partial forms unnatural and difficult to interpret. A study investigating partial forms in naturally occurring language may shed more light on the degree of difficulty for processing this variant.

Idioms with integrated concepts are processed similarly to the canonical form, but show significantly longer fixation durations, predominantly occurring as a main effect. This longer fixation time on the idiom is attributable to the extra word in the

expression. This variant only seems to occur in one interaction, with the length of the idiom. Integrated concepts take longer to process if the idiom is shorter; this may be due to fewer words and therefore less information available in the expression to facilitate the integration of the additional concept. Comparing the full variant (i.e. idiom as AOI) to the canonical form allows enough of a comparison to determine that the additional concept causes a longer processing time. However, it is more difficult to interpret the analysis when it is conducted on the manipulated word within the idiom since the manipulated word is an inserted element. In other words, is the comparison really measuring the same thing? For example, if the idiom under examination is *spill the beans* and the variant is *spill the relationship beans*, then the manipulated word is *relationship* whereas the comparison in the canonical form is *beans*. While this comparison then might insert a potential confound, it is impressive to observe that the additional word in the altered word models still only shows a main effect from the canonical form – the additional word takes longer to process than the original expected word in the canonical form. A comparison on the final words of these two expressions (i.e. *beans* in both the integrated concept variant and the canonical form), shows no significant difference, as a main effect or as an interaction. The final words are processed similarly, with no spillover effects, despite the additional concept inserted into the variant. In other words, this variant appears to be processed similarly to the canonical form, simply taking longer to process the additional, unexpected, word in the expression and integrating it into the idiomatic context.

Lexical variation is another variant which does not differ greatly from the canonical form. This variant does show longer fixations than the canonical form, but not significantly longer. Also, this variant does not occur in significant interactions with the canonical form, generally showing the same pattern as the canonical form. Significantly longer fixations are observed on the surprising word when the beginning of the idiom is altered, showing a spillover effect on the final word; however, this additional time does not result in significantly longer comprehension times for the idiom as a whole. In addition, this variant shows nearly identical number of fixations as the canonical form on the surprising word. These findings are quite different than those of Gibbs et al. (1989a), who found a significant difference in similarity ratings between the lexically altered variant (either manipulating the verb or the noun) and the literal paraphrase. The variants included in this study found

that the variant differed slightly from the canonical form, but predominantly was not significantly different from the canonical form, suggesting that Gibbs et al.'s (1989a) findings were due to preferences and semantically similarity, and not comprehension. Given that idioms can occur in "idiom sets" or "clusters" (Moon, 1998), such as *shake/quake/quiver in one's boots* or *down the drain/chute/tube/toilet*, it seems fitting that this variant did not cause too much disruption during processing.

Interestingly, formal idiom blends, as a whole, were also not processed significantly different from the canonical form. This variant was fixated on longer than the canonical form, but not significantly longer when the Area Of Interest (AOI) was the idiom. This suggests that all deviations from the canonical form observed on the altered word as AOI do not result in a significant processing cost for this variant. Fixations on the manipulated word are longer than the canonical form, showing a significant main effect. Shorter fixations were observed on the manipulated word when it occurred at the beginning of the idiom, but only for those participants who did not know the expression (i.e. the blending idiom). They did not fixate longer on the manipulated word presumably because they did not realize that it was part of a different expression. However, knowledge of a particular idiom does not enter into a significant interaction on the surprising word, suggesting that all participants fixated longer on the surprising word (i.e. the final word in the idiom when the beginning was altered), whether they knew the idiom or not. Flexibility with language (i.e. acceptability of prescriptively 'incorrect' usage) does make the switch between idioms easier, especially for all speakers processing the surprising word.

An analysis was conducted specifically on idiom blends to determine if any additional predictors contribute to their processing. In a production task, Cutting and Bock (1997) tried to induce idiom blends by presenting two idioms at the same time, which either were similar or different in their syntax or semantics, and found that participants were more likely to produce blends when the two idioms shared similar syntax or semantics. Unlike Cutting and Bock's study, the syntax of the canonical forms, or their semantics, were not predictive of fixation duration for the comprehension of idiom blends. This indicates that while these variables play some role in slips of the tongue, they do not appear to affect the comprehension of blended variants or the altered word within these variants.

The final 'variant' included in this study was the literal meaning of the idiom. The form of the idiom was not altered, but rather it occurred in a context which

attempted to convey a literal interpretation of the idiom. This variant was only included in the analysis on the idiom as an AOI, comparing the processing differences between the sequence of words processed literally versus idiomatically. The results show that the literal meaning was fixated on less than the canonical form of the idiom, and trending towards having significantly shorter fixation durations. However, participants were also significantly more likely to get the comprehension questions wrong when they occurred with the literal meaning. This could be the result of three possibilities. First, the participants may have been reading these sentences too quickly, and did not fully understand the literal meaning of the idiom. Second, the contexts may not have fully conveyed the literal meaning of the idiom. Third, some idioms like *foot the bill* have low literality ratings (i.e. they are less likely to have a plausible literal interpretation, see Titone and Connine, 1994b). These idioms may have brought the overall correctness of the comprehension questions down, as the literality of the idioms was not controlled for in this study, and unfortunately not all idioms overlapped between Titone and Connine's study to confirm this speculation.

### 4.4.2 Literal Meaning and Idiom Variants

The second research question sought to explore if the idiomatic variants were processed differently from the literal meaning of the idiom. McGlone et al. (1994) found that variants in their study were processed as quickly as a literal paraphrase of the idiom. The present study however, used the idiom in its canonical form (i.e. not a paraphrase), but placed in a different context (e.g. *While the guys were reshingling, they suddenly went through the roof*). The results show that the literal meaning is not processed significantly different from the canonical form.

In order to determine how the variants compared with the literal meaning of the idiom, the variable `Condition` was relevelled so that the literal meaning was the baseline or reference level (i.e. all factor levels would then be compared to the literal meaning and not the canonical form). The default setting of mixed-effects linear regression models (using the `lme4` package) is a 'treatment contrast', which uses one factor level as the baseline to which all other (unordered) factor levels are compared. In the models discussed in this study, that baseline was the canonical form. But to determine how the literal meaning compared to the variants, the same models were rerun with the literal meaning as the reference level. Three of the variants (i.e.

lexical variation, integrated concepts, and idiom blends) showed significantly longer fixation durations and a greater number of fixations than the literal meaning of the idiom, while partial forms showed significantly fewer fixations and shorter fixation durations than the literal meaning. In other words, the literal meaning of the idiom was processed differently from all idiomatic variants.

This is a different result from McGlone et al. (1994) who found that idiom variants were read as quickly as the literal paraphrase. Two possibilities may explain this different finding. First, this study compared the variants to the literal meaning, whereas McGlone et al. (1994) compared them to the literal paraphrase. The difference in semantics likely accounts for the different reading times between the literal meaning and the variants in this study. Second, the literal meaning and the variant sentences presented in this study differed along two dimensions: context and form. These two dimensions may also contribute to why the variants are processed significantly different from the literal meaning, and also why the literal meaning and two of the variants (i.e. lexical variation and idiom blends) are not processed significantly different from the canonical form (i.e. they only vary along one dimension).

### 4.4.3 Differences between the Variants

The third research question inquired as to whether the different idiom variants are processed differently from each other. As previously discussed, variants that are the same length as the canonical form (i.e. lexical variation and idiom blends) show a comparable number of fixations and fixation durations as the canonical form. But, in order to determine how the variants compared to each other (as opposed to the canonical form), new models were run without any interactions or covariates and a Tukey test was conducted (using the `multcomp` package) for a multiple comparison of means. The results are consistent with the larger, more complex models discussed in detail above. When the idiom is the AOI, lexical variation (e.g. *rock the canoe*) and formal idiom blends (e.g. *rock the applecart*) are not significantly different from each other, showing overlapping confidence intervals. These are the same variants which are also not significantly different from the canonical form. It appears then that idioms and idiom variants which are the same length show comparable reading times and are understood similarly.

Meanwhile, when the manipulated word or the surprising word are the AOI, the only variant which is significantly different from the others is the integrated concept (e.g. *rock the initiative boat*). This variant inserts an additional word into the idiom, adding extra information to the idiomatic meaning, and therefore takes longer to integrate and interpret. The other variants however, replace an existing word of the idiom to produce the respective variant: with a synonymous word in the lexical variation condition (e.g. *rock the canoe*), with a word from another idiom in the idiom blend condition (e.g. *rock the applecart*), or with a semantically vague word in the partial form condition (e.g. *rock things*). These 'replacement' words show overlapping confidence intervals, or similar processing times. Thus, these more specific results fit in nicely with the larger picture. Variants which are the same length as the canonical form show similar reading times, but if additional information is inserted in the idiom then comprehension takes longer in order to process this extra information.

### 4.4.4 Additional Findings and General Discussion

The results of this study are striking – variants of similar lengths to the canonical form are not processed significantly different from the canonical form. However, a number of limitations are also evident. Concerns have already been raised in regards to two variants included in this study: integrated concepts and partial forms. Specifically, measurements on the altered word for the integrated concept may not be an accurate comparison, so the results for this particular area of interest should be interpreted with caution. Meanwhile, potential confounds may have been introduced into the analyses for partial forms. Semantically vague, and possibly contextually inappropriate, verbs were used to conjoin the idiom to the sentence for alternations made to the beginning of the idiom, whereas pronouns, which tend to be skipped or simply not fixated on for as long, were used when the ending of the idiom was altered. These two features likely contribute to the results of this variant. This study controlled for different types of idiomatic variation to compare these variants to each other as well as to the canonical form, but in doing so, may not have allowed the use of idiomatic variants to be realized in a naturalistic way. Perhaps a study needs to be conducted where corpus-based instances of actual language usage are utilized as the stimuli (similar to McGlone et al., 1994, but selected based on a similar structure or form). This might then allow for more controlled comparison between the variants,

while still allowing naturalistic data to be utilized in the study. Moreover, by placing the idioms in sentence-final position in a separate clause presented on the computer screen has the limitation that regressions back to the idiom or to the preceding context are not possible, nor are spillover effects from the variant onto later portions of the sentence. Some spillover effects have been observed on the surprising word, but it is not possible to know from this study whether spillover effects are evident if the ending of the idiom is varied since the idiom is the last element in the sentence. A follow-up study may be needed to determine if and for how long spillover effects last onto subsequent text.

Despite these limitations, this study reveals several interesting findings. For example, fixation durations on the idiom as a whole show that idioms are fixated on longer if the beginning of the idiom is manipulated. This is due to longer fixations on the manipulated word and on the surprising word (i.e. the final word when the beginning of the idiom is manipulated). Thus, longer fixations on the surprising word, which have spilled over from the manipulated word, contribute to the overall longer fixations on the idiom as a whole if the beginning is altered (i.e. the verb). These results are quite different from what Gibbs et al. (1989a) found in their study on idiomatic variation, where manipulating either the beginning (the verb) or the ending (the noun) resulted in significantly lower similarity ratings with the literal paraphrase, but did not significantly differ from each other. Based on these results, they claim that altering one element in the idiom results in slower comprehension of the idiomatic variant, but there is no difference in comprehension for whether the beginning (verb) or ending (noun) is modified. Once again, these ratings may simply reflect the participants' preferences for the variants and not necessarily their ability to interpret the variants. Participants in this study are successful at interpreting the variants, shown by the correctness of the comprehension questions with the idiomatic meaning, but comprehend variants easier when alternations occur later in the expression.

Interestingly, `meanTransparencyRating` was not a significant predictor in any of the models. These mean ratings of transparency were collected in a separate study, where participants were asked to rate the transparency, or how clear the meaning of the expression is (see Chapter 2 for details). These ratings were collected instead of the semantic decomposability categories, as that distinction has proved unreliable (Titone and Connine, 1994b; Tabossi et al., 2008). Titone and Connine (1999) found

that nondecomposable idioms took additional time to process if they occurred after context. I therefore expected to see a main effect of transparency, with idioms lower in transparency showing longer fixations, since all idioms presented in this study occurred after the context. However, this variable was no longer significant once `meanVariationRating` was entered into the models. In a separate experiment, participants were asked to rate the acceptability of the sentences which occurred in this study (see Chapter 3 for details). The variants which had higher ratings of acceptability had significantly shorter fixation durations. It appears that once speaker preferences for variants are taken into account, a separate measure of transparency is superfluous.

However, the LSA scores for the idiom's paraphrase were significant, at least for measures on the whole idiom. The measures obtained for the decomposability ratings (see Gibbs and Nayak, 1989; Titone and Connine, 1994b; Tabossi et al., 2008) were also collected using the idiom's paraphrase. That is, participants in those studies were given an idiom and a literal paraphrase and asked to make a judgement about "whether the individual components of each idiom made some unique contribution to the phrase's figurative paraphrase" (Gibbs and Nayak, 1989, p.108). Based on these instructions the participants grouped the idiom into the three decomposability categories: normally decomposable, abnormally decomposable, and nondecomposable. The LSA measure for the idiom's paraphrase also seems to be targeting the similarity between the idiom and its paraphrase, based on text similarity and not subjective ratings. But in this study, higher LSA scores are predictive of longer fixations, suggestive of more difficult processing. If the idiom and its paraphrase occur in similar contexts, additional time is required to process the idiom because the contexts are less distinctive and the idiom is less predictable.

The LSA score for the idiom's definition was not significant in the models on the whole idiom, only the LSA score for the idiom's paraphrase was significant. The paraphrase of an idiom may reflect more how speakers use the idiom in context. A variant then should perhaps also reflect this same usage to be more comparable to the canonical form. The definition of an idiom however, reflects the idiom's precise meaning, which may be less important when used in context or for understanding an idiomatic variant. But this measure is significant in processing idiom blends, or at least the manipulated word of a blend (i.e. the blending idiom). This may indicate that the precise meaning of an idiom plays a role in understanding formal idiom

blends, but not necessarily for interpreting all idiomatic variants.

Length and co-occurrence frequencies of the idiom also were significant predictors of fixation duration. As these two predictors are strongly correlated, they were not included together in the models. Nevertheless, a pattern is still observed. Length of the idiom was revealed to be more predictive for analyzing the idiom as a whole. Summed total fixations were longer when the idioms themselves were longer, as expected. Length was no longer preferred when investigating the effects on the altered word within the variants. Instead, the log co-occurrence frequencies, which indirectly take the length of the expression into account, were more predictive. Overall, the altered words (both manipulated and surprising) are fixated on longer if the canonical form has higher co-occurrence frequencies. This variable does not reflect the frequency of the individual word, but instead reflects the frequency with which it co-occurs with all other elements in the expression. Therefore, if the original word occurs more often alongside the other elements (even outside of the expression), then processing an alternative word in its position is inhibitory and takes speakers longer. In sum, length of the idiom predicts fixations on the whole idiom better than frequency, but the frequencies of the words within the idiom are more predictive when interpreting an alternation to the form.

Lastly, participant-related variables are also significant in idiom processing. As anticipated, knowledge of the idiom leads to faster processing (cf. Schweigert, 1986; Popiel and McRae, 1988; Schraw et al., 1988; Titone and Connine, 1994b). This variable also interacts, particularly when the altered word is the interest area, with the speaker's acceptability of the seven prescriptively 'incorrect' sentences (cf. `PC2.LQ`). The highest loadings on this variable were the sentences containing a word blend (i.e. *irregardless*) and a variant of an idiom (i.e. *could care less*). Higher acceptability of these sentences are therefore predictive of faster processing of other idiom variants. This variable is not typically included in analyses of idiom comprehension, but as illustrated in all models in this study, is a very significant predictor – speakers fixate less on idiom variants if they are more accepting of non-standard or 'incorrect' language. Finally, the analyses for this study also explored the effects of how often people use idioms and whether they like using idioms. The questions asked at the end of the experiment were included in the analyses to determine if they were significant at predicting the processing of idiomatic variation. These variables however were not significant. It is interesting to note that despite how often one

uses idioms, or their enjoyment of these expressions, they do not affect one's ability to understand idioms or their variants.

This study has shown that idiomatic variants are not necessarily processed differently from the canonical form, as previous research has suggested. Variants which are similar in length to the canonical form (i.e. lexical variation and formal idiom blends) show comparable fixation durations to the canonical form. Additional elements incorporated into the idiom or elements removed from the idiom result in longer or shorter reading times, respectively. Thus, idioms are not simply expressions that are fixed in form, but can be modified in a variety of ways, without considerable processing delays, while still retaining their idiomatic meaning.

# CHAPTER 5

# Eliciting Idiomatic Variation

## 5.1 Introduction

Research on idioms has a long history of focusing on the canonical form. It was assumed that the idiomatic meaning was only retained when the full canonical form of the idiom was used. However, recent research has begun to explore variation of idioms within corpora (cf. Moon, 1998; Barlow, 2000; Duffley, 2013; Schröder, 2013). Moon (1998) conducted an extensive investigation of idioms and fixed expressions in corpora and found that idioms can in fact occur with a wide range of variation. For instance, lexical variation can occur with nouns (e.g. *the calm/lull before the storm*), verbs (e.g. *stick/stand out like a sore thumb*), and even conjunctions (e.g. *when/while the cat's away, the mice will play*). Truncations are also observed (e.g. *birds of a feather* [*flock together*]), as well as register variation (e.g. *knock someone dead* vs. *knock 'em dead*), dialectal variation (e.g. *wear the trousers* [BrE] vs. *wear the pants* [AmE]), and even 'erroneous' forms (e.g. *whet someone's appetite* vs. *wet someone's appetite*). Schröder (2013) further explored syntactic variation of nine idioms in two corpora. Using Horn's (2003) classification, she found that fixed idioms (e.g. *kick the bucket*) and mobile idioms (e.g. *spill the beans*) had significantly fewer tokens of syntactic variation in both corpora, compared with metaphor idioms (e.g. *make headway*). The majority of variation occurring with metaphor idioms is adjectival and adverbial modification, such as *make rapid headway.*

Barlow (2000) explored conceptual blending with two idioms, *it isn't over until the fat lady sings* and *make hay while the sun shines*, on the internet. He observed several ways in which the current situation can be conceptually integrated into the

idiom. For example, full noun phrases can be integrated into the subject position (e.g. *No competition is over until the fat lady sings*), a partial form alluding to the whole idiom can be used to signal the end of a process (e.g. *The fat lady burst into song far too early last Sunday*), and even instances where the *fat lady* is aligned with a particular individual in the current situation (e.g. *Nothing is dead until the fat lady sings, as they say. LuJuana is not fat, but she will make the final decision*). Moreover, Duffley (2013) investigated two nondecomposable idioms (i.e. *kick the bucket* and *shoot the breeze*) on the internet and found that even nondecomposable idioms can undergo variation. Among his results, he found instances of passives (e.g. *breeze is shot, piss is taken, and a bit of fun is had*), integrated concepts (e.g. *reluctant to kick his brimming bucket of life*), causatives (e.g. *before we kicked his bucket for him*), and even lexical variation and literal-scene manipulation (e.g. *wheeze the breeze with a trainer*). These examples illustrate that idioms are not nearly as fixed or 'frozen' as previously assumed, but can actually show a remarkable amount of variation.

Some studies have explored the production of idiomatic variants. Cutting and Bock (1997) conducted a series of experiments investigating the production of formal idiom blends. They presented two idioms on the computer screen, and after a two-second delay asked for one of the idioms to be produced. These experiments attempted to induce blends, which have traditionally been regarded as 'errors' or 'slips of the tongue' (cf. Fay, 1982). Among their results, they found that both the syntax and semantics of the idioms can affect the production of blends – blends are more likely to be produced if the syntax or the semantics of the two idioms are similar. More recently, Konopka and Bock (2009) investigated production of lexicalized expressions, specifically phrasal verbs which could have an idiomatic interpretation (e.g. *pull off a robbery*) or a literal interpretation (e.g. *pull off a sweater*). They primed target sentences with a sentence containing another phrasal verb to determine if the syntax of the target would be altered to the structure of the primed sentence (i.e. the particle occurring after the verb or after the object). They found that phrasal verbs were influenced by the prime and produced with a similar structure, regardless of whether they had an idiomatic or a literal meaning; thus, idiomatic phrasal verbs can be altered as much as non-idiomatic phrasal verbs.

While these studies are important for theories of speech production, they focus on slips of the tongue and do not account for deliberate, intentional variation, nor the

range of variation which can be naturally produced by speakers. This study therefore does not focus on errors we might make when speaking, but rather focuses on the range of idiomatic variation which can be spontaneously and consciously produced by speakers of English and how this variation compares with what is observed in corpora. In other words, how creative can speakers be with idioms when explicitly asked?

One way to elicit idiomatic creativity is to ask participants to produce newspaper headlines. Glancing through a local newspaper (i.e. Edmonton Metro), one can immediately see that idioms are used in headlines to attract reader's attention and interest (cf. Mostafa, 2010), such as *Birds of a feather say no to Winnipeg's weather*, *Open-air kitchen and kaboodle*, and *Legit massage therapists rubbed the wrong way*. Providing participants with an opportunity to be as creative as they like should confirm already known uses of idioms, as well as reveal novel uses. Thus, the current study explores the following research questions:

Research Questions:

1. What is the range of variation produced?
2. Are certain types of variation more frequent than others?
3. How do elicited variants compare to what is observed in corpora?
4. How creative are speakers with idioms when prompted?
5. What factors influence the use of variation?

## 5.2 Methodology

### 5.2.1 Materials

Sixty idioms were selected from the Oxford Dictionary of English Idioms (Ayto, 2009) and the Collins COBUILD Idioms Dictionary (Sinclair, 2011). Definitions for these idioms were also extracted from these dictionaries. The idioms varied in length, from two to eight words, and syntax, including verb phrase idioms (e.g. *call the shots* or *let the cat out of the bag*), prepositional phrase idioms (e.g. *down the*

*drain* or *between a rock and a hard place*), noun phrase idioms (e.g. *a needle in a haystack* or *best thing since sliced bread*), and even sentential idioms (e.g. *it's not over until the fat lady sings*). Each idiom was paired with a newspaper snippet. Each newspaper snippet ranged between three and six sentences in length. These snippets were extracted from the Strathy Corpus of Canadian English (Strathy Language Unit, 2013) and the Corpus of Contemporary American English (COCA; Davies, 2008), specifically from the newspaper genre. These corpora were selected for two main reasons. First, the Strathy Corpus represents Canadian English and since the participants would mostly be speakers of Canadian English, it seemed appropriate to provide them with contexts from Canadian newspapers. However, the Strathy Corpus is quite small by today's standards, containing 50 million words, which may not have newspaper stories for all chosen idioms. Therefore, COCA was selected to complement the Strathy Corpus when contexts were not available or not ideal.

### 5.2.2 Procedure

Participants were asked to create headlines for newspaper snippets. They were told to pretend they were the journalist who had just written the article and now must produce a headline. They were asked to include the provided idiom in their headline, but told that it did not have to be exact. They were encouraged to be creative.

The experiment was created using E-prime 2.0 Professional. The text was presented in a bold, white, Courier New 24-point font on a black background. Before each newspaper snippet, participants were asked if they knew the idiom (i.e. 'yes' or 'no'). If they did not know the idiom, a definition was provided. They then saw the newspaper snippet, along with the idiom, presented on the one screen. The snippet was positioned at the top of the screen, spanning 75% of the screen's width. The idiom appeared below the snippet, in the centre of the screen, identified by "Expression". The text box appeared below the idiom, in the bottom-half of the screen. The box had a white 5-point border, and produced white text inside the box. Participants could use the backspace if needed. They were required to write their headline in the text box, before continuing on to the next idiom-snippet pair. The experiment was self-paced. All participants had an opportunity to take a short break halfway through the experiment.

All participants saw the same 60 idiom-snippet pairs, presented in random

order. However, three versions of the experiment were created, which only differed in the example headline provided. The example idiom and snippet were the same for all versions, shown in (5.1). Three example headlines were created, shown in (5.2). The first headline (A), uses a partial form of the idiom (i.e. *mouthful of foot*) to allude to the whole idiom. The next headline (B), blends two idioms: *born with a silver spoon in one's mouth* with *put one's foot in one's mouth*. The third headline (C), utilizes lexical variation (i.e. *boot* instead of *foot*). After the example, participants proceeded to the practice, which consisted of four headlines.

(5.1) Example Snippet:
U.S. Rep. Charles Rangel told the New York Times: "Mississippi gets more than their fair share back in federal money, but who the hell wants to live in Mississippi?" It led to a firestorm of newspaper attention. "I was trying to explain why the federal government gives more to a different state," Rangel says. "You have more poor folks in Mississippi than in New York."

Example Idiom:
*put one's foot in one's mouth*

(5.2) Example Headlines:

A. Mouthful of foot: Rangel comment causes uproar
B. Rangel, born with a silver foot in his mouth
C. Rangel's comment put the whole boot in his mouth

After the experiment, participants were presented with a few additional questions. They were first asked three questions pertaining to their idiom usage: (1) How often do you use these expressions?; (2) How often do others around you use these expressions?; and (3) Do you like using these expressions? They responded to these questions using a 7-point Likert scale (i.e. 1 = 'dislike' and 7 = 'like'). They were next asked to rate the acceptability of seven prescriptively 'incorrect' sentences, shown below, using the same 7-point scale.

Language Questions (LQ):

1. The only option the school board has is to lay off a large amount of people.
2. Slot machines are thought to be more addicting than table games.
3. The document had to be signed by both Susan and I.
4. While cleaning the kitchen, Sally looked up and saw a spider on the roof.
5. I thought it could've went either way.
6. She could care less what he had to say about it.
7. You have to balance your life, irregardless of what anybody thinks.

### 5.2.3 Participants

Thirty undergraduate linguistics students from the University of Alberta participated in this experiment, ranging from 18 to 39 years of age. All participants considered themselves native or near-native speakers of English – twenty-three provided English as their first language. Four were left-handed and 26 were right-handed. All participants were compensated for their time with course credit.

## 5.3 Results

### 5.3.1 Quantitative Analysis

In order to gain insight into the kinds of variation produced by the participants, two types of analyses were performed: a quantitative analysis and a qualitative analysis. I coded the data for three measures: variation in form, variation in syntax, and variation in semantics, to determine which kinds of variation speakers prefer with idioms, as well as which types of variation are most frequent. Coding the data in too detailed a manner can lead to too many factor levels, with few instances per level, which then makes it difficult to make significant predictions. Therefore, the data were coded using a coarse-grained coding scheme to obtain minimal levels. Significant results are still found, indicating which variables are predictive of variation.

To complemement the quantitative results, headlines for select idioms will be discussed more qualitatively, to provide a more fine-grained and detailed account of the variability and creativity with which speakers can use idioms.

**Patterns and Percentages**

Any changes to the form of the expression (i.e. the words of the idiom) were coded in the variable: Variation in Form. This variable has seven levels, shown in Table 5.1, along with the frequency and percentage for each variant type. The most frequent form of the idiom in the headlines was No Variation, or the idiom's canonical form (i.e. the form of the idiom provided to the participants). Participants preferred to use the idiom in its canonical form in about 65% of the headlines, as in *Full steam ahead for Montreal Film Festival.* The most frequent type of variation, Minor Change, involves a slight modification to the idiom's form (i.e. the addition or elimination of one word in the idiom), such as the elimination of *the* in *NRA refuses to shoot breeze, targets current events instead* [shoot the breeze] or the addition of *passionate* in *Dad tells son to keep his passionate nose to the grindstone* [keep your nose to the grindstone]. The addition or elimination of one word was grouped together in this category to avoid two factor levels for a minor change to the form. As shown in Table 5.1, altering the form minimally is the most frequent type of variation, occurring in the headlines around 19%. An idiom was coded as a Partial Form when two or more words of the idiom were eliminated (i.e. only one or a few words of the idiom's form remained), as in *Prison grapevine triggers riot* [hear something through the grapevine], or when the idiom's form was split up and spread out throughout the headline, as in *Beaconsfield sees bandwagon in uniform, refuses to jump on* [jump on the bandwagon]. This type of variation was the second most frequent, around 13%.

The next three types of variation are all used quite minimally, less than 2%. The first is a combination of Lexical Variation (i.e. changing one of the words in the idiom to a synonymous word) and Minor Change, such as *Pooley advises to hedge investments* [hedge one's bets], where *bets* was altered to *investments* (Lexical Variation), but *one's* was also omitted. Since both these types of variation occurred together in the same usage, it was difficult to code as either Lexical Variation or Minor Change, therefore a new factor level was created. And as can be seen from Table 5.1, Lexical Variation occurs more often with an additional change to the form. Lexical Variation by itself, as in *Pete's actions reveal his true colours as a man of*

Table 5.1: Types of form variation and their frequency and percentages used by participants in the elicitation task

| **Type of Form Variation** | **Frequency** | **Percentage** |
|---|---|---|
| No Variation | 1157 | 64.85% |
| Minor Change | 338 | 18.95% |
| Partial Form | 225 | 12.61% |
| Lexical Variation & Minor Change | 25 | 1.40% |
| Lexical Variation | 16 | 0.90% |
| Formal Idiom Blend | 12 | 0.67% |
| No Form Used | 11 | 0.61% |
| | 1784 | 100.00% |

*true principles* [show one's true colours], occur slightly less at 0.9%. The last type of variation observed in the headlines were Formal Idiom Blends, occurring quite infrequently at 0.67%. Formal Idiom Blends involve the blending or merging of two distinct idioms, such as *Intentional spills of the bag deemed the official Chinese voting motto* blending *spill the beans* with *let the cat out of the bag*. Interestingly, this type of variation has often been regarded as a 'slip of the tongue' (Fay, 1982; Cutting and Bock, 1997; Kuiper et al., 2007), but as evident from the headlines elicited, this variant can be used intentionally as a form of creative idiomatic variation, albeit an infrequent one.

The last type of coding was No Form, which was to include the instances where no words from the idiom occurred at all in the headline. This is another interesting category, as it shows an exceptional kind of creativity by the participants, where instead of using some portion of the expression, they simply alluded to it in another way; for example, *Pete Wilson earns new nickname Green at voting* where the word *Green* is used to refer to the idiom *show one's true colours*. This last option was not frequent (only used 0.61%), but does show a unique strategy of idiomatic variation employed by the participants and was therefore included in the analyses.

The next variable, Variation in Syntax, accounted for any changes to the tense, aspect, or grammatical structure of the idiom within the headline. This variable was reduced to nine levels, shown in Table 5.2, along with the frequency and percentage for each variant. No Variation (i.e. no modification to the syntax of the idiom) is the most frequent type, at 62%. No Variation includes the exact form of the idiom,

Table 5.2: Types of syntactic variation and their frequency and percentages used by participants in the elicitation task

| **Type of Syntactic Variation** | **Frequency** | **Percentage** |
|---|---|---|
| No Variation | 1112 | 62.33% |
| Progressive | 276 | 15.47% |
| Partial Form | 158 | 8.86% |
| Inserted Modifier | 71 | 3.98% |
| Past | 50 | 2.80% |
| Other | 45 | 2.52% |
| Adjectival | 31 | 1.74% |
| Voice Alternation | 27 | 1.51% |
| Participial | 14 | 0.78% |
| | 1784 | 100.00% |

as provided to the participant, with no syntactic modifications, such as *Nations turn a blind eye to human rights in favour of security* [turn a blind eye]. It also includes third person agreement with a third person subject, as in *River Valley Health pulls the plug on Facebook use in the workplace* [pull the plug]. Third person agreement was included in No Variation since the change is simply to make the sentence grammatical (i.e. to agree with the subject), and to reduce the number of factor levels.

The most frequent type of syntactic variation is with the progressive form of the verb, as in *Company leaders tired of playing second fiddle* [play second fiddle]. This occurs in the headlines around 15%. Partial Form was included as a category of syntactic variation for when the idiom was reduced so much as to no longer have the same syntax, let alone a discernable one, such as *Recent resident counts himself lucky to still have nose* [keep one's nose to the grindstone]. This type of variation occurs in about 9% of the headlines. About 4% of the headlines have a modifier inserted directly into the idiom itself, as in *Habitat back to the tough drawing board* [back to the drawing board]. The next type of syntactic variation is Past Tense, at almost 3%, as in *Looking Back on 52 Years: Didden Kept a Tight Rein on the Bank* [keep a tight rein on something]. Other, at 2.5%, includes several types of variation which occur so infrequently that they are grouped together, such as nominalizations like *Biskind, the beans spiller* [spill the beans] and alternations like *Locals Pick Bone with City over Park Concert* [have a bone to pick with someone]. Adjectivals, like *Goodenow*

*claims No Behind the Back Dealings* [behind one's back], and Voice Alternations, like *Cigarettes, the bud that should have been nipped* [nip something in the bud], occur in less than 2% of the headlines, while Participials, like *Van Halen Claims to Have Bitten Off More than They Can Chew* [bite off more than one can chew] occur in less than 1%.

The last variable, Variation in Semantics, was much more difficult to code. Therefore fewer variable levels were created in an attempt to code for the most different or obvious examples of changes or additions to the idiom's meaning. The frequency and percentages of each variant type are shown in Table 5.3. As with the other types of variation, participants used the idiom with its intended original meaning (i.e. No Variation) in 60% of the headlines; for example *A needle in a haystack: Astronomer searches for space rock remains* [a needle in a haystack] or *Successful conductor doesn't beat around the bush* [beat around the bush].

An Integrated Concept is the most frequent type of semantic variation, occurring in 22% of the headlines. This can involve an additional word integrated into the idiom, such as *Ontario RNs Chomping at Detailed Bits of Union Agreement* [chomp at the bit], or the modification of one of the words to integrate an additional concept, as in the name *Serge*, the director of the festival in the newspaper snippet, being incorporated into this headline: *Full Serge Ahead* [full steam ahead]. The next frequent category, occurring just over 9%, is an Allusion to the idiom. This can occur from select words of the idiom (and possibly additional words included in the headline, like *scraped*), as in *Student's scraped nose pays off* [keep one's nose to the grindstone], or through only other words in the headline alluding to the idiom (i.e. when no words of the idiom are utilized), such as *Kerr's restaurant definitely not rounded* [cut corners].

Play on Literal, at 4%, coded for when the participants played on the literal meaning of one of the words in the idiom, such as *Hall of Famers pass muster, not pucks* [pass muster], where the literal meaning of *passing* is required to understand the use of *pucks* in the headline, or *NRA: shoot guns, not the breeze* [shoot the breeze], where the participant is highlighting the mission of the NRA (i.e. National Rifle Association), an organization advocating gun rights. This differs from a Play on Words, in that this latter category accounts for the creativity used with homophonous words, as in *Chicago is coming apart at the seams it seems* [come apart at the seams], or *Wise to Take a Rain Check before Handing Over Your Cheque at Trump*

Table 5.3: Types of semantic variation and their frequency and percentages used by participants in the elicitation task

| **Type of Semantic Variation** | **Frequency** | **Percentage** |
|---|---|---|
| No Variation | 1075 | 60.26% |
| Integrated Concept | 392 | 21.97% |
| Allusion to Idiom | 163 | 9.14% |
| Play on Literal | 73 | 4.09% |
| Play on Words | 54 | 3.03% |
| Idiom Not Used | 27 | 1.51% |
| | 1784 | 100.00% |

*International Hotel* [take a rain check]. Finally, a few headlines, 1.5%, seemed to not include the idiomatic meaning, where the participant presumably forgot to include the idiom or the allusion to it was less obvious, such as *Chicago neighborhoods literally cracking at the foundation* [come apart at the seams], or they used the idiom literally instead of with its figurative meaning, *Do a literal rain check before heading to the Trump International Hotel* [take a rain check].

Coding these three variables was an attempt to tease apart potential differences in variations and to distinguish which types are most frequently produced by speakers of English when specifically asked to be creative. However, as evident by the results, there is considerable overlap between these variables. For example, an idiom which was coded as a Partial Form under Variation in Form, may have also been coded as a Partial Form under Variation in Syntax. If that was the case, it was most likely coded as an Allusion under Variation in Semantics. Patterns such as that were certainly present, which is why the results show similar trends (e.g. No Variation occurring around 60% for all variables). Nevertheless, these results illustrate the prevalence of the canonical form, even when encouraged to be creative, as well as the preferred use of slight modifications to elaborate on the meaning or partial forms to allude to the idiom. Other types of variation are utilized much less frequently.

### Logistic Mixed-Effects Regression

The majority of headlines included the canonical form of the idiom; the form provided to the participants. The remainder of instances were distributed among different types of variation. In order to model such data, I created binary variables to indicate

whether variation was present in the data. Three binary response variables were analyzed (i.e. form, syntax, and semantics) using logistic mixed-effects regression, using the `lme4` package (Bates et al., 2014) in `R` (R Core Team, 2014).

Five predictor variables were significant in the three models. `Length` indicates the number of words in the idiom. `logFrequencyIdiom` is the log frequency of the idiom in its canonical form, extracted from COCA (Davies, 2008). `Version` is a factor indicating the three versions of the experiment. The versions only differed in the example headline that was provided: Version A = headline included an allusion to an idiom; Version B = headline included two idioms blended together; and Version C = headline included a lexical variant of one of the words in the idiom.

Two participant-related variables were also included in the models. `KnowIdiom` is a factor indicating whether the participant knew the idiom (i.e. 'yes' or 'no'). And `LQ7`, the participant's acceptability rating for Language Question 7 (i.e. *You have to balance your life, irregardless of what anybody thinks*). Participants were asked to rate the acceptability of seven grammatically 'incorrect' sentences in order to obtain an approximate measure of the participant's 'permissiveness' or 'flexibility' with language. This was done to determine whether a more relaxed attitude towards grammar and language usage influences the participant's spontaneous production of idiomatic variation. If the participant is more flexible with language, they may also be more likely to produce variation.

Both `logFrequencyIdiom` and `LQ7` were chosen in the analyses because they produced a lower AIC value[1] than other measures of frequency or the other Language Questions, respectively. These variables were also compared with respective Principal Components (PCs) from a Principal Components Analysis (PCA). PCA "reduces the number of dimensions required for locating the approximate position of the data points" (Baayen, 2008, p.120). Two PCAs were performed, one on the frequency variables (which consisted of 60 idioms, or data points, and 36 frequency measures, or dimensions, including the frequency of the individual words, frequency of adjacent words, and frequency of the idiom), and one on the language accept-

[1] AIC (i.e. Akaike's Information Criterion) is a measure for how well the predictor variables improve the predictiveness of the model (Akaike, 1974, 1985), and is used a method of model comparison. Given the number of parameters in each model, the difference of the entropies of two models (i.e. the negative of the entropy is often regarded as the amount of information) is measured using a log-likelihood ratio test. Thus, the model with the smallest AIC value is the preferred model.

ability questions (which consisted of the 30 participants, or data points, and the 7 language questions, or dimensions). The data points remain fixed at their locations, while new dimensions, or Principal Components (PCs), are determined, which more succintly explain the data and account for the same amount of variance as the original variables. However, `logFrequencyIdiom` and `LQ7` produced lower AIC values than PC1 or PC2 from respective PCAs and were therefore included in the models.

Three logistic regression models (i.e. Variation in Form, Variation in Syntax, and Variation in Semantics) are shown in Appendix D. These three analyses show similar findings – the same predictors are significant and show similar effects – and therefore all three models will be discussed together. Figure 5.1 shows the interaction and main effects from these models.

`Length` is only significant with Variation in Form, as both a main effect and in an interaction with `LQ7`. The top-left panel of Figure 5.1 shows the main effect of `Length`: the longer the expression, the more likely the participants are to vary the form. This of course makes sense, more options exist within the form for speakers to choose to vary, while more information is conveyed in the longer expression which may help the listener retrieve or infer the idiomatic meaning despite the variation. Participants are however less likely to vary the form the more they find `LQ7` (i.e. *irregardless*) acceptable, shown in the bottom-left panel of Figure 5.1. When these two variables interact, participants are significantly less likely to produce variations, even when the idiom is long, if they find `LQ7` acceptable.

The models for Variation in Syntax and Semantics also show this main effect of `LQ7`; participants are less likely to produce variation if they find `LQ7` acceptable. This result initially seems surprising – the greater acceptability of a non-standard form is predictive of less idiomatic variation produced by the participants. However, this form (i.e. *irregardless*) blends *regardless* and *irrespective* making the form less compositional but perhaps conveying a more obviously negative interpretation. This clearer intended meaning perhaps is what is predictive of idiomatic variation – speakers who like `LQ7` are less likely to vary the form of the idiom as they prefer forms which are more clearly interpretable.

`KnowIdiom` and `logIdiomFrequency` are significant in an interaction in both the Variation in Syntax and Variation in Semantics models. Since this interaction shows the same effect in both models, it has only been shown once in the top-right panel of Figure 5.1 from the Variation in Semantics model. The grey line indicates

that the slope for those who do not know the idiom is not significant. Participants who know the idiom are significantly less likely to produce variation when the idiom is frequent. Participants who do not know the idiom are not affected by its frequency, as expected, and produce more consistent amounts of variation. Frequency is significant in the model for Variation in Form; however, it is not significant in an interaction with `KnowIdiom`. Participants are in general less likely to produce formal variation with an idiom that is more frequent.

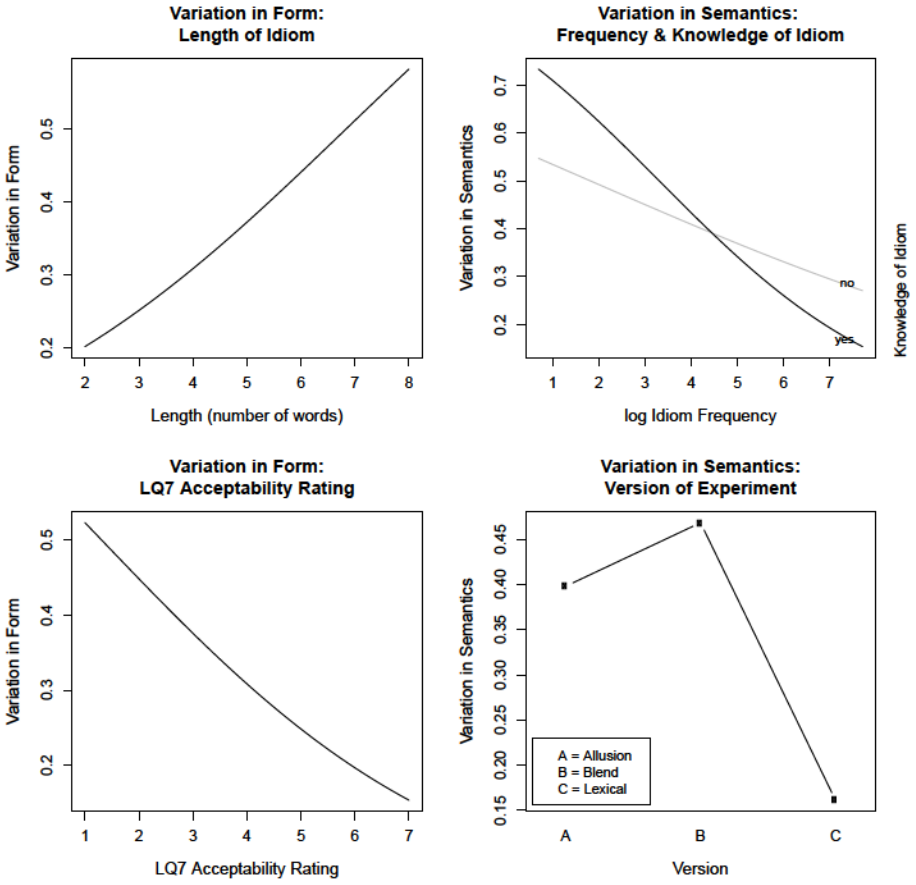


Figure 5.1: Interactions and Main Effects observed in the Mixed-Effect Logistic Regression Models for Variation in Form, Variation in Syntax, and Variation in Semantics. Lines in grey represent non-significant slopes.

Lastly, the `Version` of the experiment is significant in all three models. Participants exposed to the example headline in Version C (lexical variation) produce

significantly fewer instances of variation in their headlines than those exposed to the example headline in Versions A (allusion) and Version B (formal idiom blend). The headline examples in Versions A and B are more creative than simply changing one word in an idiom to a synonymous variant, and this creativity may in turn inspire the participants to be equally as creative.

### 5.3.2 Qualitative Analysis

In this section, I look specifically at seven select idioms and discuss the range of variation and creativity produced with each of these idioms in the elicited newspaper headlines. These seven idioms were selected based on the observed differences with the types of variation that each idiom is produced with. Each idiom demonstrates different usage patterns of variation, worthy of discussion. I also compare the experimental results to corpora. This was largely done to compare whether the same variation strategies observed in these results are also evident in corpora, or if the headlines that participants produced in this study were in some way task-dependent.

#### Call the Shots

Some idioms, such as *call the shots* and *pull the plug*, show minimal variation in the headlines. The variation which is present is largely syntactic variation, specifically the form of the verb has been modified to reflect tense or aspect. The variation evident with *call the shots* is especially minimal, occurring in either the progressive form *calling* (see Example 5.3) or in the present tense form (see Example 5.4). Approximately half of the headlines produced by the participants occur with the progressive form, while the remainder occur with the present tense form. All forms of the idiom are of the structure *call the shots*, with no alternations of the idiom besides altering the aspect of the verb.

(5.3) a. Billionaires calling the shots

b. Who is Really Calling the Shots?

c. Are Billionaire Sponsors Really Calling the Shots in our elections?

d. Cut out the Middle Men: We Know it's the Billionaires Calling the Shots

(5.4) a. Billionaires call the shots

b. Billionaires who call the shots should have to explain them

c. Billionaires Play Puppeteer and Call the Shots

d. Rich sponsors call the shots on presidential campaigns

*Call the shots* was searched in the Strathy Corpus of Canadian English (Strathy; Strathy Language Unit, 2013) and the Contemporary Corpus of American English (COCA; Davies, 2008). The lemma of *shot* was queried in the context of four words to the right or four words to the left of the lemma *call*. In Strathy, only 18 instances were found with *call* and *shot* in close proximity, and only seven of these concordance lines contained the idiom *call the shots* (cf. non-idiom uses like *whether a penalty shot was the appropriate call* or *Quebec nurses call for flu shots*). Interestingly, no instances of the idiom in its progressive form were observed; only occurrences in the past tense (i.e. *called the shots*) and the present tense (i.e. *call the shots*) were found. Meanwhile, a random sample of 100 concordance lines from COCA, showed a similar pattern to what was observed in the headlines: half occurred in the progressive form and half occurred in the present tense form. However, additional variation was also present in these concordance lines, such as omission of the determiner (e.g. *calls shots*, *calling shots*), playing on the literal meaning of the words (e.g. *Health Column "Calling the Shots", about parents who refuse vaccinations for their kids*), and even integrated concepts (e.g. *call the recycling shots*, *calling some shots*, *called the critical shots*).

**Cross that Bridge When You Come to It**

Idioms like *cross that bridge when you come to it* and *let the cat out of the bag* occur with a range of variation. The instances in Example 5.5 illustrate this array for *cross that bridge when you come to it*. More than half of the headlines occur with the full form of the idiom, shown in (5.5a). However, variation is still evident in these forms. The full form can include an additional integrated concept, such as that seen in (5.5b), or even contain an ending which has been modified in some way, as in *when he gets there* shown in (5.5c). The remainder of the headlines included a partial form of the idiom; this may be a truncated form of the idiom (i.e. *cross that bridge*), seen in (5.5d), or it may be so reduced as to only include a single word (i.e.

*bridges*), observed in (5.5e). Integrated concepts can also occur with partial forms, as in (5.5f). This example also illustrates syntactic variation, which can include passives, like the example in (5.5f), as well as tense or aspect changes (e.g. *crossed, crossing*). While the future in English does not change the form of the verb, many of the headlines include the future auxiliary *will* (or *'ll*), as one might anticipate given the meaning of the expression. Finally, one headline includes a formal idiom blend, shown in (5.5g), merging the expressions *cross that bridge when you come to it* with *one X at a time* (e.g. *one day at a time*).

(5.5) a. Retiring from the Game: David O'Brien Says He'll Cross That Bridge When He Comes to It

b. O'Brien to Cross Retirement Bridge When He Comes to It

c. Certainty is not in David O'Brien's future: He will cross that bridge when he gets there

d. No talk of contracts until they need to cross that bridge

e. David O'Brien is ready for future bridges

f. O'Brien waits for his contract bridge to be crossed

g. O'Brien crossing *one* bridge *at a time*

I searched the collocates *cross* and *bridge* in their lemmatized forms for four places to the left or right of each other in both Strathy and COCA. Interestingly, this search query resulted in few tokens with the idiomatic meaning: only eight instances in Strathy and eleven in COCA (from a random sample of 100 concordance lines). The majority of occurrences instead were with the literal meaning of 'physically crossing a bridge'. Nevertheless, a range of variation is still visible: from truncated forms (e.g. *cross that bridge*), adjectival modifiers (e.g. *little bridges to cross, an important bridge to cross*), modified endings (e.g. *when we get to it, if you come to it, when the time comes*), and even syntactic variation and alternations (e.g. *crossing / crossed that bridge*, *bridge to be crossed*, and *This is a bridge that's been crossed long ago*). Despite the limited number of occurrences in both corpora, it is interesting to note that a similar range of variation is found with this idiom.

**Mend Fences**

Maximal variation does not only occur with longer idioms, which presumably might allow for more possible variations. Shorter idioms like *mend fences* and *cut corners* can also show a range of variation. Half of the headlines show no variation, where an unaltered canonical form occurs, as shown in (5.6a). Most of these forms that occur with 'no variation' appear in a larger verb phrase with the *to*-infinitive (e.g. *attempts to mend*, *set to mend*, *trying to mend*). Syntactic variation is also noticed, altering the form of the verb or the syntax of the expression, such as the passive in (5.6b) or progressive aspect in (5.6c). Observations are also present, which play on the literal meaning of *fence*, as in Example (5.6d), incorporate an additional concept into the idiom, seen in (5.6e), or utilize lexical variation, as shown in (5.6f).

(5.6) a. Japan taking steps to mend fences with China

b. Economic fences mended between Tokyo and China

c. Japan mending fences with China

d. China and Japan mend fences by tearing them down

e. Japan attempts to mend damaged fences with China

f. Mending the border between China and Japan

COCA and Strathy show a comparable range of variation. The lemma *fence* was queried, occurring within four words to the right or four words to the left of the lemma *mend*. Variation with this idiom in the corpora appears largely with syntactic alternations: progressives (*mending fences*), past tense (*mended fences*), as well as past participles (*have mended fences*), passives (*fences have been mended*), and even infinitive alternations (*fences to mend*). However, other types of variation were also observed. Additional concepts were inserted into the idiom (e.g. *mended her Nashville fences*), which were predominantly either *some* or *political*, or both (e.g. *mend some political fences*). Nominalizations were also present (e.g. *a lot of fence mending to be done* and *Peaceful relations with others depend on the mending of fences*). Examples were even observed where the figurative use of the idiom was explicitly stated (e.g. *mend fences – the metaphorical kind*), or where the literal meaning of *fence* was played upon (e.g. *In Jasper, Texas, folks are mending fences by tearing one down*).

In an attempt to find examples of lexical variation, both *mend* and *fence* were queried without the other in Strathy. Zero instances of lexical variation were observed in approximately 600 concordance lines of *fence*, while *mend* did occur with other nouns besides *fence* to convey peace-making (e.g. *mend the divide between Europe and Washington, mend a problem caused by a previous government, mend the ethnic and political fractures, mending the constitutional wrangle*). However, the nouns in these contexts may simply be extending the semantics of *mend* into a more abstract domain and not necessarily have anything to do with the idiom *mend fences*. One striking pattern with *mend fences* was observed, its predominant use with the *to*-infinitive. This *to*-infinitive can occur with a range of different verbs (e.g. *tried to mend, hoping to mend, need to mend, wants to mend, sought to mend, attempts to mend, urged to mend, designed to mend, determined to mend, works to mend, scrambled to mend*), nouns (e.g. *an effort to mend, a chance to mend, an opportunity to mend, a plan to mend, the decision to mend, a mission to mend*), and even adjectives (e.g. *eager to mend, ready to mend, desperate to mend*). This striking use of the *to*-infinitive in corpora makes its frequent use in the headlines even more apparent, occurring in more than half of the headlines.

**Jump on the Bandwagon**

Idioms like *jump on the bandwagon* and *hear something through the grapevine* also occur with a range of variation, similar to the above idioms, but instead of occurring predominantly with syntactic variation, this idiom occurs largely with inserted concepts and a richness of creativity mostly due to the salience of *bandwagon*. The majority of the headlines utilize the full form of the idiom, as shown in Example (5.7a), although some show a minor alternation, specifically the omission of *the*. Meanwhile, half of the headlines occur with an additional concept that has been integrated into the idiom. Example (5.7b) shows the incorporation of *snazzy lookin'*, which may reflect the participant's attitude towards uniforms in school, while Example (5.7c) includes the use of *uniform* itself – the most-utilized integrated concept in these headlines. Example (5.7c) also shows an example of lexical variation, one of two instances, both of which alter the verb. This example and Example (5.7d) show syntactic variation, particularly the use of the progressive aspect, which was the most frequent type of syntactic variation.

The remaining examples show the considerable degree of creativity that par-

ticipants can utilize with idioms like *jump on the bandwagon*. The salient word *bandwagon* in (5.7e) has been split away from the rest of the idiom (i.e. *jump on*), but more importantly, blended into a rich imagery of the *bandwagon* dressed in a *uniform*, playing on the notion that bands, which traditionally rode on these wagons, dressed in band uniforms, meanwhile conveying that a particular high school chose not to implement uniforms as a dress code. Example (5.7f) utilizes a partial form of the idiom (i.e. *bandwagon*), but in a context that expresses the opposite notion of the idiom – *staying off* as opposed to *jumping on*. The third instance of a partial form, seen in (5.7g), has drastically reduced the idiom so that all that appears is a portion of the salient word (i.e. *wagons*). This usage not only conveys the idiom, specifically that support for school uniforms is spreading through the city of Montreal, but also that physical wagons decorated in uniforms might be travelling alongside the hype. The last example, (5.7h), shows an instance of a formal idiom blend, blending the expression *X is the new black* with *jump on the bandwagon*. This example nicely merges the popular, support-driven *bandwagon* with the latest fad, *the new black*. Finally, creativity is also observed in the way the participants describe the lack of *jumping on the bandwagon* by the particular school in the newspaper snippet (i.e. Beaconsfield High School). This school *refuses to jump on, stays off*, and even *misses the bandwagon*. Additionally, Beaconsfield *jumps on the bandwagon against uniforms, says no*, and simply *won't be joining*. The range of variation in describing how this school declined *jumping on the bandwagon* is itself noteworthy.

(5.7)
a. Beaconsfield High School refuses to jump on the bandwagon: students and parents voted No to introducing uniforms into the school
b. Beaconsfield High not to jump on this snazzy lookin' bandwagon
c. Beaconsfield won't be joining the uniform bandwagon
d. Montreal schools jumping on the uniform bandwagon – Beaconsfield says no
e. Beaconsfield sees bandwagon in uniform, refuses to jump on
f. Students and parents vote to stay off bandwagon
g. Uniformed wagons making way down Montreal
h. Uniforms as *the new* bandwagon?

This idiom occurs with a vast amount of variation and creativity in corpora as well. Only the word *bandwagon*, in its lemmatized form, was queried in both

Strathy and COCA. Among the results, a wide and varied range of lexical variation is observed, especially for the verb *jump*. One can *jump, join, hop, leap, board, get on* and even *be lured on* the bandwagon. The manner in which one boards the bandwagon can also be embellished (e.g. *climb, pile on, scramble, clamber*). However, one can also *miss, decline to leap on, eschew, fall off, get off*, or *be left off a bandwagon*. One can also *cling tightly to a bandwagon*. People can play an active role in creating bandwagons (e.g. *launch, start, drive, drum up*), as well as a less active role (e.g. *ride, sign on, be back on* and even *give lip service to the bandwagon*). People may feel *gleefully* or *blissful* about jumping on the bandwagon, but they can also view the jump as *hastily* or *unthinkingly*. They can jump *on, onto, into, aboard, off* and even *toward* bandwagons.

Meanwhile, bandwagons can *fill up* or *become so crowded*. They can *cruise, travel*, or *be headed* in a particular direction. Although sometimes the direction of the bandwagon does not matter (e.g. *we're not exactly sure where the bandwagon's leading* or *I want a good seat on the bandwagon*). The speed at which a bandwagon travels can also be stressed (e.g. *just beginning to roll, gained momentum, picked up astonishing speed*, and *is now unstoppable*). But bandwagons can also experience difficulties along the way (e.g. *slow down, appears to have stalled, hit a bump, derailed*, or even be *stuck on the shoulder of the road*). Bandwagons can be created (e.g. *his built-on-the-fly bandwagon*), described in terms of their size (e.g. *jumping on what has been a rather small bandwagon*) or quantified (e.g. *people will jump right back on. I may need two bandwagons*).

Numerous concepts can be integrated into the idiom, describing the type of bandwagon that it is (e.g. *victorious, political, cordless, fat-free, Twitter, Apple, guilt-trip, faith-based, eco, misdirected, profit-producing*). However, the word *bandwagon* itself can also be used as a modifier, from the more obvious *bandwagon jumpers* or *fancy bandwagon jumping*, to being used in extended metaphorical contexts, modifying people (e.g. *bandwagon man, bandwagon fans, bandwagon investors*), modifying abstract concepts (e.g. *bandwagon approach, bandwagon effect, bandwagon momentum, bandwagon proportions*), and even as an explicit metaphor (e.g. *the party is a bandwagon around one man*). As evident from the plethora of examples from COCA and Strathy, *jump on the bandwagon* is a very productive idiom, manipulated in numerous ways, for various creative and pragmatic effects.

**It's Not Over Until the Fat Lady Sings**

Some idioms have more than one salient word, such as *it's not over until the fat lady sings*, where *fat, lady*, and *sing* all appear salient. These idioms have few inserted concepts, compared with idioms like *jump on the bandwagon*, and instead tend to replace one or more of the elements in the idiom or are used with a greater array of partial forms. Full forms (i.e. *it's not over until the fat lady sings*) appear in about half of the headlines, but most occur with a slight modification to the form, such as *it's not over* becoming *it isn't over*, *it ain't over*, and even *not over*, as illustrated in Example (5.8a). Furthermore, a new noun phrase is often incorporated into the initial subject position (i.e. replacing *it* with a noun relevant to the current situation), with either *the proposal*, as shown in Example (5.8b), or *the fight*. Modification of a noun phrase is also observed with *the fat lady*, where this NP is replaced with one that is more relevant to the current situation, by linking the potential singer(s) to a specific individual(s), such as the example shown in (5.8c).

Partial forms of the idiom are also commonly used. These partial forms contain elements of varying lengths from the second half of the expression, such as *until the fat lady sings*, seen in (5.8d), or only with the elements *fat, lady,* and *sing*, as in (5.8e). This latter partial form typically includes these three elements in a negative past participial construction (i.e. *the fat lady hasn't sung*), although they can also occur in other grammatical constructions, such as *the fat lady may still sing* or *when will the fat lady sing*. Other partial forms can be utilized, such as those containing only the noun phrase *the fat lady* (see 5.8f) or only the verb *sing* (see 5.8g). In these instances, reference is made to the fact that she or the singing cannot be *heard yet*, or alternatively, that someone chooses not to *listen*. The majority of headlines convey the idea that the *fat lady* has not begun her act (i.e. *singing*) as the idiom suggests, but one headline manipulates this idea to convey that she is being ignored (i.e. *refuse to listen*), while a different headline conveys that she is still in the process of performing and not quite finished (i.e. *the fat lady is still singing*).

(5.8) a. Christian Right lobbyists claim that it isn't over until the fat lady sings

b. The proposal is not over until the fat lady sings

c. Not over until lawmakers sing, lobbyists say

d. Statehouse Regulars are keeping pressure on lawmakers until the fat lady sings

e. The fat lady hasn't sung yet, she's simply heavily sedated

f. Opponents Refuse to Listen to the Fat Lady

g. No singing heard yet at the legislative session

The variants observed in the headlines overlap with Barlow's (2000) corpus-based study, as well as the findings in Strathy and COCA. For example, Barlow reports that a new noun phrase can be integrated into the subject position in the idiom schema. This is observed in both the headlines, seen in Example (5.8b), as well as in COCA (e.g. *The game ain't over until the fat lady sings*). The expression can also be linked to specific individuals in the current situation. All three resources show examples of this, yet all three are done in a slightly different way. In Barlow's example, *the fat lady* is linked to an actual woman, although further clarification is made (cf. *Nothing is dead until the fat lady sings, as they say. LuJuana is not fat, but she will make the final decision*). In the headlines however, *the fat lady* was replaced with *lawmakers* (see Example 5.8c), more directly linking the expression to the current situation within the snippet. Meanwhile in COCA, an example is observed where an actual woman is pointed out and referenced as *the fat lady*, literally demonstrating while metaphorically implying that she has not started singing and therefore the process is not over (e.g. *We have with us The Fat Lady. Could we show you the fat lady? There she is. The fat lady is not singing.*).

Two other observations are noticed between the headlines and the corpus results. First, a partial form of the expression linking the opera world to the current situation while playing on the various stages of the singing performance (cf. *a long way from breaking into song* or *the fat lady is warming up* from Barlow, 2000). This is observed in the headlines as well, from *Sedated fat lady may still sing*, where there is uncertainty as to whether she will even sing, to *Opponents refuse to listen to the fat lady*, suggesting that she has already started singing. COCA shows even more examples of linking the two worlds through various stages of the singing performance, shown in Example (5.9a–d). The second observation is reference to *the fat lady* to signal the end of a process, often with negative consequences, such as *the fat lady burst into song far too early last Sunday* (Barlow, 2000) or *Fat lady hasn't sung regarding the proposal* from the headlines. Further examples from COCA are shown in Examples (5.9e–g), with particular emphasis on the negative consequences highlighted in (5.9f) and (5.9g).

(5.9) a. If we lose a couple more, the fat lady will start clearing her throat.

b. It's not over till the fat lady sings. She isn't singing, but she's certainly walking to the microphone.

c. I think the fat lady just sang.

d. The fat lady refused to sing when she discovered the theme.

e. The proposal isn't dead, just heavily sedated. The fat lady hasn't sung on this.

f. Who knows how long it would be before the fat lady sang to the wrong person?

g. There she is dying and she's singing her heart out. Fat Lady is surely a far better image for death than that hooded fellow with the scythe.

Some findings are only observed in the corpus results, and not in the headlines. For example, Barlow (2000, p.331) mentions that the full expression "may be followed by a conjoined phrase in which links are established to the current situation". This is seen in COCA; for example *I'm the fat lady in this town and I'm not singing yet.* However, this conjoining of sentences is not observed in the headlines, most likely due to the fact that the participants were asked to produce headlines and not necessarily sentences. Instead, the full idiom is linked with the current situation in a dependent clause (seen in Example 5.8a) or through juxtaposition of two parts through the use of a colon (e.g. *Christian Lobbyists and Legislative Opponents: It's Not Over Till the Fat Lady Sings*). In addition, explicit reference to the opera world is also not observed in the headlines. This strategy is however observed in Barlow's (2000) corpus study, as well as in the Strathy Corpus (e.g. *George Bush kept insisting, "the opera isn't over 'til the fat lady sings." On Tuesday night, she ended the first act*).

New findings however are observed in the headlines. A partial form of only *sing* was observed in the headlines (see Example 5.8g), where no mention of *the fat lady* occurred. Barlow (2000) focused his discussion specifically on the use of *the fat lady* and therefore did not present any similar examples where only *sing* occurred alluding to the idiom. Two concordance lines were discovered in a random sample of the lemmatized form of *sing* in the context of *over* in COCA: 1) *When it comes to real country music, it ain't over till the fat man sings*, and 2) *Schmid, who enjoys using the slogan, "It ain't over til the audience sings," is confident that*

*the new volume of songs will fuel the renewed fires of community singing.* However, both of these examples reference the whole idiom, and do so in a way that plays on the literalness of the word *sing*. In other words, these idioms are used for effect in these contexts to highlight the excellent singing of a particular fat man in the first context, and the excellence of the new volume of songs in the second. While these two examples illustrate variants of *the fat lady* in the idiom, they do not demonstrate the same manipulation seen in Example (5.8c), where the singing of *the lawmakers* signals the conclusion of a particular process, similar to *the fat lady* in the original.

New findings are also observed in COCA. First, this idiom can be modified to communicate that the intensity of the singing corresponds to the degree of finalization of the event, as in *We can't claim victory until the fat lady really sings.* Second, two instances are observed where *it's not over until the fat lady sings* is utilized with another expression of similar meaning: 1) *Until the fat lady sings, I'm not going to throw the towel in*; and 2) *One day you heard the fat lady sing and you knew the words were, "that's all she wrote, folks"*. These are not examples of idiom blending, where the forms are merged together, but are examples where the two expressions have been juxtaposed for pragmatic effect and emphasis.

**Down the Drain**

While prepositional phrase idioms, like *down the drain* and *in the bag*, are not manipulated to the same extent as idioms like *jump on the bandwagon* or *it's not over until the fat lady sings*, they still occur with a range of possible variation and creative uses. The majority of headlines show *down the drain* in its full canonical form, as illustrated in the examples in (5.10). However, variation appears in whether a verb is utilized with the expression, and if so, the type of verb that occurs. In Example (5.10a), the expression is used without a verb; it is simply juxtaposed next to *another season* to signal that the latest season is *down the drain.* This expression occurs with verbs in the majority of instances however, mostly with *to be*, as in Example (5.10b), or *to go*, as in Example (5.10c).

(5.10) a. Another Season Down the Drain? NHL Negotiations Continue

b. This season is completely down the drain

c. Bettman pulls the plug and a new collective bargaining agreement goes down the drain

However, *down the drain* can also be utilized with very manner-rich verbs, as illustrated in the Examples in (5.11). This expression occurs with 10 different manner-rich verbs in the headlines, including two manner-rich verbs which play on the literal meaning of the expression, *washed* and *flushed*, seen in Examples (5.11c) and (5.11d). Meanwhile, Example (5.10c) demonstrates that other information within the headline can also allude to the literal meaning of the expression. In this example, it is the use of a different idiom *pulls the plug*, which is creatively employed, to construe an intimately related literal and figurative reading of the headline. This expression can also be used with prepositions other than *down*, shown in Examples (5.11e–g). The use of a different preposition in this idiom provides a creative and imaginative way to convey the delayed, drawn-out process of the context (i.e. problematic negotiations resulting in a lost hockey season). Another way in which the headlines may allude to the drawn-out process of the negotiations is through adverbs, occassionally used, such as *slowly* in Example (5.11c).

(5.11) a. NHL season – slipping down the drain

b. Is the NHL throwing the season down the drain?

c. Possibilities of NHL season slowly washed down the drain

d. Bettman flushes the season down the drain

e. Negotiations, and the season, sliding *towards* the drain

f. Is this years NHL season headed *for* the drain?

g. Hopes of lockout-free season swirling *around* the drain as sides unable to find common ground

Searching the idiom *down the drain* in Strathy and COCA reveals similar examples of variation. For example, numerous manner-rich verbs are used with this expression in corpora as well. Four main groups of manner-rich verbs appear to be used with this expression: Pour Verbs (e.g. *pour, rinse, flush, wash, dribble, trickle*); Coil Verbs (e.g. *wind, spiral, spin, swirl*); Throw Verbs (e.g. *throw, toss, send*); and Run Verbs (e.g. *run, rush, slide, swish*). According to Levin (1993), these four groups fall into three main Verb Classes: Verbs of Putting (i.e. *Pour* and *Coil* Verbs), Verbs of Throwing (i.e. *Throw* Verbs), and Verbs of Motion (i.e. *Run* Verbs). Interestingly, these three verb classes can all describe or elaborate on the way in which something abstract becomes lost or wasted.

In addition, COCA shows examples of concepts being inserted into the idiom, such as *down the digital drain* or *down the bottomless drain*, something that was not observed in the headlines. The *drain* can also be modified using a following prepositional phrase, as in *down the drain of despair*, *down the drain of self-doubt*, or even *down the drain of bureaucracy, incompetence and corruption.* Two instances appeared in COCA where *down the drain* was blended with another idiom (i.e. *sell someone down the river*), as in *He sold me down the drain* and *Annan really sold us down the drain.* These idiom blends nicely convey the idea of extreme betrayal to the point of becoming lost or wasted. Interestingly, in Strathy, only one instance of *money* went *down the drain*, whereas in COCA, approximately one-third of the concordance lines described *money, finances*, or the *economy* as going *down the drain.* Other abstract concepts can also go *down the drain*, such as *time, years, hard work, hope, morals, dreams*, as well as *political parties, public education*, and *city services.* A quick search of *drain* was conducted in the context of any preposition. Unfortunately no instances of a different preposition were found occurring with the same idiomatic meaning as *down the drain*, within 100 concordance lines from both COCA and Strathy. The only examples noticed were metaphorical usages where something is mentally, emotionally, or financially *draining.*

**Cost an Arm and a Leg**

Some idioms like *cost an arm and a leg* and *show one's true colours* show a surprising amount of creativity and allusion through several variants. About half the headlines utilized the idiom in its full form, demonstrated in Example (5.12a), although some include a person being referenced within the idiom; the majority of such cases utilizing the pronoun *you*, shown in (5.12b). Lexical variation is also observed with this idiom, seen in (5.12c). The remaining examples illustrate just how creative participants can be. A third of the headlines occurred with a partial form. This form could be the variant *an arm and a leg* (see Example 5.12d), or the plural version of this portion of the idiom (see Example 5.12e). This partial form might only contain the verb portion of the idiom *cost*, with some additional referent – *limb* in (5.12f) or *two of each* in (5.12g). Finally, a few headlines do not contain any part of the idiom, but simply allude to the idiom using other words, typically *limbs*, as in (5.12h). These examples not only show exceptional creativity, but they also show extreme variation – reference to the idiomatic meaning with a minimal form only or by using no actual

words from the idiom whatsoever.

(5.12) a. Night on the town costs an arm and a leg
b. Dinner in San Francisco can cost you an arm and a leg
c. Date Nights Going for an Arm and a Leg in San Francisco's Civic Center
d. San Francisco date night receipt: An arm and a leg
e. Forget losing your heart – dating takes your arms and legs
f. Trip to Civic Center May Cost A Limb
g. Thankfully you have two of each: Civic Center's costly experience
h. Is a night out worth the loss of limbs?

However, this 'extreme variation' (i.e. allusion to an idiom without using any formal portion of the idiom) is likely a task effect. Both the participants and the experimenter knew the idiom being referenced, as well as the context of the newspaper snippet. Therefore, accessing the idiom's meaning from the allusion is still possible despite using no formal portion of the idiom. If, for example, (5.12h) was a real headline in a newspaper, the idiomatic meaning would presumably be lost on the audience. In fact, a search through 200 concordance lines for *limb* and *limbs* in both COCA and Strathy yields zero instances where this word appears to be alluding to the idiom *cost an arm and a leg*. This further suggests that the results from this study are only possible when all parties involved have the same 'shared knowledge', making these allusions akin to inside jokes. This shared knowledge is what allows the idiomatic meaning to still be interpretable. Regardless of the limitations of this variant, it is impressive that speakers are able to be so exceptionally creative with idioms under the right conditions.

Nevertheless, Strathy and COCA do show examples of variation with this idiom, such as lexical variation (e.g. *pay an arm and a leg*), the insertion of a pronoun into the idiom (e.g. *cost me an arm and a leg*), and even variation of the conjunction (e.g. *cost an arm or a leg*), suggesting that the event participant has the option of selecting which to sacrifice. The most common type of variation evident in the corpora however is additional hyperbole, where an extra body part is incorporated into the idiom to emphasize the expensiveness of the item. The examples in (5.13) illustrate the range of body parts, and sometimes quantity, which can be included in this expression to convey exorbitant prices.

(5.13) a. A huge variety of statement necklaces and bold cocktail rings that anywhere else would cost you an arm, a leg, and a few toes.

b. Everything's going to be okay, though this bitch is going to cost an arm, a leg, and a torso.

c. I couldn't resist them, even though they cost an arm and a leg. Well, okay, an elbow and a knee.

d. The bad news is: It's going to cost you an arm and a leg, and maybe another arm.

e. In the end it usually ends up costing the artist an arm and a leg (and a foot).

f. They cost an arm and two legs, but Sam was in his "spare no expenses" mode.

## 5.4 Discussion

This study explored the creative and spontaneous productions of idiomatic variation. Five research questions were put forward in this study. The first question sought to determine the range of variation which can be produced by speakers of English, and as illustrated through both the quantitative and qualitative analyses, that range is in fact quite extensive. Variation observed in this study includes various syntactic inflections or alternations, such as changing the verb to progressive aspect (e.g. *calling the shots*), passives (e.g. *the law is taken out of Harding's hands*), and nominalizations (e.g. *just a shot in the breeze*), omitting a single word, typically a function word (e.g. *back to drawing board*), integrating additional concepts into the idiom (e.g. *gains medical ground*), and even partial forms alluding to the idiom (e.g. *media mountain more of a molehill*). Creative instances are also noticed, such as playing on the literal meaning of a word in the idiom (e.g. *NRA: shoot guns, not the breeze*), using no formal part of the idiom to allude to the idiomatic meaning (e.g. *large appetite and small mouth* [bite off more than you can chew]), and even blending or merging two distinct expressions (e.g. *crossing one bridge at a time* [cross that bridge when you come to it & one X at a time]).

The existence of formal idiom blends in the data is an exciting discovery. These variants have traditionally been regarded as a 'slip of the tongue' or an error in speech

production (cf. Fay, 1982; Cutting and Bock, 1997; Kuiper et al., 2007). Even Langlotz (2006), who specifically researched idiomatic creativity, classified idiom blends as "non-intentional erroneous variants". For Langlotz, formal idiom blends are a momentary deviation from the original, blending two synonymous or quasi-synonymous idioms; however, he does admit that sometimes erroneous variants can gain status and become lexicalized alternatives. However, the results from this study suggest that while idiom blends are not very frequent, they are not simply 'slips of the tongue', but can be employed to produce some of the most creative instances observed in the data.[2] For example, idiom blends not only merge two idioms together, but can also integrate contextual information into the blend, as in *uniforms as the new bandwagon* (blending *jump on the bandwagon* and *X is the new black*, but also incorporating *uniforms*, the topic of the newspaper snippet, into the blend) and *design your pie and eat it too* (blending *have your fingers in every pie* and *have your cake and eat it too*, while integrating the business name *Designer's Guild*, and its objective, from the snippet into the blend). According to Kemmer (2003), lexical blends merge both formal properties and concepts, and based on the blends elicited in this study, the same can be said about idiom blends. In fact, such instances can even be observed in actual newspaper headlines; consider *Kings of the Roost* from the Edmonton Metro, which merges the idioms *king of the hill* and *rule the roost* with the name of a band *Kings of Leon*, which is what the article is about.

Formal idiom blends can sometimes appear to be instances of both lexical variation and blending. For example, the headline *Billings put his head against the wall in search for Earhart mystery* could simply be changing the verb in the idiom *bangs his head against the wall* to the verb *put*. However, it may also be a blend with *put his head in his hands* – another idiom conveying frustration and keeping in line with the content of the snippet. This continuum between lexical variation and formal idiom blending may contribute to the reasoning behind why idiom blends were simply regarded as errors. True (i.e. noticeable) lexical variation was considered intentional, but any variant which potentially incorporated another idiom was seen as an error or a deviation from the 'correct' form. But perhaps researchers should start considering idiom blends on par with word blends – some instances may be

[2] A website called *The Idiomatic* (http://theidiomatic.com) exists where the sole purpose is to mix and match different expressions together randomly, producing instances like *A leopard shouldn't bite the hand that feeds you* from *A leopard doesn't change its spots* and *Bite the hand that feeds you*, as well as *Stupid is always greener on the other side* from *Stupid is as stupid does* and *The grass isn't always greener on the other side*.

unintentional mistakes while other instances are deliberate, conscious creations (cf. Algeo, 1977).

Additionally, some scholars have found consistency in morphological blends, indicating that the blending of lexical items is not completely random (cf. Kelly, 1998; Gries, 2004a,b). For instance, shorter words tend to go first in the blend, the second word usually contributes more to the blend, and there is often phonological similarity between the words allowing the words to overlap (e.g. *fantabulous* from *fanta*[stic] and *fabulous*). A quick glance at a few idiom blends in the headlines suggests that idiom blends may follow similar strategies as word blends; for example, *spills of the bag* utilizes 1 word and 5 letters from *spill the beans* (consisting of 3 words and 13 letters) and 3 words and 8 letters from *let the cat out of the bag* (consisting of 7 words and 20 letters). A similar pattern is observed with *the new bandwagon*, which utilizes 2 words and 6 letters from *X is the new black* (4 words/13 letters) and 1 word and 9 letters from *jump on the bandwagon* (4 words/18 letters). These examples illustrate that the shorter idiom goes first and more is contributed from the second idiom, at least in terms of number of letters. Obviously a more in-depth study is required to see if this pattern upholds, but it is interesting to note that a similarity between word blends and idiom blends can already be observed.

While the finding of formal idiom blends in this study is worthy of discussion, this particular type of variant is quite infrequent. But variant preferences are certainly evident, addressing the second research question. The most utilized forms of variation in the headlines are slight modifications to the idiom, such as the addition of a word to elaborate on the current context or the omission of a word characteristic of the telegraphic style of newspaper headlines, changing the syntax of the idiom to progressive aspect to indicate a continuous or ongoing event, as well as using a partial form of the idiom to allude to the idiomatic meaning. Two of these variants may be particular to the genre of newspaper headlines or perhaps even the task. First, the omission of one word from the idiom, typically a function word like *the*, might be overly represented in this study due to the nature of headlines, especially since they tend to be shorter and do not require full sentences. Second, partial forms alluding to the idiomatic meaning may also be overrepresented in the headlines. It is difficult to know for certain whether this particular variant is overrepresented in the headlines since searching for partial forms in corpora can be challenging, but for a few idioms at least it does appear to be the case. The remaining two variants, integrated concepts

and progressive aspect, are comparable to other research. Gibbs and Nayak (1989) found that both adjectival modifiers and the present progressive are preferred variants with both decomposable and nondecomposable idioms, while Schröder (2013) obtained the most instances for adverbial/adjectival modifier variants in her study on syntatic variation. Thus, it appears that the variants most utilized preserve as much of the idiom as possible while adapting it for specific contexts or effects.

The results obtained from this study are largely comparable to what is observed in corpora, answering the third research question. This is illustrated in the qualitative section of the results. Most idioms showed corresponding variants in both Strathy (Strathy Language Unit, 2013) and COCA (Davies, 2008). Only *cost an arm and a leg* differed between the variants elicited in the experiment and what was observed in corpora. More instances of allusion, especially using no formal element of the idiom, were elicited, but were not found in random samples of concordance lines from both corpora. This suggests that the results obtained here for *cost an arm and a leg* were largely a task effect – that in other contexts and situations, alluding to an idiom through other words and no formal element of the idiom itself would be missed. Instead, variants observed in corpora for this idiom expand on the hyperbolic nature of the idiom, including additional body parts to emphasize the expensiveness of the referent.

Comparing the results obtained with specific idioms demonstrates that different idioms are utilized with different patterns of variation. Some idioms are utilized with few variants (e.g. *call the shots*), whereas others show a range of possible variants (e.g. *cross that bridge when you come to it*). More importantly, idioms show patterns as to how they are varied. Speakers prefer to preserve the salient word *bandwagon* in *jump on the bandwagon*, if they reduce the form of this expression. Otherwise they modify the expression by integrating an additional concept into the idiom or by lexically varying the verb to one which better describes the manner in which the bandwagon was boarded. Compare this to *it's not over until the fat lady sings*, which has multiple salient words: *fat, lady*, and *sing*. This idiom is utilized with a variety of partial forms, playing on one or more of the available salient words, or alternatively, changing an element within the idiom to a new concept (e.g. *the proposal isn't over* or *until lawmakers sing*). In other words, this idiom does not typically incorporate an additional concept, but rather replaces an existing one, which may possibly be due to its length, as this idiom is much longer than most. These ob-

servations demonstrate the extent to which idioms are utilized differently, especially those which have different syntactic structures. A number of syntactic types were included in this study, some of which occurred with few instances, making it difficult to determine quantitatively if specific variants are produced more with certain syntactic types.

Despite the differences between the idioms, all idioms were predominantly used with the canonical form in the headlines (i.e. the form that was provided to the participants). The participants were told to be as creative as they wanted, yet they still produced the canonical form in the majority of instances (approximately 60–65%). This usage of the canonical form coincides with what is observed in corpora (cf. Moon, 1998; Schröder, 2013), as well as converges with the findings from comprehension studies (cf. McGlone et al., 1994), which show that the canonical form is processed faster, presumably because it is utilized more, making it more entrenched and familiar. This preference for the canonical form also converges with the results from Chapter 3, where the canonical form was preferred in an acceptability ratings task, and Chapter 4, in which the canonical form shows a slight advantage in processing in an eye-tracking study.

Even though speakers utilize idioms in their canonical form in the majority of instances, very creative occurrences do emerge. The fourth research question explores how creative speakers can be with idioms when prompted and the results from this study shows that speakers can be exceptionally creative. Formal idiom blends, a form which has traditionally been considered a 'slip of the tongue', and allusions, which were perhaps exaggerated due to the task, show extreme creativity that is quite unique to this study. However, more subtle forms of creativity are also observed in the headlines, such as playing on the literal meaning of one of the words in the idiom (e.g. *Dermatology student risks skin by keeping his nose to the grindstone*). The compositionality of idioms has been largely debated in the idiom literature and the fact that speakers are able to play on the literal meaning of a word while utilizing the idiom's idiomatic meaning indicates that idioms do have some degree of compositionality, which can be accessed and manipulated for imaginative and humorous effects.

The ability to access the literal meaning of the words along with the idiomatic meaning of the whole has also been shown in priming studies (cf. Sprenger, 2003; Sprenger et al., 2006), where a related word (e.g. *foot*) to one of the words in an

idiom (e.g. *lend a hand*) has been shown to facilitate the processing and production of the idiom as a whole. Moreover, the idiom (e.g. *lend a hand*) can also facilitate the processing and production of this related word (e.g. *foot*). This interaction between the literal and figurative meanings has led some scholars to suggest a co-existence of the meanings, such as an "abstract scene comprising the idiomatic meaning is conceived against the conceptual background of a literal scene" (Langlotz, 2006, p.287). The literal meanings of the words overlap with the meanings they acquire in figurative expressions, similar to perceptual stimuli such as a vase and two profiles (Cacciari, 1993). This relationship between the literal and figurative meanings is what allows idioms to be used in various communicative and discourse purposes, in addition to being manipulated for pragmatic and contextual effects (cf. Cacciari and Glucksberg, 1991), as quite a few speakers illustrated in this study.

The last research question sought to determine which factors, if any, influence idiomatic variation. This study found that length is a significant predictor of idiomatic variation – longer idioms tend to be varied more than shorter idioms. This finding is likely due to the fact that more words belong to longer idioms; therefore, more options are available to vary. Also, speakers produced more variation depending on the version of the experiment they saw. If they were shown an example headline that contained either an allusion to an idiom or a formal idiom blend, they produced more variants. However, if they were shown an example headline which contained lexical variation, they produced less variation. Speakers tend to be more creative with idioms if they are shown a creative example; thus, creativity promotes further creativity.

However, if the speaker is more accepting of `LQ7` (e.g. *irregardless*), they are less likely to modify the idiom. This sentence includes a blended form of two words which have similar semantics (i.e. *regardless* and *irrespective*), causing the meaning of the blend to be largely unchanged while the form appears less compositional. But the form now contains two morphemes which essentially mean 'not' or 'without', which might make the meaning clearer to the participants who rated it as being more acceptable. These same participants then tend not to alter idioms because they prefer increased clarity. Lastly, speakers who know the idiom are less likely to modify it if the idiom is more frequent. The idiom is perhaps more entrenched for these speakers and therefore they prefer not to alter it.

This study complements the work of corpus-based research (Moon, 1998; Bar-

low, 2000; Duffley, 2013; Schröder, 2013), which has explored variation with both select idioms, as well as idioms in general. The results converge with similar types of variants being frequently utilized. However, participants in this study also produced instances which are difficult to search for in a corpus, such as extreme types of allusion (i.e. where no formal part of the idiom is used) and instances of partial forms or lexical variants with an idiom that contains a very common noun (e.g. *down the drain*). One would have to know which possible lexical items can be used to refer to the idiom (e.g. *limbs* for *cost an arm and a leg*), or search for this common word with additional lexical items in a context window (e.g. any preposition with *drain*), with no guarantee that an idiom variant will be found. Thus, complementing corpus-based research with an elicitation task provides new insight into how creative and flexible speakers can be when utilizing idioms, while still maintaining the idiom's meaning.

# CHAPTER 6

# Discussion

## 6.1 Discussion of Results

This dissertation investigated idiomatic variation through several different methods. Idiomatic variation has largely been explored through corpus-based studies (cf. Moon, 1998; Barlow, 2000; Langlotz, 2006; Wulff, 2008; Duffley, 2013; Schröder, 2013), while few have examined variation from an experimental perspective (cf. Gibbs and Nayak, 1989; Gibbs et al., 1989a; McGlone et al., 1994). This study has attempted to fill that gap by exploring variation using three different methodologies. A multiple methods approach can be quite revealing, shedding light on converging, as well as diverging, results (cf. Gries et al., 2005; Arppe and Järvikivi, 2007; Wulff, 2009). Different research questions and different methods employed can lead to findings which either converge or diverge. Research discussing the use of multiple methodologies has largely focused on achieving converging results, or results which appear to tell the same story. However, Arppe and Järvikivi (2007) discuss the usefulness of considering diverging results as well; for example, they claim that 'rareness' in corpus data does not necessarily mean 'unacceptable', but that participants can rate forms as acceptable even though they might be infrequent in a naturalistic sample of speech. Diverging results then are not simply contradictory findings, but require additional interpretation to understand the larger picture, especially when utilizing different modalities. The findings from this study reveal both converging and diverging results, between the experiments utilized in this study as well as with previous findings in the literature, and will be discussed systematically. The discussion will then shift towards the larger picture of idioms and idiomaticity.

### 6.1.1 The Variants

Integrated concepts were one type of variant included in this study. This variant was largely the addition of another word in the idiom (e.g. *pull the political strings*), but could have included other possible modifications in the elicitation task. Converging results are observed for this variant between the acceptability ratings study, elicitation task, and corpus results. This variant is more preferred, generally showing higher acceptability ratings than the other variants, especially by speakers who utilize idioms more often. These higher ratings for inserted concepts are consistent with the higher similarity ratings for the adjective and adverb insertion categories of Gibbs and Nayak (1989) and Tabossi et al. (2008). In addition, this variant is more likely to be produced by participants in an elicitation task, showing just over 20% when coded for semantic variation; however, this category does include other instances of conceptual integration, such as the replacement of a word with one that adds extra information (e.g. *Full Serge Ahead*). Finally, this variant is seen quite frequently with the seven idioms searched in corpora, with some idioms showing considerable use with integrated concepts (e.g. *jump on the Twitter bandwagon*). This frequent occurrence in corpora was also observed by Schröder (2013) who found that almost all idioms in her study occurred with adverbial/adjectival modification, and some idioms to a rather high degree (e.g. *make headway*). The results from the eye-tracking study however show that this variant does come with a cost – it takes longer to understand this variant than the canonical form. The results primarily show a main effect – it takes longer to interpret the variant, but it is not processed differently than the canonical form. This longer time is due to the variant being longer in length, contributing additional information to the idiomatic meaning.

Another frequently produced variant is a partial form of the idiom (e.g. *listen to the fat lady* or *through grapevined information*). This type of alternation to the form of the idiom occurred around 13% in the elicitation task when coded for formal variation, and is commonly produced with idioms which have a salient word, or salient words, in the expression, such as *jump on the bandwagon, hear things through the grapevine* and *it's not over until the fat lady sings*. Corpora searches with these expressions also reveal a variety of partial forms (e.g. *the bandwagon momentum* and *I think the fat lady just sang*), which allude to the idiom.

However, partial forms are not used to the same extent with all idioms; for

instance, *call the shots* and *mend fences* do not show examples, in the headlines or in corpora, with partial forms. This perhaps partly explains the diverging results observed in the ratings and eye-tracking studies, where all idioms were presented with a partial form variant. Partial forms were generally rated as less acceptable than the other variants, although more acceptability was observed if the idiom is longer in length (i.e. more words of the original form are still included in the variant). This variant also showed considerably longer fixation durations on the manipulated word (e.g. [*have*] *the strings*) and on the surprising word (e.g. *have the* [*strings*]) when the beginning of the idiom was altered. In order for the sentence to remain grammatical when the beginning of the idiom was altered, semantically vague verbs were used instead of the original verb in the expression (i.e. *pull*). However, these semantically vague verbs caused longer fixations on the manipulated word itself (i.e. *have*), as well as produced 'spillover effects' onto the following 'surprising' word (i.e. the first content word after the manipulated word). Idioms that had the ending altered (e.g. *spill it* [spill the beans]) did not produce significantly longer fixations on the manipulated word compared with the canonical form. These semantically vague verbs then further explain the diverging results between the elicited and corpus-collected data (i.e. naturalistic data) and the controlled stimuli created specifically for the experiments.

Interestingly, the elicitation task revealed that formal idiom blends are a possible type of idiomatic variation that speakers can employ when opting to be creative with idioms. This particular variant occurs quite infrequently, potentially indicating why scholars have traditionally regarded this variant as an error, or a slip of the tongue (cf. Fay, 1982; Cutting and Bock, 1997; Kuiper et al., 2007). But speakers are able to employ blends creativity, often incorporating additional information along with the merged idioms (e.g. *Intentional spills of the bag* blends the idioms *spill the beans* and *let the cat out of the bag*, while integrating that the secret was intentionally revealed). Formal idiom blends were therefore included in the acceptability rating and eye-tracking studies in an attempt to understand more about this variant. Speakers tend to rate idiom blends with lower acceptability, especially if they use idioms more often. Idiom blends also receive lower ratings of acceptability if the two idioms share similar semantics, and even lower ratings if the participants were not familiar with one of the expressions in the blend (i.e. the merged portion was more noticeable, and unacceptable, to these participants).

While speakers might find idiom blends with similar semantics less acceptable, semantic and even syntactic similarity does not affect the processing of idiom blends. Knowledge of the blending idiom however was predictive. Shorter fixations were observed on the manipulated word (i.e. the portion of the blending idiom included in the variant), by speakers who did not know that idiom, when the manipulated word occurred at the beginning of the idiom. That is to say, they did not recognize the other idiom included in the blend. But all participants showed longer fixations on the surprising word (i.e. the first content word after the alternation). These longer fixations were reduced however if the participants were more flexible with and accepting of non-standard or 'incorrect' usage. Despite these longer fixations on the manipulated and surprising words, summed fixation durations on the idiom as a whole do not appear significantly different from the canonical form. This suggests that even though some words may come across as surprising or unusual within a particular context, blends in general do not impair comprehension. The findings for formal idiom blends reveal that we tend to show a dispreference for using idiom blends, illustrated by their lower acceptability in the ratings task and their low frequency of occurrence in the elicitation task, but that we can in fact utilize them creatively and understand them with little difficulty. This variant is not simply a slip of the tongue, but perhaps should be thought of as similar to word blends, freely moving along a continuum between intentional and unintentional.

The last variant explored in this study was lexical variation (e.g. *reveal one's true colours*). This variant was found to be less acceptable for those who used idioms regularly, but showed almost identical ratings of acceptability as the canonical form for those who seldom use idioms. These results converge with those from the eye-tracking study, which found that this variant was processed similarly to the canonical form, showing some spillover onto the surprising word when the beginning of the idiom was altered (e.g. *tug the strings*), but generally not fixated on significantly longer than the canonical form. This variant was also seldom produced in the elicitation task, although more instances were observed when a slight modification to the form co-occurred, such as the omission of *the* (e.g. *Cape Breton facing brunt of storm* [bear the brunt of something]). In fact, even the version of the experiment which included an example of lexical variation (i.e. *put whole boot in his mouth*) was predictive of less variation produced by the participants. Perhaps the subtlety of this variant causes people to produce it less and find it less acceptable.

However, this variant was largely based off of Gibbs et al.'s (1989a) study where they defined lexical variation (or lexical flexibility) as the replacement of one of the elements within the idiom with a synonymous word. If one considers the replacement of a lexical item as an example of semantic productivity (cf. McGlone et al., 1994), then this variant is produced more often than it initially appears. Instances of integrated concepts (e.g. *Full Serge Ahead*) and even formal idiom blends (e.g. *Billings puts his head against the wall* [bang one's head against the wall]) could have been considered examples of semantically productive lexical variation. In fact, the examples of verb changes observed in corpora for *jump on the bandwagon* does in fact convey a great deal of semantic productivity, from *join, leap, climb, scramble*, etc. Perhaps then, instances of lexical variation used in the acceptability ratings study received lower ratings by those who utilize idioms more often because these synonymous forms do not contribute additional information to the variant, making them less common and less preferred. Whereas instances of semantically productive lexical variation are more utilized, favoured, and possibly seen as purposefully creative idiomatic variation.

The literal meaning of the idiom was also explored, to determine how it differed from the idiomatic meaning of the idiom. Interestingly, the literal meaning of the idiom (i.e. used in a different context than the idiomatic meaning, but in the canonical form) was one of the variants rated with the lowest acceptability, especially if the participants indicated that they utilize idioms regularly. In other words, if speakers are more familiar with its idiomatic meaning, they do not find its literal use very acceptable. It is interesting to note however, that longer idioms tend to be more accepted with a literal reading, possibly because there is more information available in the longer expression to support a literal interpretation (cf. *the last/final straw* with the original *it's the last straw that breaks the camel's back*; from Moon, 1998).

Idioms are rarely used with their literal meaning, as evident from the elicitation task. With contexts and idioms that allow a more literal reading, several patterns were noticed, such as the use of the word *literally* in the headline (e.g. *Beware of skating on this ice, literally*), repetition of the idiom to convey an idiomatic and a literal reading separately (e.g. *Thrillseekers are skating on thin ice if they choose to skate on thin ice*), and even a altered form of the expression to better convey a literal reading (e.g. *Skaters warned of thin ice*). However, while the use of the literal

interpretation is not preferred, playing on the literal reading is certainly a variation strategy that some speakers enjoy, as evident in these examples: *Ontario nurses are chomping at the bit and keep being put out to pasture*, *Chinese let the cat out of the bag before it's even in it*, and *Gingrich better like the taste of crow – he'll be eating it the rest of his career*. These headlines nicely illustrate Langlotz's (2006) 'topic related literal-scene manipulation' strategy, where both the literal and figurative scenes are modified for effect. But while the use of the literal meaning of the idiom is less preferred, and infrequently produced, it is still able to be interpreted, showing similar fixation times as the canonical form.

Finally, this study has shown that idioms, which have traditionally been regarded as a homogeneous class of fixed phrasal forms, actually occur with different distributions and patterns of variation. This was largely illustrated in the headlines for the seven idioms from the elicitation task, and the corresponding corpus results. Some idioms were utilized with very little variation and deviation from the canonical form (e.g. *call the shots*), whereas others showed much more variation and creativity in their usage (e.g. *jump on the bandwagon*). The type of variation and creativity even varied between the idioms: *down the drain* maintained its form but was creatively employed with manner-rich verbs, while *it's not over until the fat lady sings* was often reduced to a partial form, alluding to the idiomatic meaning, and containing one or all of *fat, lady*, and *sing*. Distributional patterns were also evident, such as *mend fences* occurring to a considerable degree with a *to*-infinitive (e.g. *attempt to mend fences*). This variability was also evident in the random effects structure for the acceptability ratings and the eye-tracking study, where by-Item random slopes for condition were significant. These slopes represent the different preferences for each idiom in each condition. But while idioms can be produced with variation, they tend to be utilized in their canonical form the majority of the time. The canonical form was used approximately 60–65% in the elicitation task and was preferred in the acceptability ratings task. This preference for the canonical form presumably facilitates the learning of idioms and leads to their faster recognition.

### 6.1.2 Predictor Variables

Previous literature on idioms has placed a great deal of emphasis on the compositionality of the individual words within the idiomatic expression. However, the most

frequent method of investigating the role of compositionality is through a classification task of semantic decomposability, where participants are asked to categorize idioms for the amount of information each component part contributes to the meaning of the whole (cf. Gibbs and Nayak, 1989; Titone and Connine, 1994b; Tabossi et al., 2008). Idioms whose consitutents contribute directly to the meaning of the whole are 'normally decomposable idioms' (e.g. *pop the question*), whereas idioms whose constituents contribute indirectly, or metaphorically, are 'abnormally decomposable idioms' (e.g. *carry a torch*). Meanwhile, idioms whose constituents do not contribute to the meaning are 'nondecomposable idioms' (e.g. *kick the bucket*). However, studies have shown that speakers do not reliably group idioms into these semantic decomposability categories and in fact perform at chance (Titone and Connine, 1994b; Tabossi et al., 2008). Furthermore, results are inconsistent with this measure. Gibbs and Nayak (1989) found that normally decomposable idioms were rated as more semantically similar to their literal paraphrases, than both abnormally decomposable and nondecomposable idioms, when they were syntactically altered, but Tabossi et al. (2008) found that semantic similarity ratings did not differ between normally decomposable and nondecomposable idioms when syntactically modified. Due to these inconsistent and unreliable results, a different strategy for accounting for the compositionality of idioms was employed in this study.

One such measure was the Latent Semantic Analysis (LSA) scores, obtained from the English Lexicon Project (Balota et al., 2007).[1] LSA is a measure of similarity, used here between the words of the idiom and either its definition or a literal paraphrase. This analysis uses a pairwise comparison of two strings of text (i.e. the idiom and its definition or the idiom and its paraphrase), which compares the local context of these two strings to obtain a measure of similarity (Landauer et al., 1998). When the LSA score is higher, the texts share more similar contexts (e.g. *stop something in its tracks* 'stop something' = 0.87) and the words of the idiom contribute more to the meaning of the expression, but if the LSA score is smaller (e.g. *cut the mustard* 'be acceptable' = 0.07) fewer contexts are shared and the individual words contribute much less to the meaning of the expression.

The LSA measure, which is an objective measure of the compositionality of the expression, is predictive of fixation duration on the idiom as a whole. Longer fixations were observed on idioms that had higher LSA scores between the paraphrase and the

[1] http://lsa.colorado.edu/

idiom. In other words, longer reading times were observed for idioms which are more semantically similar and occur in similar contexts as their paraphrase. These idioms require additional time to determine the usage and interpretation since the words of the idiom are not distinctive for the context. The LSA score between the idiom and its paraphrase is predictive on the First Fixation Duration of the manipulated word, showing shorter fixations for idioms with higher LSA scores, suggesting that if the manipulated word is more similar to how the idiom is used (i.e. its paraphrase), then the word is easier to process. However, this initial advantage does not persist and is no longer predictive in the Total Fixation Duration. The fact that the LSA scores are primarily predictive of the whole idiom, and are not predictive on the altered words, is a finding similar to Titone and Libben (2014) who only found an effect of decomposability on the target in a priming study 1000ms after the presentation of the idiom. The fact that these measures are showing up on the whole idiom or after the presentation of the idiom may suggest that these are actually measures of analyzability, not compositionality. In addition, Titone and Libben (2014) found that nondecomposable idioms were more primed than decomposable idioms, a similar result to what was discovered here – idioms whose constituents are more semantically related to the meaning of the whole are processed slower as the similar contexts make the idiomatic meaning less distinctive.

Interestingly, the LSA measures were not predictive of the amount of idiomatic variation produced in the elicitation task. Cutting and Bock (1997) also did not find an effect of decomposability in their study where they attempted to induce the production of idiom blends. These results further indicate that the compositionality of the idiom does not affect the degree of variation produced, either intentional or unintentional. This is perhaps due to the fact that all idioms can be varied and are not dependent on the relationship between the literal and figurative meanings (cf. Duffley, 2013). The LSA measures were also not a significant predictor of the acceptability of variation in the ratings task. Instead, the mean transparency ratings were predictive. A measure of transparency was collected in Chapter 2, which asked participants to rate the transparency of each idiom included in this study. These ratings were collected for inclusion in all analyses on idiomatic variation, in an attempt to determine whether it was the compositionality or overall transparency of the expression which contributed to the understanding and usage of idioms and idiomatic variation. This measure was only found to be predictive for the acceptability

ratings – the more transparent an idiom was felt to be, the more acceptable speakers found the expression. But transparency was not predictive of variation (i.e. idioms were not found to be more acceptable with variation if they were more transparent). This finding is also in line with the notion that all idioms can be varied, as long as the overall meaning is interpretable, and is not dependent on the compositionality, or transparency, of the expression.

Interestingly, an idiom was rated as more transparent if the LSA score between the idiom and its definition was higher. In other words, idioms which share similar contexts with its definition are regarded as more transparent. But in order to understand an idiom, and an idiom variant, contexts which are uniquely distinctive allow the idiom to be understood faster. The transparency of the expression is not facilitative for processing, instead the contexts which allow the idiomatic meaning to be more predictable, despite the actual words utilized in the expressions or the type of variation (i.e. the LSA scores were not significant in an interaction with condition), are facilitative. In addition, the mean variation ratings were also predictive of the comprehension of idiomatic variation. The mean rating for each idiom in each condition (e.g. *spill the beans* with lexical variation) was highly significant, such that idioms which were rated as more acceptable were fixated on significantly less. This was true for the idiom as a whole, as well as the variants.

The length of the idiom was found to be predictive for idiomatic variation; a variable which is surprisingly not often discussed in the literature on idioms (cf. Fanari et al., 2010). Longer idioms are rated as being more transparent, especially if the speaker enjoys using idioms. Longer idioms were also noted previously to be more acceptable with a literal interpretation – perhaps these two findings are linked, such that longer expressions provide more information about the relationship between the literal and the figurative meanings, making a literal reading more plausible and an idiomatic reading more transparent. But longer idioms occurring in their canonical form in an idiomatic context tend to be rated as less acceptable than shorter idioms. Perhaps once the idiom is learned and the appropriate connections are made, a shorter reduced version is preferred, making a literal reading less likely. This may explain why idioms get shortened or truncated over time (e.g. *hair of the dog* [*that bit you*] or [*he who pays the piper*] *calls the tune*), as well as why idioms become more opaque to newer generations who have not learned the full expression. This longer, more transparent form may also indicate why more variation is observed with longer

idioms in the elicitation task – there is still plenty of context available in the lengthy expression to convey the meaning or to assist in integrating additional meaning (cf. integrated concepts have fixation durations more comparable to the canonical form when they are longer).

The frequency of the expression is also predictive of variation, particularly for speakers who are familiar with the idiom. Speakers are less likely to vary an expression, more likely to rate the expression as transparent, and are quicker at rating the idiom (i.e. both transparency and acceptability) when the frequency of the idiom is higher. Thus, the idiom becomes more entrenched with a particular form, making it easier to recognize and interpret (cf. Schweigert, 1986). Interestingly, frequency is not predictive of acceptability ratings for idiomatic variation. Speakers are less likely to vary an idiom when it is more frequent, but listeners do not necessarily mind if they do. But the processing of this alternation will come at a cost; longer fixations and a greater number of fixations are observed on the manipulated or surprising word when the idiom has higher co-occurrence frequencies. Thus, when a sequence of words which typically occur together, and therefore display effects of frequency, has been modified, additional time is required to interpret the new sequence as the advantage it typically receives due to predictability is no longer available.

The idiom variants included in this study had either the beginning (i.e. the verb, such as *tug the strings*) or the ending (i.e. the noun, as in *spill the peas*) modified to determine whether a difference in placement affected either the comprehension or acceptability of the variants. Gibbs et al. (1989a) modified either the noun or the verb in their study and found no difference between the modifications. But both were rated as less semantically similar to a literal paraphrase than the canonical form, and more similar than a variant which altered both the verb and the noun (e.g. *punt the pail* [kick the bucket]). Following their example, this study also manipulated the noun or the verb to determine if similar findings are observed. No difference in acceptability was observed – speakers do not have a preference for whether the beginning or the ending of the idiom is altered. However, longer fixations are observed on the idiom as a whole if the beginning was altered. This is likely due to the fact that longer fixations are observed on the manipulated word, both when it occurred at the beginning (verb) or the ending (noun) of the idiom, as well as on the surprising word (i.e. the final word when the beginning was altered). These findings show a different result from what Gibbs et al. (1989a) found. In an

attempt to determine any differences in comprehension for variants, they utilized a ratings task, comparing the variant to a literal paraphrase of the idiomatic meaning. They likely collected participants' preferences for the variants rather than the degree of processing difficulty. When speakers are asked to rate the acceptability (or similarity) of the variants, variants are rated as less acceptable than the canonical form, but speakers have no preference between modifications to the verb or the noun. When a speaker understands these variants however, a processing cost is observed when modifications are made to the beginning of the expression; thus, comprehension takes a little longer.

The findings in this study seem to initially contradict those of Wulff (2008), who found that modifications to the verb were more common than to the noun. However, she only included morpho-syntactic variants (e.g. tense, voice) or lexico-grammatical variants, which added an element to the idiom. She did not include alternations which modified the actual lexical items of the idiom, such as lexical variation or formal idiom blends. Thus, modifications to the verb, such as tense, aspect, mood, or person and number agreement on the verb, are the most frequent types of modification to the verb. But when the changes are greater, such as the replacement of an actual lexical item, then understanding the variant results in a slower processing time than alternations made to the noun. These results are likely due to the time-dependent nature of idiom processing (Titone and Libben, 2014). As one advances through the idiom, the predictability of the idiom becomes greater as the idiomatic meaning slowly accumulates, resulting in greater priming effects on later words (e.g. final word) or post-idiom offset. Therefore, if meaning slowly accrues as we advance through the idiom, then it seems reasonable that changes made later in the expression will come with less of a cost – the meaning is more predictable and one has an easier time interpreting this meaning even if alternations are present. This time-dependent nature may also explain why integrated concepts show longer fixation times on concepts integrated later in the phrase (cf. *kept his overwhelmed nose to the grindstone* vs. *burn a hole in his careless pocket*). Since these concepts are expanding and elaborating upon the context, they integrate new information which is not predictable, which then seems more surprising later in the phrase. When presented earlier, the accumulation of the idiomatic meaning is not as great and therefore is easier to integrate.

This study has been unique in including several participant-related variables.

First, knowledge of the idiom has been shown to be a very important predictor. Previous studies in the idiom literature have typically used a measure of familiarity (cf. Schweigert, 1986; Schraw et al., 1988; Titone and Connine, 1994b; Nippold and Taylor, 2002), which is either used as a control variable separating high vs. low familiarity idioms or as a mean rating of subjective familiarity collected by a different set of participants. This study however, collected a binary response (i.e. 'yes' or 'no') from every participant as to whether they know each expression. Since idioms were being altered in this study, it was important to include a variable which indicated whether the participant was actually familiar with each idiom, and not just a general measure of familiarity. The results from this variable are consistent with the literature, such that known idioms are fixated on less (i.e. are understood faster) than unknown or unfamiliar idioms. In addition, idioms are rated as more transparent and more acceptable in their canonical form when they are known, as well as responded to quicker than unknown idioms. Participants who did not know the idiom found no difference in acceptability between the variants and the canonical form (i.e. no effect of condition was evident); that is, they did not notice variation present in the idiom. Lastly, participants who knew the idiom were more likely to vary the idiom in the elicitation task, but only if the idiom had a lower frequency.

Familiarity with the idiom was not the only participant-related predictor variable – variables specific to the speakers' personality were also found to be predictive of idioms and idiomatic variation. For example, flexibility with language was found to be a very important predictor. Participants rated the acceptability of seven prescriptively incorrect sentences, which were either used directly in the model or indirectly via a Principal Component. The ratings for a specific sentence were most often predictive. LQ7 (i.e. *irregardless*) was predictive for the transparency ratings (those participants who rated LQ7 as more acceptable showed higher ratings of transparency for unknown idioms) and the elicitation task (those participants who rated LQ7 as more acceptable produced less variation). LQ3 (i.e. *by Susan and I*) was predictive for the acceptability ratings, showing greater acceptability of variation the more the participant was accepting of LQ3. LQ5 (i.e. *could have went*) was also predictive of the transparency ratings – the more participants accepted this variant, the less transparent they rated unknown idioms. And finally, PC2 from a Principal Components Analysis, which had the highest loadings for LQ7 and LQ6 (i.e. *could care less*), was most significant in predicting fixation duration – shorter fixations

were observed on the whole idiom and on altered words the more the participant was accepting of 'incorrect' language usage in general.

These results begin to make sense if one considers the type of error in these sentences and the response variable that it is predicting. LQ7, which is a word blend combining the words *regardless* and *irrespective*, becomes a wordform that is less compositional. But this blend now contains two morphemes (i.e. *ir-* and *-less*) which convey negativity, both meaning 'not, without, lack of'. The presence of these two morphemes then strengthens the intended meaning of the word, perhaps making its meaning more obvious. Therefore, if participants find this usage more acceptable, they are more likely to find idioms more transparent, as they are less concerned with the compositionality of the form and place more importance on the interpretability and intended meaning. In addition, they are less likely to modify the form of the idiom in order to maintain this clarity.

Meanwhile, LQ3 and LQ5 utilize alternative words to convey the same meaning. LQ3 utilizes a different noun form (i.e. *I* instead of *me*), whereas LQ5 uses a different verb form (i.e. *have went* instead of *have gone*). Speakers who are more accepting of LQ3 are also more accepting of idiomatic variants, both of which use different forms to convey a similar meaning. For a similar reason, higher acceptability of LQ5 is predictive of lower transparency ratings. This sentence uses a different form of the verb to indicate the past participle. This 'incorrect' usage does not add clarity to the intended meaning. Acceptability of LQ5 then, which is acceptance of a variant form to convey a similar meaning, is predictive of lower transparency ratings because altering the form leads to the contruction becoming less obvious. However, why is LQ3 predictive of the acceptability ratings while LQ5 is predictive of transparency? This result may simply reflect the preferences of the different speakers who participated in each study – certain participants may have stronger opinions about one form over another. But it may also reflect latent information about the noun and verb. The distributional properties of the verb have been proposed to contribute more to idiomaticity (Wulff, 2008), which may explain why the sentence with the alternate verb form was predictive for transparency. Meanwhile, alternations to nouns are easier to understand, due to the predictability of the idiomatic meaning as one progresses through the idiom (Titone and Libben, 2014); therefore, variant forms of any noun may be more predictive of acceptability, reflecting this easier interpretability.

Finally, one's overall flexibility with language – PC2.LQ, which is a variable combining the acceptability ratings for all seven of these prescriptively incorrect sentences – is predictive of fixation duration. The comprehension of language is ultimately about the successful interpretation of an utterance, despite the potential for a variety of errors which could be present in the language signal. Therefore, reading time is shorter for those participants who are more flexible with language in general, irregardless of the form.

Interestingly, whether the participant enjoys using idioms appears to be predictive of idiomatic transparency ratings. Speakers who enjoy using these types of expressions view them as more transparent, especially if they are not familiar with the expression. They also take longer to rate unknown idioms, spending extra time considering the meaning and interpretability. Thus, if speakers enjoy using idioms, they seem more willing to analyze them and establish a relationship between the words and the meaning, in turn making these expressions clearer and more obvious in meaning. In addition, how often the participant utilizes idioms is predictive of the acceptability ratings of idiomatic variation. All variants are viewed as less acceptable if the speaker uses idioms often, but the variants are rated more similar to the canonical form if the participant seldom uses idioms. This finding suggests that those participants who utilize idioms frequently develop more entrenched and established relationships between the constituent words of the canonical form and the idiomatic meaning. They are then less accepting of altered forms. These two variables however are not predictive of the comprehension or production of idiomatic variation; it does not matter if one uses idioms often or enjoys using idioms, we all understand idioms and idiom variants equally well and can produce variants at a comparable rate when explicitly encouraged to be creative.

## 6.2 Idioms & Idiomatic Variation: An Implicit Grammar Approach

### 6.2.1 Traditional Approaches

The current widely-accepted view of idioms is that they are stored and accessed whole, activated when the language user reads or hears the expression. One account

for activating these idioms is by accessing the 'idiom key', or the idiomatic configuration indicating the listener has received sufficient input (Cacciari and Tabossi, 1988). But as illustrated in this study, idioms are utilized with a range of possible variants, from lexical variation to inserted concepts to partial forms. How would this separately stored representation be activated when one encounters an idiomatic variant? For example, when encountering a variant such as *leap on the bandwagon*, where the lexical alternative *leap* is used instead of the more canonical form *jump*, how is the configuration for this idiom reached if one of the elements of the idiom, especially in the beginning of the idiom, is different? Furthermore, what if the variant is only a partial form, *the fat lady is warming up*, instead of the much longer version: *it's not over until the fat lady sings* – how does one know that it is the *fat lady* in the operatic world who is about to sing, alluding to the idiom in reference to the conclusion of the event, and not a specific person?

In some models, like the Superlemma Theory, these variants are explicitly not allowed (Sprenger et al., 2006, p.177). This would mean then that each variant must have its own superlemma, linking the components with their respective words. But what about the idiom's meaning – would these new variant superlemmas be linked to the same conceptual representation or would they have their own representation? This would presumably differ depending on the type of variant. For variants adding personal or stylistic flavour, such as Duffley's (2013) poetic or extravagant alternation (e.g. *she had died, flipped the bucket*), the conceptual representation would be the same. But for variants which are semantically productive, adding novel information into the expression (e.g. *shatter the ice*), this representation would have to be different in order to account for the semantic modification. Constructionist approaches might better account for this variation. In Construction Grammar for example, partially filled idioms like *pull someone's leg* have their own construction, and another construction (e.g. *her, his, Bob's, the teacher's*) could fill that open slot (i.e. someone). Maybe more idioms share this partially-filled representation (e.g. *make X headway*). Or, perhaps all idioms are represented in a schematic representation (cf. Moon, 1998; Taylor, 2012), where all possible forms are listed along with their frequencies in order to know what variants are most probable, and which variants can fill the schema and utilize its associated metaphors. These associated variants would include all possible variants, even partial forms.

Numerous studies have found effects of frequency in language; effects for individual words, idioms, collocations of varying lengths or 'n-grams' like *I don't know why*, and even lexical bundles such as *in the middle of* (cf. Bybee, 2006; Arnon and Snider, 2010; Shaoul and Westbury, 2011; Tremblay et al., 2011). These effects are in line with the notion of storing frequency information along with the variant in the idiom schema. However, concerns have been raised about the plausibility of storing such information (cf. Baayen et al., 2013). For example, in true exemplar-based theories, every token of a word or n-gram would be stored; this storage would logically begin with the first instance otherwise how else would one know when they have reached a certain threshold? But storing every token of every possible n-gram would lead to a combinatorial explosion. Baayen et al. (2013, p.335) present an illustrative example of this combinatorial explosion: after collecting all 4-grams available from CELEX (Baayen et al., 1995), using the 47 final words included in Arnon and Snider's (2010) study of the frequency effects of four-word phrases, they generated "337,069 different phrases covering 7,494 distinct meanings". This suggests that storing every n-gram or construction, along with their respective frequencies, would lead to longer processing times (not faster ones) as speakers would have to search through a vast amount of stored information to find the specific word/n-gram/construction.

Furthermore, Baayen (2010) found that n-gram probability (i.e. sentential and morphological context) was a more robust predictor of lexical decision latencies than repetition frequency counts. This n-gram probability has been observed in priming studies as well, where the combination of words leads to certain predictions or expectations (see Elman, 2011, for a review). For example, subject-verb combinations lead to specific predictions about upcoming objects, such that *The lumberjack cuts...* primes *wood* whereas *The surgeon cuts...* primes *bone*. Predictions are also observed for the type of semantic theme that follows, based on voice in a sentence fragment (e.g. *She arrested* primes *crook*, but *She was arrested* primes *cop*). Similar expectations (or predictability) are also observed with idioms; idioms which have a higher cloze probability have an idiomatic meaning that is available earlier, such as on the penultimate word (cf. Cacciari and Tabossi, 1988; Titone and Connine, 1994a; Fanari et al., 2010). It is this aspect of predictability that lead Tabossi and Zardon (1993, p.156) to indicate that the recognition of an idiom (i.e. accessing the idiom key) is based on probability, or "the point in the string after which the probability of the fragment to continue idiomatically is very high".

But literal language processing does not stop after activation of the idiom key, as Cacciari and Tabossi (1988) initially proposed. Studies have shown that the literal meaning of words within the idiom are still active (cf. Cutting and Bock, 1997; Sprenger et al., 2006), and can play a role in idiom variation (cf. Cacciari and Glucksberg, 1991; McGlone et al., 1994; Langlotz, 2006). This led researchers to propose that the literal meaning of an idiom's constituents are accessed and used where applicable, either contributing to the idiom's interpretation or for use in discourse. Models such as the Idiom Decomposition Hypothesis and the Hybrid Model incorporate the processing of the literal meaning, which is facilitative for decomposable idioms whose constituents contribute to the meaning of the whole (Gibbs et al., 1989b; Titone and Connine, 1999). Since the constituents of nondecomposable idioms do not contribute meaning to the whole, activation of their literal meaning leads to an inhibitory effect, accounting for why nondecomposable idioms take longer to process. But the literal meaning of nondecomposable idioms can contribute to the meaning, especially when the literal scene of the idiom has been manipulated in a way that adds contextual relevance to the idiomatic meaning, such as *All we know with certainty is that Titian died in 1576... He may have been over, or under, 90 years old when he kicked the paint can* (Duffley, 2013). This example comments on the fact that both the painter passed away and that his career is officially over.

McGlone et al. (1994) attempted to account for variation by stating that the idiom can either be retrieved directly or generated using ordinary linguistic processing. They suggest that speakers employ different strategies during the processing of variants in order to understand them and that these strategies may differ depending on the variant. For example, if encountering a variant with an added quantifier, speakers may access the idiom directly and modify the quantity implied by the expression. Whereas if a variant containing a different lexical item is encountered, speakers may access the meaning of the constituent words and interpret the structural relations between the meanings. McGlone et al. also indicate that the constituents can develop polysemous meanings from their use in the idiom for use in contexts outside the idiom, such as *spill* being used as 'reveal' in contexts other than *spill the beans* (see also Geeraerts, 1995). Therefore, if an idiom is used frequently and develops polysemous senses, then the variant can be generated quicker. But if the canonical form and the variants can be processed incrementally using ordinary linguistic processing, then the idiom need not be stored. One advantage for storing

the idiom whole would be for faster processing of the idiomatic meaning when it occurs in its canonical form, as this meaning can be accessed directly as opposed to generated (McGlone et al., 1994). However, given the previous findings in regards to n-gram probability and the predictability of idioms (cf. Tabossi and Zardon, 1993; Titone and Connine, 1994a; Baayen, 2010; Shaoul et al., 2014), perhaps idioms are simply processed like the rest of language (i.e. ordinary linguistic processing), but are interpreted more quickly if they are more predictable. Studies have shown that the syntax of idioms is not stored separately but uses the same principles as literal language (Konopka and Bock, 2009; Tabossi et al., 2009b). Therefore, if idioms can be processed like literal language, and processed more quickly when their n-gram probabilities are higher, then storing the idiom separately seems superfluous. Perhaps then, idioms can be better understood within a different framework, one which takes n-gram probabilities and predictability into account.

### 6.2.2 Implicit Grammar

The effects of predictability and frequency have led scholars to consider other options for storing and processing language, specifically the availability of lexical knowledge and information without a designated 'lexicon' (cf. Elman, 2004, 2011; Baayen et al., 2011; Baayen and Ramscar, 2015). Listeners have much more information available to them than just the meaning of a word. They are able to draw upon past experiences, cultural norms, event and world knowledge, and even feelings of the speaker to interpret the meaning being communicated. In these recent approaches to language, words no longer possess meaning, but instead are 'cues to meaning', modulated by experience and context. An example of this was already illustrated – specific predictions are able to be made about upcoming elements in an utterance, such as the subject-verb combination *The lumberjack cuts* leading to expectations of the object *wood* (Elman, 2011). This information would not be available at the verb, if language was understood in a truly compositional way. In fact, Baayen (2015) illustrates this lack of compositionality using the verb *finished*. In each of the three sentences, the verb has a different meaning: *The author finished the book* means 'writing'; *The student finished the book* means 'reading', while *The goat finished the book* implies 'eating'. If the verb actually contributed meaning in a compositional linear manner, the meaning of the verb *finished* should be the same. A related point has been made by Gibbs (1995) and Gibbs and Colston (2012) about the impossibility to identify

what is meant by 'literal' language. Gibbs and Colston (2012, p.50) provide an example of a 'literal' question *Have you eaten?* If this sentence was understood in a truly literal way, the interpretation would be something like 'have you eaten ever', but when used in context it means 'have you eaten today'. But pragmatics and real-world knowledge can change this meaning even further to refer to a specific meal, such as 'have you eaten dinner' if uttered say around seven in the evening. These examples illustrate that words do not have a one-to-one mapping between meaning and form in the traditional sense, but are instead used as cues to meaning.

One recent approach is Implicit Grammar (see Baayen and Ramscar, 2015, for details), in which language is regarded as a signal meant to reduce the listener's uncertainty by conveying cues that are discriminative for a particular outcome. For example, Libben and Titone (2008) found that infrequent verbs lead to higher predictability of idiomatic sequences, these specific verbs then would be discriminative for the idiomatic meaning. This approach is based on learning theory, where certain behaviours or patterns are discriminative from other potential cues for predicting specific events. For instance, Pavlov (1927) trained dogs to salivate at the ring of a bell by presenting the ringing bell alongside the presentation of food. But the simple existence of an association is not enough for learning. If a light was flashed at the same time as the bell was rung, the light would not become a predictor for the presentation of food, because the bell is an already reliable predictor of the food (a phenomenon known as 'blocking'; see Baayen and Ramscar, 2015). Unless the light on its own is learned to predict the presentation of food, this cue will remain an uninformative cue for the food. Implicit Grammar is based on this 'co-learning'. Learning occurs when cues (e.g. bells, lights) successfully predict outcomes (e.g. food), but also when predictions fail to result in those outcomes.

Learning theory has been applied to language in Implicit Grammar through the Naive Discriminative Learner (NDL; Baayen et al., 2011, 2013; Baayen and Ramscar, 2015). The NDL uses a two-layer network, where cues in the form of letter unigrams and bigrams (or trigrams) are the input cues in the model, and are used to predict outcomes, or meaning (i.e. lexical, grammatical, pragmatic). This model utilizes letter unigrams and bigrams (as opposed to morphemes or words) due to evidence of sublexical information. Phonaesthemes (e.g. *gl* in *glitter, glisten, gleam, glow* 'light' or *sn* in *snout, snore, sneeze, snot* 'mouth, nose') have been shown to produce priming effects similar to morphemes (Bergen, 2004), suggesting that meaning may

reside at a sublexical level. The NDL utilizes the Rescorla-Wagner equations for learning (Wagner and Rescorla, 1972), which adjust the weights associated with each cue-outcome pair. When a cue is present and successfully predicts an outcome, the weights are increased. But if a cue is present and the outcome is absent, then the weights are decreased. The weights are not adjusted for a cue that is not present (see Baayen et al., 2015a, for details). Since multiple cues may be present at once, cues then compete for outcomes. Similarly, cues can be used to predict multiple outcomes, therefore outcomes also compete for cues. All cues are linked to all outcomes, but only the weights for active cues (to all outcomes) are adjusted. The sum of the weights for all cues predicting an outcome are referred to as the activation weights. High activation of an outcome reflects a high degree of confidence that the cues which discriminate this particular outcome are present in the external world (Baayen et al., 2015a, pg.6). In addition, frequency effects can be accounted for in this model through the activation weights and n-gram probabilities of cues, without stored representations being posited to explain these effects (cf. Baayen et al., 2013).

Learning a language then is about learning which cues (i.e. sounds, morphemes, words, contexts) are informative, or discriminative, for a particular outcome (i.e. meaning). For example, the letter bigram *q-a* for the scrabble word *qaid* 'tribal chieftain' is quite distinctive for the word, unlike the letter bigram *i-d*, which is found in numerous other words, such as *hid, quid,* and *said*. Therefore, *q-a* appears with a higher weight and is more discriminative in predicting the word *qaid* (Baayen and Ramscar, 2015). In this approach, "morphology and syntax are implicit in the distribution of cues and outcomes, which jointly shape a network that is continuously updated with usage" and therefore this approach is referred to as Implicit Grammar (Baayen and Ramscar, 2015, p.111).

The outcomes in Implicit Grammar are referred to as 'lexomes', which are "a theoretical construct at the interface of language and a world that is in constant flux with the flow of experience" (Baayen et al., 2015a, p.5). That is, lexomes are dimensions in a system of knowledge which speakers acquire and continually modify and can be invoked through language or real-world experiences. Following De Saussure (1966), lexomes are like chess pieces in that they depend where on the board they are located and where they are in relation to the other chess pieces. Similarly, the value of a lexome depends on the other lexomes encoded in the signal and on the other lexomes and experiences available in one's system of knowledge. Finally, these

lexomes can be invoked by any systematic lexicalized contrast, including grammatical and pragmatic information.

The signal can still discriminate idiomatic knowledge even if the cues are different (i.e. idiomatic variation). The value of the lexome for a variant will be similar to the original (although subject to updating from the experience) as long as the other lexomes encoded in the signal and in one's system of knowledge discriminate it to be similar. Part of discriminative learning is to discern which cues in the language signal are encoding relevant information and experiences and which ones are not (Baayen and Ramscar, 2015, p.111). So when an additional concept is integrated into the idiom, for example, these relevant cues discriminate the corresponding lexomic experiences. But if the variant is an unintentional slip of the tongue, the cues will likely not be discriminative and therefore deemed irrelevant, even if only discriminated as irrelevant by the listener. In addition, the signal can discriminate overlapping lexomes or multiple dimensions within one's system of knowledge. This is likely what happens when using idioms as more information is conveyed with an idiom than a literal paraphrase (cf. Gibbs, 1993, 2007). Not only does one discriminate lexomic information about 'revealing a secret' with the signal *spill the beans*, but other lexomic experiences as well, such that this secret was revealed unintentionally as opposed to confidentially disclosing it.

For the purposes of illustration, the Rescorla-Wagner equations in the `ndl` package (Arppe et al., 2014) in `R` (R Core Team, 2014) were run on the idiom variants and literal phrases of *mend fences* in Example 6.1. The variants are examples from the elicited headlines in Chapter 5, while the literal phrases are made up to contrast a different lexome. In this demonstration, all idiomatic variants are cues for the lexome 'make peace' while all literal phrases are cues for the lexome 'repair fence'. In order to illustrate higher activations (i.e. higher weights) for a particular lexome, *fences* was only used with the lexome 'make peace', while *fence* was used for the lexome 'repair fence'. All cues were fed into the model an equal number of times. The activation weights from the NDL model are shown in Table 6.1.

(6.1) a. *mend fences* 'make peace'
b. *mend damaged fences* 'make peace'
c. *restoring fences* 'make peace'
d. *fences mended* 'make peace'

e. *fix the fence* 'repair fence'

f. *mending a fence* 'repair fence'

g. *the fence is mended* 'repair fence'

As anticipated, the cues *fences* and *fence* show high weights for the lexomes 'make peace' (0.61) and 'repair fence' (0.69) respectively. However, other cues show high activations for the lexomes as well. *Mend* and *restoring* show higher weights for the lexome 'make peace', while *the* and *is* show higher activations for 'repair fence'. These results of course make sense. Both *is* and *the* only occur with 'repair fence' in this data sample, so they intuitively should be discriminative for the lexome 'repair fence'. Meanwhile, *mend* only occurs in its bare form with 'make peace' – *mended* and *mending* are the cues that occur with 'repair fence'.

Table 6.1: Activation Weights for 'Mend Fences'

| | MakePeace | RepairFence |
|---|---|---|
| a | -0.12 | 0.13 |
| damaged | -0.09 | -0.07 |
| fence | 0.18 | 0.69 |
| fences | 0.61 | -0.16 |
| fix | 0.11 | 0.19 |
| is | 0.18 | 0.37 |
| mend | 0.48 | 0.23 |
| mended | -0.09 | -0.20 |
| mending | -0.12 | 0.13 |
| restoring | 0.41 | 0.18 |
| the | 0.30 | 0.56 |

*Restoring* is discriminative for 'make peace' as well, but this cue is particularly noteworthy when compared with the weight for *damaged*. Both of these cues are alternations of the original idiom *mend fences* and have been discussed throughout this dissertation using various methodologies. The results here, in this very simplistic model, already seem to converge with previous results. Integrated concepts (e.g. *mend damaged fences*) were found to be produced more in the elicitation task and were the most preferred variant in the acceptability ratings task. The weights on *damaged* are quite small, likely due to the fact that both *mend* and *fences*, which are discriminative for the lexome, are both present in the signal. This is an example of blocking, where *damaged* does not become discriminative because other cues in the signal are successful at discriminating the lexome. This is probably why this

type of variant is preferred and utilized more – it does not result in considerable relearning as the other cues in the signal are still able to discriminate the meaning. However, lexical variation (e.g. *restoring fences*) replaces one of the elements in the canonical form with a synonymous word. This results in greater co-learning, as the cues used to discriminate the lexome have changed which causes the weights for *mend* to decrease and the weights for *restoring* to increase. This greater co-learning may explain why lexical variation is less preferred and produced less frequently, and why speakers prefer to use the canonical form even when encouraged to be as creative as they would like. This greater co-learning may also account for how some variants can "gain the status of usual variants" or become an institutionalized formal variant (Langlotz, 2006, p.201); that is, if they are encountered enough as cues to discriminate a lexome, they will continue to be used as cues for that lexome.

While this model is far from the complete picture of understanding idioms and idiomatic variation, it does illustrate two important points. First, experiential and lexical knowledge can be accessed through contrastive learning of distributional patterns of sublexical cues. Thus, distributional properties of idioms and their contexts can serve as discriminative cues of their meaning. Second, varying these cues comes with a cost – the variation can make formerly relevant cues less relevant (cf. Baayen et al., 2015b). This likely explains why there is a preference to use the canonical forms of idioms, reserving variants for elaboration or personal flare. It may also explain how truncation of idioms happen, certain cues are more discriminative of the lexome allowing them to continue to be utilized in the signal while the irrelevant ones begin to be left out. Adjustments to the weights will also be observed if societal and cultural changes lead to the loss of experiental knowledge for motivating idioms (e.g. *kick the bucket*). While new interpretations can be created to explain the idiom, the contexts in which it occurs will presumably start to diminish resulting in a restrictive context for its use. This may explain why more opaque idioms are rather infrequent in corpora and appear with more limited syntactic flexibility (Wulff, 2008; Schröder, 2013).

## 6.3 What is Idiomaticity?

Several attempts have been made to define idiomaticity. Wulff (2008) proposed that idiomaticity is the distributional behaviour of idioms which characterizes the idioms

and places them along the idiom-collocation continuum. This continuum increases in distributional rigidity, where those expressions at the extreme end of the continuum are the 'pure' idioms and are less likely to occur with, for example, changes in mood or voice. However, if we assume an Implicit Grammar approach, then the distributional properties which discriminate each expression would be accounted for in the activation weights between cues and outcomes. Furthermore, Wulff (2008) assumes idioms are a type of construction, in which the morpho-syntactic, semantic, and pragmatic information associated with the idiom is stored (Fillmore et al., 1988). But as discussed previously, storing all this information along with each construction would lead to a combinatorial explosion. But if it is not the distributional properties, what is so unique about these expressions that makes them different from the rest of language? We know that second language learners have difficulty with these expressions, they can take longer to understand them and can get the form incorrect (cf. Schraw et al., 1988; Schmitt and Underwood, 2004), but the ultimate goal of speaking a language fluently is to 'speak idiomatically' (Taylor, 2012). In other words, there is something special about these expressions that makes it difficult for non-native speakers to learn.

If the distributional patterns and behaviours are accounted for through the association weights (i.e. reflecting usage patterns and n-gram probabilities), then idiomaticity must be semantic rather than distibutional. Nunberg et al. (1994) also proposed that idiomaticity was semantic in nature, but according to them the syntactic distribution of idioms was determined by the semantics. While it is certainly plausible to creatively utilize an idiom based on the semantic motivation, this does not always explain distributional patterns. For example, verbs can show surprisingly uneven distributions with certain inflections (e.g. *mean, suppose*, and *thank* occur over 90% in their base form in casual conversation; see Newman, 2008). These verbs can certainly be motivated to occur with other inflections, but they rarely are. These inflectional islands then are also experientially motivated (Rice and Newman, 2005).

In addition, idiomaticity is not simply constructional polysemy, or polysemy at the phrasal level. Research on idioms has largely focused on the fact that idioms can have two possible meanings: a 'literal meaning' or the meaning associated with each individual word that makes up the sequence, and an 'idiomatic meaning', or the meaning associated with the whole string. Idioms are created by employing non-conventional metaphors and metonymies, which become conventionalized and

accepted into the language (Langlotz, 2006). This figurative relationship is still interpretable, allowing these idioms to be seen as analyzable and transparent. But these expressions can evolve with time; they may become truncated in form, the figurative relationship may no longer be accessible especially if society and culture has changed, or maybe the expression has become used in new ways rendering the original ties unrecognizable. These opaque idioms, which no longer have a recognizable connection, could be akin to homonyms (i.e. words that have the same form but unrelated meanings, such as *bank* 'financial institution' and *bank* 'river's edge'). Research on lexical ambiguity has found that words with polysemous, or related meanings, are understood quicker than homonyms, which have unrelated meanings (cf. Klepousniotou, 2002; Klepousniotou and Baum, 2007). The related senses are facilitative in accessing the appropriate meaning, and probably are facilitative for idioms as well – transparent idioms are facilitated by their related meanings, while opaque idioms are not.

However, not all idioms have a plausible literal meaning. Titone and Connine (1994b) collected literality ratings for 171 idioms and found that some idioms (e.g. *foot the bill, be on cloud nine, pay through the nose*) have extremely low literality ratings (i.e. they have a very low probability that they would be used with their literal meaning). In fact, studies have found that activation of the idiomatic meaning is greater for those idioms which have low literality ratings (cf. Titone and Connine, 1994a; Titone and Libben, 2014). Secondly, those idioms which do have a plausible literal meaning tend to use different distributional patterns to convey that meaning. Moon (1998, p.183) presents syntactic structures, along with their respective frequencies, for the literal and idiomatic uses of *break the ice*. More instances were observed with the idiomatic meaning, 201 tokens, whereas only 65 instances were observed with the literal meaning. The figurative sense predominantly occurs as *break the ice* (152) or *the ice is broken* (31), whereas the literal meaning had a more varied distribution, such as with an adverb particle (e.g. *break the ice with*), without a determiner (e.g. *break ice* or *break ice into*), or with the *ice* as a subject in the active voice (e.g. *the ice breaks* or *the ice breaks apart*). Thus, idioms with plausible literal interpretations show different distributional patterns that distinguish their usage (i.e. different cues discriminate the respective lexomes).

What makes idioms so unique then, and different from more 'literal' language, is the potentially vast amount of information conveyed by the idiom. Idioms do not

simply mean their literal paraphrase (Nayak and Gibbs, 1990; Gibbs, 1993, 2007), but are used to convey culturally-relevant metaphors and metonymies, pragmatic and situational information, experiential information and cultural norms, folk etymologies and visual scenarios, etc. A considerable amount of information is communicated through a minimal signal. Consider this comment: *Well folks! I made it 14 days of pretty constant travel with my parents and it finally happened. I blew a gasket.*[2] Upon reaching the idiom in this particular context, we understand a great deal about what happened. We know she was bottling up her emotions (for 14 days to be exact), until the pressure was too unbareable and there was a sudden release of this pressure. We know this release was unintentional, and that the trigger of this release was not the cause of the anger. We know from our own experiences that she was probably yelling and said things she now regrets. And we also, depending on our own experiences, can emotionally relate as to why this incident even happened. Thus, the signal *I blew a gasket* conveys a substantial amount of information, much more than simply *I got angry*. Idiomaticity then is a richness of information. It is the ability to communicate maximal information with a minimal signal.

[2]I thank Pattra Winner for allowing me to use her Facebook status as an example.

# References


Aitken, R. C. B. (1969). A growing edge of measurement of feelings. *Proceedings of the Royal Society of Medicine*, 62:989–993.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, AC-19(6):716–723.

Akaike, H. (1985). Prediction and entropy. In Atkinson, A. C. and Fienberg, S. E., editors, *A celebration of statistics: The ISI centenary volume*, pages 1–24. Springer, New York.

Algeo, J. (1977). Blends, a structural and systematic view. *American Speech*, 52:47–64.

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

Arppe, A., Hendrix, P., Milin, P., Baayen, R. H., and Shaoul, C. (2014). ndl: Naive Discriminative Learning. R package version 0.2.16.

Arppe, A. and Järvikivi, J. (2007). Every method counts – combining corpus-based and experimental evidence in the study of synonymy. *Corpus Linguistics and Linguistic Theory*, 3(2):131–160.

Ayto, J., editor (2009). *From the horse's mouth: Oxford dictionary of English idioms*. Oxford University Press, Oxford.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press, Cambridge.

Baayen, R. H. (2010). Demythologizing the word frequency effect: A discriminative learning perspective. *The Mental Lexicon*, 5(3):436–461.

Baayen, R. H. (2015). Implicit grammar. Workshop slides at the Spring Training in Experimental Psycholinguistics (STEP), University of Alberta.

Baayen, R. H., Hendrix, P., and Ramscar, M. (2013). Sidestepping the combinatorial explosion: An explanation of n-gram frequency effects based on naive discriminative learning. *Language and Speech*, 56(3):329–347.

Baayen, R. H., Milin, P., Filipovic-Durdevic, D., Hendrix, P., and Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review*, 118:438–481.

Baayen, R. H., Piepenbrock, R., and Gulikers, L. (1995). The CELEX lexical database (CD-ROM). University of Pennsylvania: Linguistic Data Consortium.

Baayen, R. H. and Ramscar, M. (2015). Abstraction, storage and naive discriminative learning. In Dabrowska, E. and Divjak, D., editors, *Handbook of cognitive linguistics*, pages 100–120. De Gruyter Mouton, Berlin.

Baayen, R. H., Shaoul, C., Willits, J., and Ramscar, M. (2015a). Comprehension without segmentation: A proof of concept with naive discriminative learning. *Language, Cognition, and Neuroscience*. In Press.

Baayen, R. H., Tomaschek, F., Gahl, S., and Ramscar, M. (2015b). The Ecclesiastes principle in language change. In Hundt, M., Mollin, S., and Pfenninger, S., editors, *The changing English language: Psycholinguistic perspectives*. Cambridge University Press, Cambridge, UK. In Press.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., and Treiman, R. (2007). English lexicon project. *Behavior Research Methods*, 39(3):445–459.

Barlow, M. (2000). Usage, blends and grammar. In Barlow, M. and Kemmer, S., editors, *Usage-based models of language*, pages 315–345. CSLI Publications, Stanford, CA.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2014). lme4: Linear mixed-effects models using eigen and S4.

Bates, E., Dale, P. S., and Thal, D. (1995). Individual differences and their implications for theories of language development. In Fletcher, P. and MacWhinney, B., editors, *Handbook of child language*. Basil Blackwell, Oxford.

Bergen, B. K. (2004). The psychological reality of phonaesthemes. *Language*, 80(2):290–311.

Bobrow, S. and Bell, S. (1973). On catching on to idiomatic expressions. *Memory and Cognition*, 1:343–346.

Bowdle, B. F. and Gentner, D. (2005). The career of metaphor. *Psychological Review*, 112(1):193–216.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, 82(4).

Cacciari, C. (1993). The place of idioms in a literal and metaphorical world. In Cacciari, C. and Tabossi, P., editors, *Idioms: Processing, structure, and interpretation*, pages 27–55. Lawrence Erlbaum, New Jersey.

Cacciari, C. and Glucksberg, S. (1991). Understanding idiomatic expressions: The contribution of word meanings. In Simpson, G. B., editor, *Understanding word and sentence*, pages 217–240. Elsevier, Amsterdam, NL.

Cacciari, C. and Tabossi, P. (1988). The comprehension of idioms. *Journal of Memory and Language*, 27:668–683.

Cutting, J. C. and 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.

Dabrowska, E. (2014). Words that go together: Measuring individual differences in native speakers' knowledge of collocations. *Mental Lexicon*, 9(3):401–418.

Davies, M. (2008). The corpus of contemporary American English: 450 million words, 1990-present. Available online at http://corpus.byu.edu/coca/.

de Boer, A. G. E. M., van Lanschot, J. J. B., Stalmeier, P. F. M., van Sandick, J. W., Hulscher, J. B. F., de Haes, J. C. J. M., and Sprangers, M. A. G. (2004). Is a single-item visual analogue scale as valid, reliable and responsive as multi-item scales in measuring quality of life? *Quality of Life Research*, 13:311–320.

De Saussure, F. (1966). *Course in general linguistics*. McGraw, New York.

Duffley, P. J. (2013). How creativity strains conventionality in the use of idiomatic expressions. In Borkent, M., Dancygier, B., and Hinnell, J., editors, *Language and the creative mind*, pages 49–61. CSLI Publications, Stanford, CA.

Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Sciences*, 8(7):301–306.

Elman, J. L. (2011). Lexical knowledge without a lexicon. *The Mental Lexicon*, 6(1):1–33.

Fanari, R., Cacciari, C., and Tabossi, P. (2010). The role of idiom length and context in spoken idiom comprehension. *European Journal of Cognitive Psychology*, 22(3):321–334.

Fay, D. (1982). Substitutions and splices: A study of sentence blends. In Cutler, A., editor, *Slips of the tongue and language production*, pages 163–195. Mouton de Gruyter, Amsterdam.

Fillmore, C. J., Kay, P., and O'Connor, M. C. (1988). Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language*, 64(3):501–538.

Freyd, M. (1923). The graphic rating scale. *The Journal of Educational Psychology*, 14:83–102.

Funke, F. and Reips, U.-D. (2012). Why semantic differentials in web-based research should be made from visual analogue scales and not from 5-point scales. *Field Methods*, 24(3):310–327.

Geeraerts, D. (1995). Specialization and reinterpretation in idioms. In Everaert, M., van der Linden, E.-J., Schenk, A., and Schreuder, R., editors, *Idioms: Structural and psychological perspectives*, pages 57–73. Lawrence Erlbaum, New Jersey.

Geeraerts, D. (2009). Prisms and blends: 'Digging one's grave' from two perspectives. In Oleksy, W. and Stalmaszczyk, P., editors, *Cognitive approaches to language and linguistic data: Studies in honor of Barbara Lewandowska-Tomaszczyck*, pages 87–104. Peter Lang, Frankfurt am Main.

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

Gibbs, R. W. (1986). Skating on thin ice: Literal meaning and understanding idioms in conversation. *Discourse Processes*, 9(1):17–30.

Gibbs, R. W. (1993). Why idioms are not dead metaphors. In Cacciari, C. and Tabossi, P., editors, *Idioms: Processing, structure, and interpretation*, pages 57–77. Lawrence Erlbaum, New Jersey.

Gibbs, R. W. (1994). *The poetics of mind: Figurative thought, language, and understanding*. Cambridge University Press.

Gibbs, R. W. (1995). Idiomaticity and human cognition. In Everaert, M., van der Linden, E.-J., Schenk, A., and Schreuder, R., editors, *Idioms: Structural and psychological perspectives*, pages 97–116. Lawrence Erlbaum, New Jersey.

Gibbs, R. W. (2007). Idioms and formulaic language. In Geeraerts, D. and Cuyckens, H., editors, *The Oxford handbook of cognitive linguistics*, pages 697–725. Oxford University Press, Oxford.

Gibbs, R. W. and Colston, H. L. (2012). *Interpreting figurative meaning*. Cambridge University Press, Cambridge/New York.

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

Gibbs, R. W. and Nayak, N. P. (1989). Psycholinguistic studies on the syntactic behavior of idioms. *Cognitive Psychology*, 21:100–138.

Gibbs, R. W., Nayak, N. P., Bolton, J. L., and Keppel, M. E. (1989a). Speakers' assumptions about the lexical flexibility of idioms. *Memory & Cognition*, 17(1):58–68.

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

Gift, A. G. (1989). Visual analogue scales: Measurement of subjective phenomena. *Nursing Research*, 38(5):286–288.

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

Gries, S. T. (2004a). Isn't that fantabulous? How similarity motivates intentional morphological blends in English. In Achard, M. and Kemmer, S., editors, *Language, culture, and mind*, pages 415–428. CSLI Publications, Stanford, CA.

Gries, S. T. (2004b). Shouldn't it be breakfunch? A quantitative analysis of blend structure in English. *Linguistics*, 42(3):639–667.

Gries, S. T., Hampe, B., and Schönefeld, D. (2005). Converging evidence: Bringing together experimental and corpus data on the association of verbs and constructions. *Cognitive Linguistics*, 16(4):635–676.

Hamblin, J. L. and Gibbs, R. W. (1999). Why you can't kick the bucket as you slowly die: Verbs in idiom comprehension. *Journal of Psycholinguistic Research*, 28(1):25–39.

Hayes, M. H. S. and Patterson, D. G. (1921). Experimental development of the graphic rating scale. *Psychology Bulletin*, 18:98–99.

Hirotani, M., Frazier, L., and Rayner, K. (2006). Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language*, 54:425–443.

Horn, G. M. (2003). Idioms, metaphors and syntactic mobility. *Journal of Linguistics*, 39(2):245–273.

Hothorn, T., Hornik, K., and Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics Statistics*, 15(3):651–674.

Kelly, M. H. (1998). To "brunch" or to "brench": Some aspects of blend structure. *Linguistics*, 36(3):579–590.

Kemmer, S. (2003). Schemas and lexical blends. In Cuyckens, H., Berg, T., Dirven, R., and Panther, K.-U., editors, *Motivation in language: Studies in honor of Günter Radden*, pages 69–97. John Benjamins, Amsterdam/Philadelphia.

Keysar, B. and Bly, B. (1995). Intuitions of the transparency of idioms: Can one keep a secret by spilling the beans? *Journal of Memory and Language*, 34:89–109.

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

Klepousniotou, E. (2002). The processing of lexical ambiguity: Homonymy and polysemy in the mental lexicon. *Brain and Language*, 81:205–223.

Klepousniotou, E. and Baum, S. R. (2007). Disambiguating the ambiguity advantage effect in word recognition: An advantage for polysemous but not homonymous words. *Journal of Neurolinguistics*, 20:1–24.

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

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

Kuperman, V. and Van Dyke, J. A. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language*, 65:42–73.

Landauer, T. K., Foltz, P. W., and Laham, D. (1998). Introduction to latent semantic analysis. *Discourse Processes*, 25:259–184.

Langlotz, A. (2006). *Idiomatic creativity: A cognitive-linguistic model of idiom-representation and idiom-variation in English*. John Benjamins.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. University of Chicago Press, Chicago and London.

Libben, M. R. and Titone, D. A. (2008). The multidetermined nature of idiom processing. *Memory & Cognition*, 36(6):1103–1121.

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

Miller, G. A. and Charles, W. G. (1991). Contextual correlates of semantic similarity. *Language and Cognitive Processes*, 6(1):1–28.

Moon, R. (1998). *Fixed expressions and idioms in English*. Oxford University Press, Oxford.

Mostafa, M. (2010). Variation in the V+the+N idioms: The modification of idioms expresses the sponaneity of language in use. *English Today*, 26(4):37–43.

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

Newman, J. (2008). Aiming low in linguistics: Low-level generalizations in corpus-based research. In *Proceedings of the 11th international symposium on Chinese languages and linguistics*, Hsinchu, Taiwan. National Chiao Tung University.

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

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

Nunberg, G., Sag, I. A., and Wasow, T. (1994). Idioms. *Language*, 70(3):491–538.

Oxford English Dictionary (2015). *OED Online*. Oxford University Press, available at www.oed.com.

Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex (translated by G. V. Anrep)*. Oxford University Press, London.

Pawley, A. (1985). Lexicalization. In Tannen, D. and Alatis, J. E., editors, *GURT '85: Languages and linguistics. The interdependence of theory, data, and application.*, pages 98–120, Washington, DC. Georgetown University Press.

Popiel, S. J. and McRae, K. (1988). The figurative and literal senses of idioms, or all idioms are not used equally. *Journal of Psycholinguistic Research*, 17(6):475–487.

R Core Team (2014). R: A language and environment for statistical computing.

Ramscar, M. and Baayen, R. H. (2013). Production, comprehension, and synthesis: A communicative perspective on language. *Frontiers in Psychology*.

Rayner, K., Kambe, G., and Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology*, 53A(4):1061–1080.

Rice, S. and Newman, J. (2005). Inflectional islands. In *9th International Cognitive Linguistics Conference*, Seoul, South Korea.

Schmitt, N. and Underwood, G. (2004). Exploring the processing of formulaic sequences through a self-paced reading task. In Schmitt, N., editor, *Formulaic sequences: Acquisition, processing, and use*, pages 173–189. John Benjamins, Amsterdam/Philadelphia.

Schraw, G., Trathen, W., Reynolds, R. E., and Lapan, R. T. (1988). Preferences for idioms: Restrictions due to lexicalization and familiarity. *Journal of Psycholinguistic Research*, 17(5):413–424.

Schröder, D. (2013). *The syntactic flexibility of idioms: A corpus-based approach.* AVM, Munich, DE.

Schweigert, W. A. (1986). The comprehension of familiar and less familiar idioms. *Journal of Psycholinguistic Research*, 15(1):33–45.

Shaoul, C., Baayen, R. H., and Westbury, C. F. (2014). N-gram probability effects in a cloze task. *The Mental Lexicon*, 9(3):437–472.

Shaoul, C. and Westbury, C. (2011). Formulaic sequences: Do they exist and do they matter? *The Mental Lexicon*, 6(1):171–196.

Sinclair, J., editor (2011). *Collins COBUILD idioms dictionary.* Harper Collins.

Skoufaki, S. (2009). Investigating the source of idiom transparency intuitions. *Metaphor and Symbol*, 24:20–41.

Sprenger, S. (2003). *Fixed expressions and the production of idioms.* Radboud University Nijmegen, MPI Series in Psycholinguistics, vol. 21. Nijmegen, NL.

Sprenger, S. A., Levelt, W. J. M., and Kempen, G. (2006). Lexical access during the production of idiomatic phrases. *Journal of Memory and Language*, 54:161–184.

Strathy Language Unit (2013). Strathy corpus of Canadian English: 50 million words, 1970-2010. Available online at http://corpus2.byu.edu/can/.

Swinney, D. A. and Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behaviour*, 18:523–534.

Tabossi, P., Fanari, R., and Wolf, K. (2008). Processing idiomatic expressions: Effects of semantic compositionality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(2):313–327.

Tabossi, P., Fanari, R., and Wolf, K. (2009a). Why are idioms recognized fast? *Memory & Cognition*, 37(4):529–540.

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

Tabossi, P. and Zardon, F. (1993). The activation of idiomatic meaning in spoken language comprehension. In Cacciari, C. and Tabossi, P., editors, *Idioms: Processing, structure, and interpretation*, pages 145–162. Lawrence Erlbaum, New Jersey.

Tabossi, P. and Zardon, F. (1995). The activation of idiomatic meaning. In Everaert, M., van der Linden, E.-J., Schenk, A., and Schreuder, R., editors, *Idioms: Structural and psychological perspectives*, pages 274–282. Lawrence Erlbaum, New Jersey.

Taylor, J. R. (2012). *Mental corpus: How language is represented in the mind.* Oxford University Press, Oxford/New York.

Titone, D. A. and Connine, C. M. (1994a). Comprehension of idiomatic expressions: Effects of predictability and literality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(5):1126–1138.

Titone, D. A. and Connine, C. M. (1994b). Descriptive norms for 171 idiomatic expressions: Familiarity, compositionality, predictability, and literality. *Metaphor and Symbolic Activity*, 9(4):247–270.

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

Titone, D. A. and Libben, M. R. (2014). Time-dependent effects of decomposability, familiarity and literal plausibility on idiom meaning activation: A cross-modal priming investigation. *The Mental Lexicon*, 9(3):473–496.

Tremblay, A., Derwing, B., Libben, G., and Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning*, 61(2):569–613.

Urdang, L. (1988). *The whole ball of wax and other colloquial phrases.* Perigee Books, New York.

Wagner, A. and Rescorla, R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In Black, A. H. and Prokasy, W. F., editors, *Classical conditioning II*, pages 64–99. Appleton-Century-Crofts, New York.

Wray, A. (2002). *Formulaic sequences and the lexicon.* Cambridge University Press.

Wulff, S. (2008). *Rethinking idiomaticity: A usage-based approach.* Continuum, London/New York.

Wulff, S. (2009). Converging evidence from corpus and experimental data to capture idiomaticity. *Corpus Linguistics and Linguistic Theory*, 5(1):131–159.

# APPENDIX A

# Linear Mixed-Effects Models for the Transparency Ratings of Idioms

## A.1 Transparency Ratings

Table A.1: Fixed Effects for the Transparency Ratings of Idioms

| | Estimate | Std. Error | t value |
|---|---|---|---|
| Intercept | 20.65 | 6.81 | 3.03 |
| KnowIdiom=Yes | 7.89 | 3.50 | 2.25 |
| logFrequencyIdiom | 0.72 | 0.95 | 0.75 |
| LQ5 | -9.97 | 3.76 | -2.65 |
| LQ7 | 7.90 | 3.55 | 2.22 |
| LikeUsingIdioms | 7.48 | 3.49 | 2.14 |
| Length | 1.98 | 0.87 | 2.26 |
| LSA.Score.Definition | 13.86 | 5.90 | 2.35 |
| I(logFrequencyIdiom\|KnowIdiom=Yes) | 4.81 | 0.87 | 5.50 |
| I(LQ5\|KnowIdiom=Yes) | 9.68 | 1.33 | 7.28 |
| I(LQ7\|KnowIdiom=Yes) | -3.15 | 1.14 | -2.77 |
| I(LikeUsingIdioms\|KnowIdiom=Yes) | -9.28 | 1.08 | -8.58 |
| I(LikeUsingIdioms\|Length) | 0.61 | 0.27 | 2.22 |

Table A.2: Random Effects for the Transparency Ratings of Idioms

| | Groups | Name | Std.Dev | Corr |
|---|---|---|---|---|
| 1 | Idiom | Intercept | 10.38 | |
| 2 | | KnowIdiom=Yes | 8.75 | 0.06 |
| 3 | Subject | Intercept | 12.35 | |
| 4 | Subject | TrialScaled | 3.60 | |
| 5 | Subject | logFrequencyIdiom | 1.10 | |
| 6 | Residual | | 20.48 | |

## A.2 Transparency Rating Reaction Times

Table A.3: Fixed Effects for the RTs to Rate the Transparency of Idioms

| | Estimate | Std. Error | t value |
|---|---|---|---|
| Intercept | 7.47 | 0.11 | 65.51 |
| KnowIdiom=Yes | -0.03 | 0.06 | -0.49 |
| logFrequencyIdiom | 0.01 | 0.01 | 0.42 |
| LikeUsingIdioms | 0.14 | 0.10 | 1.32 |
| TrialScaled | -0.14 | 0.02 | -9.01 |
| I(logFrequencyIdiom\|KnowIdiom=Yes) | -0.04 | 0.02 | -2.33 |
| I(LikeUsingIdioms\|KnowIdiom=Yes) | -0.09 | 0.02 | -3.91 |

Table A.4: Random Effects for the RTs to Rate the Transparency of Idioms

| | Groups | Name | Std.Dev |
|---|---|---|---|
| 1 | Idiom | Intercept | 0.09 |
| 2 | Subject | TrialScaled | 0.06 |
| 3 | Subject | Intercept | 0.45 |
| 4 | Residual | | 0.52 |

# APPENDIX B

# Linear Mixed-Effects Models for Acceptability Ratings of Idiomatic Variation

## B.1 Acceptability Ratings for those who Know the Idiom

Table B.1: Fixed Effects for the Acceptability Ratings for those who Know the Idiom

| | Estimate | Std. Error | t value |
|---|---|---|---|
| Intercept | 89.03 | 5.64 | 15.78 |
| Condition=Concept | -19.89 | 8.61 | -2.31 |
| Condition=Blend | -27.83 | 9.61 | -2.90 |
| Condition=Lexical | -18.29 | 10.19 | -1.80 |
| Condition=Literal | -50.99 | 10.02 | -5.09 |
| Condition=Partial | -44.10 | 10.41 | -4.24 |
| HowOftenUseIdioms | 4.82 | 1.67 | 2.89 |
| Length | -2.65 | 1.30 | -2.03 |
| meanTransparencyRating | 4.17 | 1.12 | 3.71 |
| LQ3 | 4.02 | 1.15 | 3.49 |
| I(HowOftenUseIdioms\|Condition=Concept) | -2.83 | 1.92 | -1.47 |
| I(HowOftenUseIdioms\|Condition=Blend) | -9.16 | 1.91 | -4.79 |
| I(HowOftenUseIdioms\|Condition=Lexical) | -7.62 | 1.94 | -3.93 |
| I(HowOftenUseIdioms\|Condition=Literal) | -9.24 | 1.90 | -4.85 |
| I(HowOftenUseIdioms\|Condition=Partial) | -5.06 | 1.94 | -2.60 |
| I(Length\|Condition=Concept | 1.48 | 2.03 | 0.73 |
| I(Length\|Condition=Blend) | 1.81 | 2.25 | 0.81 |
| I(Length\|Condition=Lexical) | 0.83 | 2.38 | 0.35 |
| I(Length\|Condition=Literal) | 6.66 | 2.34 | 2.84 |
| I(Length\|Condition=Partial) | 4.65 | 2.42 | 1.92 |

Table B.2: Random Effects for the Acceptability Ratings for those who Know the Idiom

| | Groups | Name | Std.Dev | Corr | | | | |
|---|---|---|---|---|---|---|---|---|
| 1 | Subject | Intercept | 8.36 | | | | | |
| 2 | Subject | TrialScaled | 3.77 | | | | | |
| 3 | Idiom | Intercept | 4.38 | | | | | |
| 4 | | Condition=Concept | 8.64 | 0.11 | | | | |
| 5 | | Condition=Blend | 11.74 | -0.12 | 0.26 | | | |
| 6 | | Condition=Lexical | 13.26 | -0.11 | 0.8 | -0.07 | | |
| 7 | | Condition=Literal | 12.49 | -0.48 | 0.22 | 0.45 | 0.13 | |
| 8 | | Condition=Partial | 13.36 | 0.56 | -0.15 | 0.08 | 0.05 | -0.33 |
| 9 | Residual | | 27.61 | | | | | |

## B.2 Acceptability Ratings for those who do Not Know the Idiom

Table B.3: Fixed Effects for the Acceptability Ratings for those who do Not Know the Idiom

| | Estimate | Std. Error | t value |
|---|---|---|---|
| Intercept | 50.76 | 1.75 | 29.08 |
| HowOftenUseIdioms | -3.50 | 1.47 | -2.39 |
| meanTransparencyRating | 4.85 | 1.27 | 3.81 |
| LQ3 | 3.98 | 1.57 | 2.54 |

Table B.4: Random Effects for the Acceptability Ratings for those who do Not Know the Idiom

| | Groups | Name | Std.Dev |
|---|---|---|---|
| 1 | Subject | Intercept | 9.59 |
| 2 | Subject | TrialScaled | 3.89 |
| 3 | Idiom | Intercept | 5.51 |
| 4 | Residual | | 29.29 |

## B.3 Acceptability Ratings for Formal Idiom Blends

Table B.5: Fixed Effects for the Acceptability Ratings of Idiom Blends (difference contrast)

| | Estimate | Std. Error | t value |
|---|---|---|---|
| Intercept | 58.75 | 4.49 | 13.07 |
| meanTransparencyRating | 3.95 | 1.99 | 1.99 |
| KnowExpIdiom=Yes | -4.85 | 3.61 | -1.34 |
| KnowBlendingIdiom=Yes | 1.90 | 3.65 | 0.52 |
| PairedIdiomSemantics=Similar | -16.94 | 5.44 | -3.12 |
| I(PairedIdiomSemantics=Similar\|KnowExpIdiom=Yes) | 13.51 | 5.01 | 2.69 |
| I(PairedIdiomSemantics=Similar\|KnowBlendingIdiom=Yes) | 10.04 | 5.05 | 1.99 |

Table B.6: Fixed Effects for the Acceptability Ratings of Idiom Blends (significance of slope in each condition of 'Know Experimental Idiom')

| | Estimate | Std. Error | t value |
|---|---|---|---|
| Intercept | 55.63 | 4.23 | 13.15 |
| meanTransparencyRating | 4.01 | 2.00 | 2.01 |
| KnowBlendingIdiom=Yes | 6.98 | 2.62 | 2.67 |
| KnowExpIdiom=Yes | -5.58 | 3.61 | -1.55 |
| I(PairedIdiomSemantics=Similar\|KnowExpIdiom=No) | -11.77 | 4.79 | -2.46 |
| I(PairedIdiomSemantics=Similar\|KnowExpIdiom=Yes) | 4.03 | 3.80 | 1.06 |

Table B.7: Fixed Effects for the Acceptability Ratings of Idiom Blends (significance of slope in each condition of 'Know Blending Idiom')

| | Estimate | Std. Error | t value |
|---|---|---|---|
| Intercept | 55.25 | 4.30 | 12.84 |
| meanTransparencyRating | 3.87 | 1.98 | 1.96 |
| KnowExpIdiom=Yes | 1.61 | 2.72 | 0.59 |
| KnowBlendingIdiom=Yes | 0.74 | 3.65 | 0.20 |
| I(PairedIdiomSemantics=Similar\|KnowBlendingIdiom=No) | -9.89 | 4.76 | -2.08 |
| I(PairedIdiomSemantics=Similar\|KnowBlendingIdiom=Yes) | 3.27 | 3.80 | 0.86 |

Table B.8: Random Effects for the Acceptability Ratings of Idiom Blends (difference contrast)

| | Groups | Name | Std.Dev |
|---|---|---|---|
| 1 | Subject | Intercept | 14.34 |
| 2 | Idiom | Intercept | 10.35 |
| 3 | Residual | | 26.88 |

## B.4 RTs to Rate the Acceptability

Table B.9: Fixed Effects for the RTs to Rate the Acceptability

| | Estimate | Std. Error | t value |
|---|---|---|---|
| Intercept | -0.11 | 0.09 | -1.21 |
| Condition=Concept | 0.17 | 0.07 | 2.42 |
| Condition=Blend | -0.03 | 0.07 | -0.36 |
| Condition=Lexical | -0.05 | 0.07 | -0.67 |
| Condition=Literal | 0.05 | 0.07 | 0.77 |
| Condition=Partial | -0.15 | 0.07 | -2.21 |
| KnowIdiom=Yes | -0.23 | 0.06 | -3.58 |
| HowOftenUseIdioms | -0.16 | 0.06 | -2.59 |
| PC1.logFrequency | -0.02 | 0.01 | -2.41 |
| TrialScaled | -0.20 | 0.03 | -7.52 |
| IsL1English=No | 0.38 | 0.13 | 2.92 |
| I(Condition=Concept\|KnowIdiom=Yes) | 0.25 | 0.09 | 2.85 |
| I(Condition=Blend\|KnowIdiom=Yes) | 0.21 | 0.09 | 2.48 |
| I(Condition=Lexical\|KnowIdiom=Yes) | 0.30 | 0.08 | 3.58 |
| I(Condition=Literal\|KnowIdiom=Yes) | 0.22 | 0.08 | 2.58 |
| I(Condition=Partial\|KnowIdiom=Yes) | 0.22 | 0.08 | 2.59 |

Table B.10: Random Effects for the RTs to Rate the Acceptability

| | Groups | Name | Std.Dev |
|---|---|---|---|
| 1 | Subject | Intercept | 0.45 |
| 2 | Subject | TrialScaled | 0.21 |
| 3 | Idiom | Intercept | 0.18 |
| 4 | Residual | | 0.73 |

## B.5 RTs to Rate the Acceptability of Formal Idiom Blends

Table B.11: Fixed Effects for the RTs to Rate the Acceptability of Idiom Blends

| | Estimate | Std. Error | t value |
|---|---|---|---|
| Intercept | -0.15 | 0.08 | -1.89 |
| HowOftenUseIdioms | -0.15 | 0.07 | -2.19 |
| PC1.logFrequency | -0.03 | 0.01 | -2.49 |
| TrialScaled | -0.17 | 0.03 | -5.31 |
| IsL1English=No | 0.38 | 0.14 | 2.66 |

Table B.12: Random Effects for the RTs to Rate the Acceptability of Idiom Blends

| | Groups | Name | Std.Dev |
|---|---|---|---|
| 1 | Subject | Intercept | 0.45 |
| 2 | Subject | TrialScaled | 0.16 |
| 3 | Idiom | Intercept | 0.18 |
| 4 | Residual | | 0.69 |

# APPENDIX C

# Linear Mixed-Effects Models for Eye-Tracking Study on Variation

## C.1 Idiom as AOI: Total Fixation Duration

Table C.1: Fixed Effects for the Total Fixation Duration on the Idiom as an Area Of Interest

| | Estimate | Std. Error | t-value |
|---|---|---|---|
| Intercept | 6.71 | 0.09 | 75.97 |
| Condition=Concept | 0.49 | 0.10 | 5.04 |
| Condition=Blend | 0.08 | 0.10 | 0.75 |
| Condition=Lexical | 0.01 | 0.10 | 0.05 |
| Condition=Literal | -0.19 | 0.10 | -1.94 |
| Condition=Partial | -0.75 | 0.16 | -4.80 |
| KnowIdiom=Yes | -0.18 | 0.04 | -4.32 |
| Length | 0.11 | 0.02 | 6.76 |
| PortionIdiomAltered=Ending | -0.06 | 0.02 | -2.52 |
| PC2.LQ | -0.07 | 0.03 | -2.42 |
| LSA.Score.Paraphrase | 0.24 | 0.07 | 3.49 |
| meanVariationRating | -0.06 | 0.01 | -7.23 |
| Gender=Male | -0.17 | 0.08 | -2.17 |
| TrialScaled | -0.04 | 0.01 | -3.78 |
| I(KnowIdiom=Yes\|Condition=Concept) | 0.06 | 0.05 | 1.16 |
| I(KnowIdiom=Yes\|Condition=Blend) | 0.08 | 0.06 | 1.42 |
| I(KnowIdiom=Yes\|Condition=Lexical) | 0.08 | 0.06 | 1.52 |
| I(KnowIdiom=Yes\|Condition=Literal) | 0.03 | 0.06 | 0.55 |
| I(KnowIdiom=Yes\|Condition=Partial) | 0.17 | 0.06 | 2.75 |
| I(Length\|Condition=Concept) | -0.05 | 0.02 | -2.62 |
| I(Length\|Condition=Blend) | -0.01 | 0.02 | -0.36 |
| I(Length\|Condition=Lexical) | 0.00 | 0.02 | 0.20 |
| I(Length\|Condition=Literal) | 0.02 | 0.02 | 1.04 |
| I(Length\|Condition=Partial) | 0.08 | 0.03 | 2.48 |

Table C.2: Random Effects for the Total Fixation Duration on the Idiom as an Area Of Interest

| | Groups | Name | Std.Dev | Corr | | | | |
|---|---|---|---|---|---|---|---|---|
| 1 | Subject | Intercept | 0.26 | | | | | |
| 2 | Subject | TrialScaled | 0.08 | | | | | |
| 3 | Idiom | Intercept | 0.08 | | | | | |
| 4 | | Condition=Concept | 0.06 | -0.46 | | | | |
| 5 | | Condition=Blend | 0.10 | -0.34 | 0.9 | | | |
| 6 | | Condition=Lexical | 0.09 | -0.49 | 0.73 | 0.57 | | |
| 7 | | Condition=Literal | 0.06 | -0.48 | 0.72 | 0.52 | 0.99 | |
| 8 | | Condition=Partial | 0.24 | -0.19 | 0.62 | 0.25 | 0.45 | 0.52 |
| 9 | Residual | | 0.32 | | | | | |

## C.2 Idiom as AOI: First Fixation Duration

Table C.3: Fixed Effects for the First Fixation Duration on the Idiom as an Area Of Interest

| | Estimate | Std. Error | t-value |
|---|---|---|---|
| Intercept | 5.95 | 0.07 | 86.64 |
| Condition=Concept | 0.29 | 0.07 | 4.00 |
| Condition=Blend | -0.02 | 0.07 | -0.21 |
| Condition=Lexical | -0.06 | 0.07 | -0.86 |
| Condition=Literal | -0.09 | 0.07 | -1.21 |
| Condition=Partial | -0.89 | 0.07 | -11.90 |
| Length | 0.15 | 0.01 | 10.30 |
| PortionIdiomAltered=Ending | -0.00 | 0.03 | -0.03 |
| KnowIdiom=Yes | -0.02 | 0.02 | -1.01 |
| PC2.LQ | -0.02 | 0.02 | -0.89 |
| meanVariationRating | -0.02 | 0.01 | -3.31 |
| LSA.Score.Paraphrase | 0.20 | 0.06 | 3.36 |
| Gender=Male | -0.12 | 0.05 | -2.29 |
| I(Length\|Condition=Concept) | -0.02 | 0.02 | -1.22 |
| I(Length\|Condition=Blend) | 0.01 | 0.02 | 0.49 |
| I(Length\|Condition=Lexical) | 0.02 | 0.02 | 1.12 |
| I(Length\|Condition=Literal) | 0.01 | 0.02 | 0.49 |
| I(Length\|Condition=Partial) | 0.07 | 0.02 | 4.50 |
| I(PortionIdiomAltered=Ending\|Condition=Concept) | -0.03 | 0.04 | -0.73 |
| I(PortionIdiomAltered=Ending\|Condition=Blend) | 0.01 | 0.04 | 0.38 |
| I(PortionIdiomAltered=Ending\|Condition=Lexical) | -0.01 | 0.04 | -0.16 |
| I(PortionIdiomAltered=Ending\|Condition=Literal) | 0.02 | 0.04 | 0.51 |
| I(PortionIdiomAltered=Ending\|Condition=Partial) | 0.25 | 0.04 | 6.96 |
| I(PC2.LQ\|KnowIdiom=Yes) | -0.03 | 0.01 | -2.40 |

Table C.4: Random Effects for the First Fixation Duration on the Idiom as an Area Of Interest

| | Groups | Name | Std.Dev |
|---|---|---|---|
| 1 | Subject | Intercept | 0.17 |
| 2 | Subject | TrialScaled | 0.04 |
| 3 | Idiom | Intercept | 0.06 |
| 4 | Residual | | 0.30 |

## C.3 Idiom as AOI: Fixation Count

Table C.5: Fixed Effects for the Fixation Count on the Idiom as an Area Of Interest

| | Estimate | Std. Error | z-value | p-value |
|---|---|---|---|---|
| Intercept | 1.22 | 0.09 | 13.63 | 0.00 |
| Condition=Concept | 0.56 | 0.09 | 5.89 | 0.00 |
| Condition=Blend | 0.09 | 0.10 | 0.86 | 0.39 |
| Condition=Lexical | 0.03 | 0.10 | 0.27 | 0.79 |
| Condition=Literal | -0.20 | 0.11 | -1.90 | 0.06 |
| Condition=Partial | -0.61 | 0.11 | -5.56 | 0.00 |
| Length | 0.12 | 0.02 | 6.47 | 0.00 |
| PortionIdiomAltered=Ending | -0.06 | 0.02 | -2.49 | 0.01 |
| KnowIdiom=Yes | -0.13 | 0.02 | -6.13 | 0.00 |
| PC2.LQ | -0.06 | 0.02 | -2.71 | 0.01 |
| meanVariationRating | -0.07 | 0.01 | -8.42 | 0.00 |
| LSA.Score.Paraphrase | 0.19 | 0.07 | 2.72 | 0.01 |
| TrialScaled | -0.05 | 0.01 | -4.72 | 0.00 |
| I(Length\|Condition=Concept) | -0.05 | 0.02 | -2.28 | 0.02 |
| I(Length\|Condition=Blend) | 0.01 | 0.02 | 0.33 | 0.74 |
| I(Length\|Condition=Lexical) | 0.01 | 0.02 | 0.60 | 0.55 |
| I(Length\|Condition=Literal) | 0.03 | 0.02 | 1.29 | 0.20 |
| I(Length\|Condition=Partial) | 0.08 | 0.02 | 3.48 | 0.00 |

Table C.6: Random Effects for the Fixation Count on the Idiom as an Area Of Interest

| | Groups | Name | Std.Dev |
|---|---|---|---|
| 1 | Subject | Intercept | 0.22 |
| 2 | Subject | TrialScaled | 0.06 |
| 3 | Idiom | Intercept | 0.07 |

## C.4 Idiom as AOI: Formal Idiom Blends

Table C.7: Fixed Effects for the Total Fixation Duration of Formal Idiom Blends on the Idiom as an Area Of Interest

| | Estimate | Std. Error | t-value |
|---|---|---|---|
| Intercept | 6.76 | 0.09 | 72.44 |
| KnowIdiom=Yes | -0.10 | 0.04 | -2.45 |
| PortionIdiomAltered=Ending | -0.09 | 0.04 | -2.33 |
| Length | 0.12 | 0.02 | 7.11 |
| PC2.LQ | -0.08 | 0.03 | -2.58 |
| meanVariationRating | -0.09 | 0.02 | -4.13 |
| TrialScaled | -0.05 | 0.02 | -2.65 |

Table C.8: Random Effects for the Total Fixation Duration of Formal Idiom Blends on the Idiom as an Area Of Interest

| | Groups | Name | Std.Dev |
|---|---|---|---|
| 1 | Subject | Intercept | 0.29 |
| 2 | Subject | TrialScaled | 0.09 |
| 3 | Idiom | Intercept | 0.11 |
| 4 | Residual | | 0.30 |

## C.5 Manipulated Word as AOI: Total Fixation Duration

Table C.9: Fixed Effects for the Total Fixation Duration on the Manipulated Word in the Idiom as an Area Of Interest

| | Estimate | Std. Error | t-value |
|---|---|---|---|
| Intercept | 5.70 | 0.06 | 98.48 |
| Condition=Concept | 0.47 | 0.06 | 8.28 |
| Condition=Blend | 0.15 | 0.06 | 2.67 |
| Condition=Lexical | 0.09 | 0.06 | 1.54 |
| Condition=Partial | 0.30 | 0.07 | 4.61 |
| PortionIdiomAltered=Ending | 0.27 | 0.06 | 4.49 |
| KnowIdiom=Yes | -0.04 | 0.03 | -1.29 |
| PC2.LQ | -0.10 | 0.03 | -3.12 |
| PC1.logFrequency | 0.03 | 0.01 | 4.70 |
| meanVariationRating | -0.07 | 0.02 | -4.27 |
| TrialScaled | -0.04 | 0.01 | -2.79 |
| I(PortionIdiomAltered=Ending\|Condition=Concept) | -0.12 | 0.08 | -1.46 |
| I(PortionIdiomAltered=Ending\|Condition=Blend) | -0.09 | 0.08 | -1.17 |
| I(PortionIdiomAltered=Ending\|Condition=Lexical) | -0.02 | 0.08 | -0.26 |
| I(PortionIdiomAltered=Ending\|Condition=Partial) | -0.40 | 0.09 | -4.42 |
| I(PC2.LQ\|KnowIdiom=Yes) | 0.06 | 0.02 | 2.27 |

Table C.10: Random Effects for the Total Fixation Duration on the Manipulated Word in the Idiom as an Area Of Interest

| | Groups | Name | Std.Dev | Corr | | | |
|---|---|---|---|---|---|---|---|
| 1 | Subject | Intercept | 0.23 | | | | |
| 2 | Subject | TrialScaled | 0.07 | | | | |
| 3 | Idiom | Intercept | 0.13 | | | | |
| 4 | | Condition=Concept | 0.18 | -0.57 | | | |
| 5 | | Condition=Blend | 0.14 | -0.11 | -0.11 | | |
| 6 | | Condition=Lexical | 0.18 | -0.64 | 0.5 | -0.29 | |
| 7 | | Condition=Partial | 0.22 | -0.04 | 0.25 | -0.19 | -0.14 |
| 8 | Residual | | 0.54 | | | | |

## C.6 Manipulated Word as AOI: First Fixation Duration

Table C.11: Fixed Effects for the First Fixation Duration on the Manipulated Word in the Idiom as an Area Of Interest

| | Estimate | Std. Error | t-value |
|---|---|---|---|
| Intercept | 4.98 | 0.05 | 95.24 |
| Condition=Concept | 0.10 | 0.05 | 1.96 |
| Condition=Blend | 0.04 | 0.05 | 0.82 |
| Condition=Lexical | 0.00 | 0.05 | 0.07 |
| Condition=Partial | 0.13 | 0.05 | 2.54 |
| PortionIdiomAltered=Ending | 0.36 | 0.05 | 7.00 |
| KnowIdiom=Yes | -0.00 | 0.03 | -0.06 |
| PC2.LQ | 0.02 | 0.03 | 0.89 |
| LSA.Score.Paraphrase | -0.22 | 0.06 | -3.63 |
| I(PortionIdiomAltered=Ending\|Condition=Concept) | -0.28 | 0.07 | -3.98 |
| I(PortionIdiomAltered=Ending\|Condition=Blend) | -0.10 | 0.07 | -1.43 |
| I(PortionIdiomAltered=Ending\|Condition=Lexical) | -0.13 | 0.07 | -1.77 |
| I(PortionIdiomAltered=Ending\|Condition=Partial) | -0.12 | 0.07 | -1.63 |
| I(PC2.LQ\|KnowIdiom=Yes) | -0.06 | 0.03 | -2.30 |

Table C.12: Random Effects for the First Fixation Duration on the Manipulated Word in the Idiom as an Area Of Interest

| | Groups | Name | Std.Dev |
|---|---|---|---|
| 1 | Subject | Intercept | 0.13 |
| 2 | Idiom | Intercept | 0.00 |
| 3 | Residual | | 0.56 |

## C.7 Manipulated Word as AOI: Fixation Count

Table C.13: Fixed Effects for Fixation Count on the Manipulated Word in the Idiom as an Area Of Interest

| | Estimate | Std. Error | z-value | p-value |
|---|---|---|---|---|
| Intercept | 0.44 | 0.07 | 6.56 | 0.00 |
| Condition=Concept | 0.54 | 0.06 | 8.59 | 0.00 |
| Condition=Blend | 0.19 | 0.07 | 2.83 | 0.00 |
| Condition=Lexical | 0.17 | 0.07 | 2.49 | 0.01 |
| Condition=Partial | 0.28 | 0.07 | 4.17 | 0.00 |
| PortionIdiomAltered=Ending | -0.01 | 0.07 | -0.12 | 0.91 |
| PC1.logFrequency | 0.03 | 0.01 | 2.44 | 0.01 |
| KnowIdiom=Yes | -0.12 | 0.04 | -2.97 | 0.00 |
| PC2.LQ | -0.07 | 0.03 | -2.64 | 0.01 |
| meanVariationRating | -0.09 | 0.02 | -5.29 | 0.00 |
| TrialScaled | -0.03 | 0.02 | -2.09 | 0.04 |
| I(PortionIdiomAltered=Ending\|Condition=Concept) | 0.07 | 0.09 | 0.86 | 0.39 |
| I(PortionIdiomAltered=Ending\|Condition=Blend) | -0.02 | 0.09 | -0.18 | 0.86 |
| I(PortionIdiomAltered=Ending\|Condition=Lexical) | 0.03 | 0.09 | 0.28 | 0.78 |
| I(PortionIdiomAltered=Ending\|Condition=Partial) | -0.41 | 0.10 | -4.33 | 0.00 |
| I(PC1.logFrequency\|Condition=Concept) | 0.01 | 0.02 | 0.44 | 0.66 |
| I(PC1.logFrequency\|Condition=Blend) | -0.01 | 0.02 | -0.32 | 0.75 |
| I(PC1.logFrequency\|Condition=Lexical) | 0.00 | 0.02 | 0.05 | 0.96 |
| I(PC1.logFrequency\|Condition=Partial) | -0.04 | 0.02 | -2.42 | 0.02 |

Table C.14: Random Effects for Fixation Count on the Manipulated Word in the Idiom as an Area Of Interest

| | Groups | Name | Std.Dev |
|---|---|---|---|
| 1 | Subject | Intercept | 0.21 |
| 2 | Subject | TrialScaled | 0.07 |
| 3 | Idiom | Intercept | 0.07 |

## C.8 Manipulated Word as AOI: Formal Idiom Blends

Table C.15: Fixed Effects for Total Fixation Duration on the Manipulated Word in Formal Idiom Blends as an Area Of Interest

| | Estimate | Std. Error | t-value |
|---|---|---|---|
| Intercept | 5.93 | 0.14 | 43.26 |
| KnowBlendingIdiom=Yes | 0.25 | 0.10 | 2.37 |
| PortionIdiomAltered=Ending | 0.41 | 0.12 | 3.32 |
| PC2.LQ | 0.03 | 0.05 | 0.62 |
| meanVariationRating | -0.08 | 0.04 | -2.20 |
| LSA.Score.Definition | -0.55 | 0.18 | -3.07 |
| I(PortionIdiomAltered=Ending\|KnowBlendingIdiom=Yes) | -0.36 | 0.14 | -2.62 |
| I(PC2.LQ\|KnowBlendingIdiom=Yes) | -0.11 | 0.05 | -2.18 |

Table C.16: Random Effects for Total Fixation Duration on the Manipulated Word in Formal Idiom Blends as an Area Of Interest

| | Groups | Name | Std.Dev |
|---|---|---|---|
| 1 | Subject | Intercept | 0.25 |
| 2 | Idiom | Intercept | 0.13 |
| 3 | Residual | | 0.57 |

## C.9 Surprising Word as AOI: Total Fixation Duration

Table C.17: Fixed Effects for the Total Fixation Duration on the Surprising Word in the Idiom as an Area Of Interest

| | Estimate | Std. Error | t-value |
|---|---|---|---|
| Intercept | 5.97 | 0.07 | 84.62 |
| Condition=Concept | 0.24 | 0.06 | 3.75 |
| Condition=Blend | 0.13 | 0.06 | 2.01 |
| Condition=Lexical | 0.11 | 0.05 | 2.08 |
| Condition=Partial | 0.23 | 0.07 | 3.44 |
| PortionIdiomAltered=Ending | 0.00 | 0.07 | 0.01 |
| KnowIdiom=Yes | -0.10 | 0.03 | -2.91 |
| PC2.LQ | -0.07 | 0.03 | -2.24 |
| PC1.logFrequency | 0.03 | 0.01 | 4.46 |
| meanVariationRating | -0.05 | 0.02 | -2.57 |
| TrialScaled | -0.03 | 0.02 | -2.11 |
| I(PortionIdiomAltered=Ending\|Condition=Concept) | 0.15 | 0.09 | 1.69 |
| I(PortionIdiomAltered=Ending\|Condition=Blend) | -0.04 | 0.09 | -0.48 |
| I(PortionIdiomAltered=Ending\|Condition=Lexical) | -0.01 | 0.08 | -0.18 |
| I(PortionIdiomAltered=Ending\|Condition=Partial) | -0.30 | 0.09 | -3.26 |

Table C.18: Random Effects for the Total Fixation Duration on the Surprising Word in the Idiom as an Area Of Interest

| | Groups | Name | Std.Dev | Corr | | | |
|---|---|---|---|---|---|---|---|
| 1 | Subject | Intercept | 0.29 | | | | |
| 2 | Subject | TrialScaled | 0.10 | | | | |
| 3 | Idiom | Intercept | 0.22 | | | | |
| 4 | | Condition=Concept | 0.24 | -0.79 | | | |
| 5 | | Condition=Blend | 0.23 | -0.48 | 0.33 | | |
| 6 | | Condition=Lexical | 0.13 | -0.45 | 0.3 | -0.29 | |
| 7 | | Condition=Partial | 0.24 | -0.33 | 0.32 | 0.45 | -0.23 |
| 8 | Residual | | 0.54 | | | | |

## C.10 Surprising Word as AOI: First Fixation Duration

Table C.19: Fixed Effects for the First Fixation Duration on the Surprising Word in the Idiom as an Area Of Interest

| | Estimate | Std. Error | t-value |
|---|---|---|---|
| Intercept | 5.08 | 0.05 | 100.78 |
| Condition=Concept | -0.09 | 0.05 | -1.78 |
| Condition=Blend | 0.09 | 0.05 | 1.70 |
| Condition=Lexical | 0.07 | 0.05 | 1.31 |
| Condition=Partial | -0.07 | 0.05 | -1.32 |
| PortionIdiomAltered=Ending | 0.17 | 0.05 | 3.22 |
| KnowIdiom=Yes | -0.00 | 0.03 | -0.08 |
| PC2.LQ | 0.03 | 0.03 | 0.96 |
| PC1.logFrequency | 0.01 | 0.00 | 2.10 |
| I(PortionIdiomAltered=Ending\|Condition=Concept) | -0.09 | 0.07 | -1.31 |
| I(PortionIdiomAltered=Ending\|Condition=Blend) | -0.14 | 0.07 | -1.95 |
| I(PortionIdiomAltered=Ending\|Condition=Lexical) | -0.18 | 0.07 | -2.46 |
| I(PortionIdiomAltered=Ending\|Condition=Partial) | 0.08 | 0.08 | 1.08 |
| I(PC2.LQ\|KnowIdiom=Yes) | -0.07 | 0.03 | -2.70 |

Table C.20: Random Effects for the First Fixation Duration on the Surprising Word in the Idiom as an Area Of Interest

| | Groups | Name | Std.Dev |
|---|---|---|---|
| 1 | Subject | Intercept | 0.17 |
| 2 | Idiom | Intercept | 0.04 |
| 3 | Residual | | 0.57 |

## C.11 Surprising Word as AOI: Fixation Count

Table C.21: Fixed Effects for Fixation Count on the Surprising Word in the Idiom as an Area Of Interest

| | Estimate | Std. Error | z-value | p-value |
|---|---|---|---|---|
| Intercept | 0.65 | 0.07 | 9.35 | 0.00 |
| Condition=Concept | 0.33 | 0.06 | 5.61 | 0.00 |
| Condition=Blend | 0.08 | 0.06 | 1.22 | 0.22 |
| Condition=Lexical | 0.01 | 0.06 | 0.24 | 0.81 |
| Condition=Partial | 0.27 | 0.06 | 4.25 | 0.00 |
| PortionIdiomAltered=Ending | -0.22 | 0.08 | -2.98 | 0.00 |
| PC1.logFrequency | 0.03 | 0.01 | 1.89 | 0.06 |
| KnowIdiom=Yes | -0.15 | 0.04 | -3.80 | 0.00 |
| PC2.LQ | -0.07 | 0.03 | -2.54 | 0.01 |
| meanVariationRating | -0.05 | 0.02 | -2.80 | 0.01 |
| I(PortionIdiomAltered=Ending\|Condition=Concept) | 0.30 | 0.08 | 3.59 | 0.00 |
| I(PortionIdiomAltered=Ending\|Condition=Blend) | 0.12 | 0.09 | 1.29 | 0.20 |
| I(PortionIdiomAltered=Ending\|Condition=Lexical) | 0.20 | 0.09 | 2.18 | 0.03 |
| I(PortionIdiomAltered=Ending\|Condition=Partial) | -0.37 | 0.09 | -4.02 | 0.00 |
| I(PC1.logFrequency\|Condition=Concept) | 0.01 | 0.02 | 0.83 | 0.41 |
| I(PC1.logFrequency\|Condition=Blend) | -0.01 | 0.02 | -0.44 | 0.66 |
| I(PC1.logFrequency\|Condition=Lexical) | 0.00 | 0.02 | 0.26 | 0.79 |
| I(PC1.logFrequency\|Condition=Partial) | -0.04 | 0.02 | -2.58 | 0.01 |

Table C.22: Random Effects for Fixation Count on the Surprising Word in the Idiom as an Area Of Interest

| | Groups | Name | Std.Dev |
|---|---|---|---|
| 1 | Subject | Intercept | 0.25 |
| 2 | Subject | TrialScaled | 0.08 |
| 3 | Idiom | Intercept | 0.14 |

## C.12 Surprising Word as AOI: Formal Idiom Blends

Table C.23: Fixed Effects for Total Fixation Duration on the Surprising Word in Formal Idiom Blends as an Area Of Interest

| | Estimate | Std. Error | t-value |
|---|---|---|---|
| Intercept | 6.01 | 0.06 | 96.84 |
| PC2.LQ | -0.09 | 0.04 | -2.11 |

Table C.24: Random Effects for Total Fixation Duration on the Surprising Word in Formal Idiom Blends as an Area Of Interest

| | Groups | Name | Std.Dev |
|---|---|---|---|
| 1 | Subject | Intercept | 0.37 |
| 2 | Idiom | Intercept | 0.24 |
| 3 | Residual | | 0.55 |

# APPENDIX D

# Linear Mixed-Effects Models for Idiomatic Variation in the Elicitation Task

## D.1 Variation in Form

Table D.1: Fixed Effects for the Variation in Form

| | Estimate | Std. Error | z-value | p-value |
|---|---|---|---|---|
| Intercept | 0.60 | 1.02 | 0.59 | 0.56 |
| Length | 0.44 | 0.11 | 4.04 | 0.00 |
| LQ7 | -0.14 | 0.13 | -1.13 | 0.26 |
| logFrequencyIdiom | -0.46 | 0.11 | -4.14 | 0.00 |
| Version=B | 0.14 | 0.51 | 0.27 | 0.79 |
| Version=C | -1.33 | 0.53 | -2.50 | 0.01 |
| I(Length\|LQ7) | -0.04 | 0.02 | -2.34 | 0.02 |

Table D.2: Random Effects for the Variation in Form

| | Groups | Name | Std.Dev |
|---|---|---|---|
| 1 | Idiom | Intercept | 0.83 |
| 2 | Subject | Intercept | 1.10 |

## D.2 Variation in Syntax

Table D.3: Fixed Effects for the Variation in Syntax

| | Estimate | Std. Error | z-value | p-value |
|---|---|---|---|---|
| Intercept | 0.69 | 0.60 | 1.15 | 0.25 |
| LQ7 | -0.22 | 0.06 | -3.59 | 0.00 |
| Version=B | -0.18 | 0.32 | -0.56 | 0.58 |
| Version=C | -1.00 | 0.33 | -3.04 | 0.00 |
| KnowIdiom=Yes | 1.44 | 0.40 | 3.62 | 0.00 |
| logFrequencyIdiom | -0.07 | 0.11 | -0.63 | 0.53 |
| I(logFrequencyIdiom\|KnowIdiom=Yes) | -0.27 | 0.08 | -3.22 | 0.00 |

Table D.4: Random Effects for the Variation in Syntax

| | Groups | Name | Std.Dev |
|---|---|---|---|
| 1 | Idiom | Intercept | 1.06 |
| 2 | Subject | Intercept | 0.64 |

## D.3 Variation in Semantics

Table D.5: Fixed Effects for the Variation in Semantics

| | Estimate | Std. Error | z-value | p-value |
|---|---|---|---|---|
| Intercept | 1.34 | 0.60 | 2.25 | 0.02 |
| LQ7 | -0.26 | 0.08 | -3.26 | 0.00 |
| Version=B | 0.29 | 0.41 | 0.69 | 0.49 |
| Version=C | -1.23 | 0.43 | -2.87 | 0.00 |
| KnowIdiom=Yes | 0.98 | 0.41 | 2.37 | 0.02 |
| logFrequencyIdiom | -0.17 | 0.09 | -1.86 | 0.06 |
| I(logFrequencyIdiom\|KnowIdiom=Yes) | -0.22 | 0.09 | -2.51 | 0.01 |

Table D.6: Random Effects for the Variation in Semantics

| | Groups | Name | Std.Dev |
|---|---|---|---|
| 1 | Idiom | Intercept | 0.79 |
| 2 | Subject | TrialScaled | 0.25 |
| 3 | Subject | Intercept | 0.87 |

# APPENDIX E

# Frequency Measures in the Principal Components Analysis

Frequency of the Idiom
Frequency of Word 1
Frequency of Word 2
Frequency of Word 3
Frequency of Word 4
Frequency of Word 5
Frequency of Word 6
Frequency of Word 7
Frequency of Words 1 & 2
Frequency of Words 2 & 3
Frequency of Words 3 & 4
Frequency of Words 4 & 5
Frequency of Words 5 & 6
Frequency of Words 6 & 7
Frequency of Words 1 through 3
Frequency of Words 1 through 4
Frequency of Words 1 through 5
Frequency of Words 1 through 6
Frequency of Words 1 through 7
Frequency of Words 2 through 4
Frequency of Words 2 through 5
Frequency of Words 2 through 6
Frequency of Words 2 through 7
Frequency of Words 3 through 5
Frequency of Words 3 through 6
Frequency of Words 3 through 7
Frequency of Words 4 through 6
Frequency of Words 4 through 7
Frequency of Words 5 through 7