

THE DIFFERENCE BETWEEN BUCKET-KICKING AND KICKING THE BUCKET:  
UNDERSTANDING IDIOM FLEXIBILITY

by

Megan Schildmier Stone

---

Copyright © Megan Schildmier Stone 2016

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF LINGUISTICS

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2016

ProQuest Number: 10248559

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10248559

Published by ProQuest LLC (2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 - 1346

THE UNIVERSITY OF ARIZONA  
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Megan Schildmier Stone, titled *The Difference Between Bucket-Kicking and Kicking the Bucket: Understanding Idiom Flexibility* and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

\_\_\_\_\_  
Heidi Harley Date: August 10, 2016

\_\_\_\_\_  
Andrew Carnie Date: August 10, 2016

\_\_\_\_\_  
Mike Hammond Date: August 10, 2016

\_\_\_\_\_  
Simin Karimi Date: August 10, 2016

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

\_\_\_\_\_  
Dissertation Director: Heidi Harley Date: August 10, 2016

#### STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of the requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that an accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Megan Schildmier Stone

## ACKNOWLEDGEMENTS

When one undertakes a work of this magnitude, and especially when so many years are spent in its creation, it seems foolish to even endeavor to thank all of those who have helped bring it to fruition. Yet more foolish indeed is the one thinks herself above expressing gratitude. And so I try, though I will surely fail.

First and foremost, I offer my sincerest thanks to Heidi Harley, my dissertation advisor. From my first bewildering weeks as a fledgling linguist in the graduate program to my final moments as a PhD Candidate, Heidi has been steadfast and encouraging. Her genuine enthusiasm for my research interests has been surpassed only by her care for my growth and development as a person. It is exceedingly rare to find such kindness and such intellect intertwined in one delightful person. Anyone who knows Heidi will see her mark throughout this dissertation, but her influence on me has extended far beyond my intellectual enterprises and into all aspects of my life. I'm certain that this dissertation would not exist were it not for Heidi's guidance; I'm equally certain that I am a better person for having had her in my life.

I also owe much to my remaining committee members. Mike Hammond's rigorous commentary pushed me to think more deeply and convey my analysis more clearly. He offered a refreshing perspective on the theoretical modelling, and he also helped significantly improve the discussion of the statistical analyses (though he is not to be blamed for any remaining oddities or oversights in that regard). Andrew Carnie taught me so much, both in one-on-one discussions and in his classes, about appreciating other models of language and about the structure of linguistic argumentation. Simin Karimi encouraged me to consider the implications of my analysis for Persian and other languages; although much of that discussion did not make it into this dissertation, she has given me plenty to do for the next few years!

This work was also influenced, in big ways and small, by discussions with colleagues and professors too numerous to list here, including Diana Archangeli, Andy Barss, David Beaver, John Beavers, Dane Bell, Jaehoon Choi, Katrin Erk, Christiane Fellbaum, Amy Fountain, Will Harwood, Hyun-Kyoung Jung, Greg Key, Lionel Mathieu, Dave Medeiros, Rosamund Moon, Jorge Muriel, Ryan Nelson, Janet Nicol, Diane Ohala, Jaime Parchment, Jessamyn Schertz, Sylvia Reed Schreiner, Rebecca Sharp, Ryan Walter Smith, Jae-Hyun Sung, Alex Trueman, Adam Ussishkin, Natasha Warner, Steve Wechsler, and Samantha Wray. I'm especially grateful to Jeffrey Punske for collaboration on ideas that are central to this dissertation; to Zach Childers and Kevin Schluter for extensive discussion and encouragement; and to Colin Gorrie for holding me to a higher standard than I ever would have held myself. I also owe a debt of gratitude to Aron Aji, Lee Garver, Marshall Gregory, Andrew Levy, and Travis Ryan, who introduced me to analytical thinking and intellectual rigor as an undergraduate at Butler University; and especially to Bill Watts, my undergraduate thesis advisor, who gave me my first taste of linguistics.

I would be remiss if I didn't thank the many friends who have brought balance to my life, including Hilary Andersen, Clarissa Correa, Sue Hargreaves, Kaitlin Petras, Lisa

Rivers, Jenna Roberts, Julie Scott, Kat Scott, Lisa Sramek, Janel Tasker and everyone at Equine Rehabilitation of Central Texas and Hopeful Hearts Therapeutic Riding; and Teresa Chin, Julie Hamsa, Rachel Holland, Paul McCaughtry, CJ Neagle, Alec Rains, Catie Sauer and the rest of my krav maga family at The Mat Martial Arts and Fitness. Some days, going for a quiet ride at the barn or hitting a bag at the gym was the only thing that got me through!

Finally, I want to express my sincerest appreciation to my family. To my mom Ann, my dad Mike, my brother John, and my sister Kristen—you have given me the kind of unconditional love that transcends space and time. To my cats Pova and Terty—you teach me anew every day the value of patience (and snuggles). To my faithful dog Leah—you have reminded me that walks are essential, even on the busiest of days. And to my husband Ted—you have given me the confidence to dream great dreams and the space to pursue them.

*Behind every great achievement is a dreamer of great dreams.*      ~Robert K. Greenleaf

## DEDICATION

*To Heidi – mentor, inspiration, friend.*

## TABLE OF CONTENTS

LIST OF FIGURES .....	10
LIST OF TABLES .....	10
ABSTRACT.....	11
CHAPTER 1: INTRODUCTION.....	12
1.1 Overview .....	12
1.2 What exactly <i>is</i> an idiom? .....	18
1.2.1 Should idiom be defined syntactically or semantically?.....	22
1.2.2 Must idioms be ambiguous? .....	24
1.2.3 Can idioms include meaningless elements?.....	25
1.2.4 Can monomorphemic words or compounds be idioms?.....	26
1.2.5 Should the class of idioms include larger expressions such as proverbs or clichés?.....	29
1.2.6 How can, or should, idioms be distinguished from collocations, polysemy, light verb constructions, and metaphors?.....	30
1.2.7 Summary .....	38
1.3 What constrains idiomaticity? .....	38
1.4 What are some characteristics of idiom behavior?.....	53
1.4.1 The range of idiom flexibility .....	53
1.4.2 Experimental approaches to idiom flexibility.....	63
1.4.3 Corpus work on idiom flexibility.....	66
1.5 Conclusion and overview of the dissertation .....	67
CHAPTER 2: EXPERIMENTAL INVESTIGATIONS OF IDIOM FLEXIBILITY .....	71
2.1 Introduction .....	71
2.2 Experiment One.....	76
2.2.1 Methodology .....	77
2.2.2 Predictions.....	81
2.2.3 Results.....	81
2.2.4 Discussion .....	82
2.3 Experiment Two.....	89
2.3.1 Methodology .....	90
2.3.2 Predictions.....	97
2.3.3 Results.....	100
2.3.4 Discussion .....	102



2.4	Conclusion.....	106
CHAPTER 3: CORPUS INVESTIGATIONS OF IDIOM FLEXIBILITY .....		108
3.1	Introduction .....	108
3.2	Previous work.....	113
3.3	Google Books study .....	115
3.3.1	Methodology .....	115
3.3.2	Results.....	123
3.4	COCA study .....	137
3.4.1	Methodology .....	137
3.4.2	Results.....	143
3.4.3	Summary .....	156
3.5	Discussion .....	157
3.5.1	Comparing corpora .....	157
3.5.2	Implications of results.....	161
3.6	Conclusion.....	162
CHAPTER 4: A THEORETICAL MODEL OF IDIOM FLEXIBILITY .....		164
4.1	Introduction .....	164
4.2	Inadequate approaches to idiom flexibility .....	165
4.2.1	Overview.....	165
4.2.2	Model Type I: Predicts no regularities.....	166
4.2.3	Model Type II: Predicts exhaustive subset relationships.....	168
4.2.4	Model Type III: Predicts a cline of flexibility .....	171
4.2.5	Model Type IV: Predicts two classes of idioms .....	176
4.2.6	Model Type V: Anything is possible .....	194
4.2.7	Summary .....	202
4.3	The Structural Requirement Model.....	203
4.3.1	Details of the model .....	203
4.3.2	Re-examining the verbal functional structure.....	213
4.3.3	Summary .....	221
4.4	Conclusion.....	221
CHAPTER 5: CONCLUSION .....		224
APPENDIX A: STIMULI FOR EXPERIMENT ONE .....		229
APPENDIX B: STIMULI AND FILLERS FOR EXPERIMENT TWO .....		236
B.1	Stimuli .....	236

B.2	Fillers.....	244
APPENDIX C: ANALYZING THE COCA RESULTS AS BINARY .....		246
C.1	Fisher’s Exact Test for COCA .....	246
C.2	Multiple Correspondence Analysis for COCA .....	248
REFERENCES .....		251

## LIST OF FIGURES

Figure 2.1: Histogram of responses to Passive items .....	82
Figure 2.2: Histogram of responses to all items .....	83
Figure 2.3: Histograms of responses to Active, Passive, and Gerund items .....	87
Figure 2.4: Mean reading times for Congruent and Incongruent items .....	101
Figure 2.5: Histogram of responses to Passive ~ Congruent items .....	102
Figure 3.1: Multiple Correspondence Analysis (MCA) for Google Books.....	132
Figure 3.2: Scree Plot for Google Books MCA .....	135
Figure 3.3: Principal Component Analysis (PCA) for COCA.....	153
Figure 3.4: Scree Plot for COCA PCA .....	155
Figure C.1: Multiple Correspondence Analysis (MCA) for COCA .....	248
Figure C.2: Scree Plot for COCA MCA .....	250

## LIST OF TABLES

Table 2.1: Means of responses by Syntax.....	87
Table 2.2: Idiomatic/literal clause pairings.....	93
Table 2.3: Mean reading times for all items in milliseconds.....	101
Table 2.4: Mean reading times for two individual items in milliseconds.....	105
Table 3.1: Variation types.....	120
Table 3.2: Results of Google Books corpus search .....	125
Table 3.3: Conditional probabilities $P(A B)$ for Google Books .....	127
Table 3.4: Contingency table for Google Books data.....	129
Table 3.5: Fisher's Exact Test results for Google Books .....	130
Table 3.6: Multiple Correspondence Analysis (MCA) for Google Books .....	134
Table 3.7: Results of COCA corpus search .....	145
Table 3.8: Conditional probabilities $P(A B)$ for COCA .....	148
Table 3.9: Kendall's Tau rank correlations for COCA.....	150
Table 3.10: Principal Component Analysis (PCA) for COCA .....	154
Table 4.1: Decomposability as a predictor of idiomatic flexibility .....	191
Table 4.2: Patterns of idiomaticity predicted by the structural requirement model.....	207
Table C.1: Fisher's Exact Test results for COCA.....	247
Table C.2: Multiple Correspondence Analysis (MCA) for COCA .....	249

## ABSTRACT

The question of how to integrate idioms into standard theories of grammar has been a matter of investigation since at least the beginning of generative grammar. Idioms are uniquely positioned at the interface between the lexicon and the syntax, demonstrating properties of both words and phrases. On the one hand, idioms behave like stored units, arbitrary correspondences between sound and meaning that must simply be memorized by speakers of the language. In this way, they are similar to words, which have long been recognized as arbitrary sound-meaning pairs (cf. Saussure (1986)'s *arbitrariness of the sign*). On the other hand, idioms in the traditional sense are multiword units, often with some degree of syntactic flexibility, ranging from tense inflection (e.g. *Eli kicked the bucket yesterday* vs. *I'm pretty sure Eli's going to kick the bucket tomorrow*) to passivizability (e.g. *Lisa spilled the beans* vs. *The beans were spilled (by Lisa)*), and beyond. This places idioms in the purview of the syntax, where the combination and manipulation of multiword units is typically assumed to take place. Idioms, then, bridge the gap between the lexicon and the syntax, challenging traditional assumptions about grammar.

This dissertation provides a proposal for dealing with just such issues. I provide an account of idiomatic representations that is consistent with theoretical and empirical research in the field. I explore what kinds of structures are licensed to have special idiomatic interpretations, and I present novel experimental and corpus results that bear on the issue of how idioms are represented. Ultimately, I argue that the structural requirement model alone is able to sufficiently account for the data.

## CHAPTER 1: INTRODUCTION

### 1.1 Overview

Idiomatic expressions have long been problematic in linguistic theory, primarily because they appear to be complex syntactic objects which nonetheless behave semantically as listed lexical items. Lexical items are not typically syntactically complex; syntactically complex structures are not typically listed. Thus idioms exist at an uneasy nexus in language, and linguistic theory has struggled to provide an adequate and predictive model of idiom behavior—indeed, as we shall see, it has even been unclear which expressions actually count as idioms and which do not.

The goal of this dissertation is to investigate the syntactic behavior of idioms, in particular their syntactic flexibility. Which idioms can appear in structures other than their canonical ones, and what are these noncanonical structures? Are there any relationships between different types of syntactic flexibility? If such relationships exist, why do they exist—that is, what kind of model of idiomaticity might be able to predict the observed distributions?

The findings here confirm that syntactic flexibility is not a binary property of verb-object idioms, such that some idioms can appear in noncanonical structures while others cannot; rather, there are more and less flexible idioms, as well as completely fixed ones. The primary contribution, however, is that the variation I document is not characterizable as a simple cline. Rather, there are subset relationships between different noncanonical structures. I present evidence that, if an idiom can appear in nominal

gerundization, it can also appear in object incorporation and passivization. If it is attested in object incorporation, passivization is always possible. Finally, there are idioms that can only appear in passivization.

Importantly, I demonstrate that the subset relationships observed are completely unexpected within most well-known models of idiom flexibility, but are completely expected in the model adopted here. I show that these noncanonical variations relate to independently motivated syntactic hypotheses about the internal structure of the verb phrase, and consequently I argue for a syntactic characterization of the constrained domain of an idiom. The more functional syntactic structure needed to license an idiomatic interpretation, the less syntactic flexibility that idiom will demonstrate.

Before we can tackle the problem of idiom behavior, however, we might well wonder what an idiom even *is*. Open any collection, such as the Oxford Dictionary of English Idioms (Ayto 2010) or <http://idioms.thefreedictionary.com> (Idioms and Phrases - TheFreeDictionary), and you are likely to find a wide range of phenomena, from verb-particle constructions such as *look up* to noun phrases such as *the Big Apple*, from verb phrases such as *make a mountain out of a mole hill* to proverbs such as *Don't look a gift horse in the mouth*. Even single words with unusual metaphorical interpretations grace the pages (or pixels) of these tomes. Yet, although there are many popular conceptions of *idiom*, there is really no consensus from a technical perspective, and linguists approach these expressions with an equal measure of delight and dread. On the one hand, idioms are pervasive in language and have interesting properties which can give insight into other aspects of the grammar; on the other, they are challenging for many theories of

grammar because they do not fall neatly in line with the properties of either words or phrases.

An idiom is loosely conceived of as a group of words which has an unexpected meaning given the meanings of the individual words that make it up. It is thus seen as being in some sense both phrasal and word-like: it is phrasal in that it is composed of more than one word, but it is word-like in that there seems to be a single, memorized meaning assigned to it.

There is an additional facet of idioms that makes them particularly challenging from the researcher's point of view: They are especially troublesome when it comes to the most common methodology employed to probe speakers' linguistic representations. Ask native speakers of English whether *The bullet was bitten by Gemini* can mean that Gemini did something unpleasant, and you will get a whole lot of puzzled looks, shrugged shoulders, and furrowed brows in reply. This is likely due to the nuanced nature of idiomatic judgments, which require subjects to distinguish between syntactic well-formedness and semantic anomaly. In this case the question is not whether the string of words *The bullet was bitten by Gemini* is a grammatically acceptable sentence of English; the question is whether or not this string can receive an idiomatic interpretation. This task, while similar to standard semantic ambiguity tests, is more challenging because the string's two possible meanings are not computed in the same way; the literal meaning is straightforwardly compositional, while the figurative meaning is accessed via some

special mechanism.<sup>1</sup> The fact that not all idioms have a semantically well-formed literal interpretation is an additional confounding factor.<sup>2</sup>

Many researchers have observed that idioms are difficult to work with. As noted by Fraser (1970:23), “idioms, more than most aspects of language, vary enormously from speaker to speaker.” There is evidence for this even in the linguistics literature, where the same idiom may be assigned different judgments by different researchers. For instance, Fraser (1970:40) claims that *throw in the sponge* can undergo particle movement (i.e. *throw the sponge in*), while Newmeyer (1974:335) claims that it cannot. Similarly, Katz & Postal (1963:281) find *kick the bucket* acceptable in the imperative (i.e. *Kick the bucket!*), while Fraser (1970:34–35) does not. Examples like these are abundant and serve to elucidate the fact that idiomatic data are messy and variable.

Given this, it is natural for researchers to turn to experimental and corpus-based investigations. Well-designed experimental protocols may help to either eliminate or quantify this variability: If the variability results from the nature of these specific kinds of judgment tasks, experimental design may be able to circumvent this difficulty. It is possible for a controlled experiment measuring reading times or tracking eye movements to access a subject’s intuitions about the acceptability of a phrase without asking about those intuitions directly. If the variability is not merely an artefact of the task,

---

<sup>1</sup> The model of idioms advocated for in this dissertation assumes some level of compositionality for the figurative meaning; hence, compositionality alone cannot differentiate the literal and figurative meanings. See Section 1.3 for additional discussion.

<sup>2</sup> This could partially account for the fact that idiom flexibility correlates with lack of ambiguity in idioms, i.e. idioms with no plausible literal interpretation are more likely to be acceptable in noncanonical forms (Gibbs & Nayak 1989).



experiments might still help to filter out some of the “noise” in subject responses to elucidate the full range of variation, for example by providing a larger sample size of subjects giving feedback on a fixed set of carefully constructed stimuli. Additionally, corpus-based studies could give insight into how idioms are actually being used without the need to directly ask speakers about idiom usage.

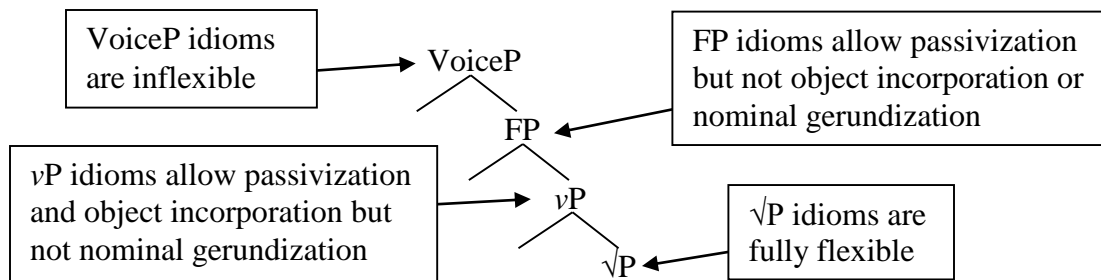
The present studies contribute to the growing body of experimental and corpus work on idioms, revealing a previously unrecognized pattern of behavior: If a given idiom occurs in nominal gerundization, it also occurs in object incorporation; and if it occurs in object incorporation, it also occurs in passivization. This kind of patterned subset behavior is largely unexplored in the literature and provides a new empirical dimension for theories of idioms to model.

Only one extant theory of idiom representation can naturally account for this behavior: the structural requirement approach of Folli & Harley (2007), Stone (2009), Harley & Stone (2013), and Punske & Stone (2014). According to this approach, idioms behave differently based on the amount of functional structure each requires to license its idiomatic interpretation. As such, the hierarchical configuration of the verb phrase provides an independent mechanism by which this subset behavior is generated. Some idioms require minimal structure—just the verb and its object at the  $\sqrt{P}$  level—and are therefore maximally flexible. Other idioms specify much more structure, requiring the functional heads up to VoiceP (and everything below), which results in a highly inflexible (or syntactically “frozen”) idiom. And, I argue, the layers of functional

structure between  $\sqrt{P}$  and VoiceP provide (potential) licensing domains to which idioms could be restricted with intermediate levels of flexibility.

The studies reported in this dissertation provide evidence for four distinct idiom types. The tree in (1) shows each idiom type with its associated functional structure.

(1) Idiom types with functional structure



One strength of this approach lies in the fact that the subset behavior is “built in” to the theory in virtue of the fact that the verb phrase is also hierarchical in nature. Idioms which are licensed higher in the tree are highly restricted; all lower structure is necessitated by the presence of the higher projection. Consider, for instance, VoiceP. Leaving aside FP for the moment<sup>3</sup>, VoiceP is only licensed in structures where  $vP$  is already present. As such, idioms fixed at the VoiceP level will allow no flexibility in either Voice or  $v$ . This results in an idiom that is completely “frozen”, at least with respect to passivization and gerundization. An idiom which is licensed lower in the tree, at the  $vP$  level, will be somewhat freer. This is because  $vP$  is compatible with different kinds of VoiceP heads (e.g. Voice<sub>DO</sub> and Voice<sub>PASSIVE</sub>). However, the  $vP$  is fixed in such idioms, meaning that structures lacking a  $vP$ , such as nominal gerunds, will not allow the

<sup>3</sup> See discussion in Section 4.3.2.

idiomatic interpretation. Finally, idioms at the  $\sqrt{P}$  level will combine freely with different Voice and category-changing heads, and thus allow both passivization and gerundization. In this way, the hierarchical nature of the verb phrase interplays seamlessly with the subset behavior of flexibility observed in idioms.

It is important to emphasize that the observed idiom types correspond to levels of structure that have been independently shown to be relevant for the variations under investigation. For instance, the theory models idioms that cannot undergo passivization as being fixed at the VoiceP level, which is the locus for manipulations of voice (such as passivization). Such an approach not only has the theoretical machinery to account for the data, but it in fact predicts their existence and provides a straightforward account of these particular subset relationships.

In the remainder of this chapter, I discuss some foundational notions about idioms and their behavior. I begin by reviewing the prevailing ideas about what constitutes an idiom (Section 1.2) and what kinds of constraints on idiomaticity have been proposed (Section 1.3), then demonstrating the characteristic flexibility of (certain) idioms (Section 1.4). After that, I end with an overview of the dissertation (Section 1.5).

## **1.2 What exactly *is* an idiom?**

In order to investigate how idioms are mentally represented in the lexicons and grammars of the speakers who use them, it is relevant to consider what kind of object an idiom actually is: Is the class of elements that have ever been called “idioms” a homogenous or heterogeneous one? Although most people have a general intuition about what counts as

an idiom, there is surprisingly little consensus in the theoretical literature. A survey of definitions culled from scholarly sources reveals almost as many conceptions of *idiom* as researchers who have written about the topic, although some patterns do emerge. These definitions run the gamut from ostensive, which define by example (e.g. “all expressions of the same kind as *kick the bucket*, *let the cat out of the bag*, *take the bull by the horns*, *break the ice*, and so on” (Egan 2008:381)) to theory-laden (e.g. “certain lexemes [that]...lend themselves to incorrect decoding” (Makkai 1972:72)). It is somewhere in between these two extremes that we find the most useful, operationizable definitions—those that have found the most traction in the literature as the basis for theoretical models. And yet even here there is little consensus. The most common models adopt one (or more) of the following lines of thinking: (a) idiomaticity is a form of polysemy; (b) idiomaticity is a form of selection; (c) idiomaticity is the locus of confluence between the syntax and the lexicon; or (d) idiomaticity is a lack of compositionality. A representative example of how each type defines idiomaticity is provided in (2) below.

(2) Some working definitions of *idiom*

- (a) An idiom is “a phraseological unit that involves at least two polysemous constituents, and in which there is a reciprocal contextual selection of subsenses.” (Weinreich 1969:42)
- (b) “In these expressions, the combination of the verb and the direct object presents in one form or another an exceptional property, one that has been

taken as arising out of strict subcategorization and/or selectional restrictions.” (Ruwet 1991:171)

- (c) “[I]dioms are the syntactically complex constituents in a language that the semantic component of the grammar treats as lexical items” (Katz 1973:359)
- (d) “What [idioms] have in common is that their meanings are not deducible from the ordinary meanings of their components by the usual rules of compositional semantics.” (Pulman 1993:249)

Although these are all clearly variations on a theme, each emphasizes a different aspect of what is commonly understood as idiomaticity. For Weinreich (1969), idiomaticity is a bidirectional selectional relationship between two independent lexical items. Take, for instance, idiomatic *bury the hatchet* ‘make peace’. For Weinreich, *bury* selects for the subsense of *hatchet* meaning ‘peace’, and *hatchet* selects for the subsense of *bury* meaning ‘make’.<sup>4</sup> Ruwet (1991), on the other hand, sees the same relationship more asymmetrically, as one lexical item imposing restrictions on (i.e. “selecting”) another. Thus, for Ruwet, it is *bury* which selects for a special subsense of *hatchet* in this idiomatic expression, and the idiomatic meaning arises out of this particular selectional relationship. In contrast, Katz (1973) conceives of an idiom as a single lexical item—at least semantically—rather than the semantic composition of separate lexical items.

---

<sup>4</sup> Note that these subsenses are only available in this particular idiomatic context. More on this in Section 1.2.6 below.

Similarly, Pulman (1993) focuses on the (apparent) lack of compositionality in idiomatic expressions. For Katz and Pulman, then, *bury the hatchet* is semantically impenetrable.

Of the many idiom definitions, the type exemplified by (2d) is the most oft-repeated and oft-cited. The concept of compositionality is for many essential to the understanding of idioms. However, even compositionality can be understood in at least two different ways. The traditional notion of bottom-up compositionality holds that in a compositionally interpreted expression, the independently identifiable meanings of the parts combine to establish the meaning of the whole. However, Nunberg, Sag & Wasow (1994) advocate for a view in which there is a kind of post-hoc compositionality even within idioms, whereby the meanings of the parts, although not identifiable outside of the expression, can be established after the meaning of the whole is understood. This latter notion is sometimes also called *analyzability*.

The conceptualization of *idiom* adopted here most closely resembles that expressed by Weinreich in (2a). For reasons that are elucidated below (see, in particular, the discussion of aspectual compositionality in Section 1.3), I reject the dominant view of idioms as wholly noncompositional expressions, as well as the position that idioms are stored lexical items. Idioms are not monolithic expressions; rather, idiomaticity is a conspiracy of polysemous lexical items that have mutually-activated subsenses. When two such items appear together, the idiomatic meaning results.<sup>5</sup>

---

<sup>5</sup> For a more detailed discussion of this view, see Section 1.5 and the references cited there.

Given these many and varied approaches, it should be clear that the notion of *idiom*, which seems like a fairly intuitive concept, is actually quite fraught as soon as it is considered in any degree of technical detail. The remainder of this section highlights several other points of contention to be considered when attempting to operationalize a definition of *idiom*.

### ***1.2.1 Should idiom be defined syntactically or semantically?***

One broad issue in the literature is the degree to which syntactic considerations should come into play in constraining the notion of *idiom*. Few people deny that syntactic “frozenness” is a key property of idiomaticity, yet there is little consensus on precisely how much consideration “frozenness” should be given when considering what “counts” as an idiom.

As an extreme example, Ruwet (1991:176) defines two separate classes, syntactic idioms and semantic idioms. A syntactic idiom is an expression with a conventionalized syntactic idiosyncrasy. For example, in the French idiom *rendre justice à NP* ‘do justice to NP, give NP his due’, the object NP is missing an article where French syntax typically requires one. A semantic idiom is one in which “the meaning of the words does not allow an interpretation of the combination” (Ruwet 1991:176), e.g. *break the ice* ‘relieve tension by starting a conversation’, where the figurative meaning has nothing to do with literal *breaking* or literal *ice*. (The latter of course encompasses the standard notion of noncompositionality discussed above, cf. Katz & Postal (1963), Fraser (1970), Healey (1968).)

Whereas Ruwet separates the notions of semantic and syntactic idioms, Wood (1986) combines them. For her, idioms are both “wholly non-compositional in meaning and wholly non-productive in form” (Wood 1986:2). That is, in addition to being semantically impenetrable, Wood claims that (true) idioms must also be syntactically *fixed* (though not necessarily syntactically *idiosyncratic*, as was important for Ruwet). According to Wood, both compositionality and formal productivity are continua rather than binary absolutes. Instead of imposing an arbitrary point along either continuum at which something qualifies as an idiom, Wood contends that only expressions which are both fully noncompositional and fully nonproductive are “true” idioms. Thus, for Wood, *shoot the bull* is not a true idiom as it has a lexical variant, *shoot the breeze*. In contrast, *fly off the handle* is a true idiom; in addition to being wholly noncompositional, it allows no syntactic or lexical variation.

Wood’s approach is more extreme than most. The majority of researchers in this area acknowledge that “frozenness” is a feature of idiomaticity, but they are reticent to exclude all expressions that are not wholly “frozen” from the domain of idioms (cf. Chafe 1968; Katz & Postal 1963; Weinreich 1969; Fraser 1970; Newmeyer 1974, etc.). I take the latter approach as well. Rather than excluding these expressions, it is the question of how best to model varying degrees of syntactic “frozenness” that is at the heart of this work.



### 1.2.2 *Must idioms be ambiguous?*

The remaining considerations, while less directly relevant for the central thesis of this work, are no less vital to an exhaustive understanding of the notion of *idiom*. For instance, the question of whether an expression must be ambiguous to be an idiom is highly contentious. On the one hand, some researchers argue that an essential property of idioms is their ambiguity between a literal and a figurative interpretation. Others contend that any expression that meets the basic semantic definition qualifies as an idiom.

Taking the former approach, Makkai (1972) and Weinreich (1969) insist that expressions lacking ambiguity do not qualify as idioms. Weinreich addresses cases like *blind alley*, where the literal interpretation is nonsensical and would be ruled out in some semantic frameworks. For him, the literal interpretation is possible but strongly dispreferred. Expressions that are unambiguous because they contain meaningless elements (e.g. binomials like *spic and span*; see discussion in Section 1.2.3) are labeled by Makkai as “pseudo-idioms”. He argues that these are actually strong collocations rather than idioms. In a stable collocation, one constituent predicts the next with a very high probability, but the expression is otherwise semantically transparent (see discussion in Section 1.2.6). Idioms, on the other hand, represent a strong selectional restriction on a polysemous subsense of a particular lexical entry (rather than selecting for a given lexical item). In this way, Weinreich and Makkai are able to restrict the class of idioms to ambiguous expressions (see also Chafe 1968).

Wood observes that ambiguity is a frequent contributor to opacity, but she argues that it is too strong a requirement for idiomaticity. She provides many examples of

expressions that satisfy her definition of idiom but that are not ambiguous (e.g. *by and large*, *hell for leather*, *come a cropper*).

The position adopted here aligns with that of Wood (1986). There is nothing intrinsic in the basic notion of *idiom* that should preclude unambiguous but semantically opaque expressions. In fact, the ambiguity requirement seems to be an extra stipulation and therefore something to be avoided if possible.

### ***1.2.3 Can idioms include meaningless elements?***

As foreshadowed above, closely related to the question of ambiguity is the issue of meaningless elements: Should expressions that contain meaningless elements (variously termed “caboodle items”, “cran-morphs”, or “cranberry morphs”) be included in the domain of idioms?

It should be obvious that those who require ambiguity for idiomaticity must reject idioms with meaningless elements because such expressions are inherently unambiguous. Meaningless elements, such as *gamut* in *run the gamut*, by definition have no interpretation in and of themselves; they are only interpretable in the appropriate surrounding context. Weinreich (1969) and others who reject unambiguous idioms argue that this is not true idiomaticity. For Makkai (1972), such expressions are formally identified as “pseudo-idioms”.

Wood (1986), in sharp contrast, contends that an expression containing a cran-morph must be an idiom in virtue of the fact that the meaning of the expression cannot be compositionally constructed, since no independent meaning can be assigned to

the cran-morph and, therefore, its meaning cannot contribute to the meaning of the larger expression. For similar views, see Chafe (1968), Healey (1968), Higgins (1981), Pulman (1993), and Harley & Stone (2013).

These two positions stand in stark contrast and have clear implications for what kinds of expressions “count” as idioms. In the former case, expressions like *run the gamut* cannot be idioms because they are unambiguous—there is no literal interpretation, only a figurative one; thus they either force the creation of a new category, “pseudo-idiom”, or disappear into another category, such as *collocations* (discussed in Section 1.2.6). In the latter case, these expressions are idiomatic in virtue of the fact that they are not and cannot be wholly compositional, given the presence of a cran-morph. It is the latter position that I adopt in this dissertation.

#### ***1.2.4 Can monomorphemic words or compounds be idioms?***

Another consideration is whether or not words themselves can be idioms, a question which has two parts: Can monomorphemic words be considered idioms, and can compounds be considered idioms?

For those who adhere strictly to a semantic definition of *idiom*, words typically satisfy that definition and are therefore considered idioms. Fraser (1970), for instance, distinguishes three classes of idioms: words (monomorphemic words); *lexical idioms* (polymorphemic words with noncompositional lexical entries, e.g. *kucklehead*); and *phrasal idioms* (polylexemic expressions with more complex constituent structures, e.g. *buy the farm*). Wood (1986) points out that the distinction between compounds and

phrasal idioms is in some cases arbitrary; compare *topsoil*, *top sheet*, *top-heavy*. There is no principled reason why *topsoil* should be considered a word and *top sheet* a phrase. For this reason, she includes compounds in the domain of idioms. Others who share a similar view include Katz & Postal (1963) and Healey (1968).

Taking this view to its logical conclusion, Hockett (1958) includes not just monomorphemic words but all morphemes as idioms.<sup>6</sup> An idiom is any expression whose meaning cannot be deduced from the meanings of its parts and the way those parts are combined. Morphemes (and monomorphemic words) are expressions whose parts have no independently deducible meaning; therefore, they are idioms.

On the other hand, some linguists simply stipulate that words should be excluded from the domain of idioms. Bolinger (1977), for instance, includes a stipulation that idioms must consist of two or more words. Makkai (1972) has a more nuanced view: He excludes monomorphemic expressions from the domain of idioms because, although they satisfy his basic semantic definition, they do not meet his additional requirement that idioms be “misleading” (see further discussion of ambiguity in Section 1.2.2). Compounds, however, can be misleading in this sense and can therefore be idiomatic.

Nunberg (1978) takes a still more subtle view in which some monomorphemic words are considered idioms and others are not. For him, “only uses that are analyzed as in some measure licensed by some other use will qualify as idioms” (Nunberg 1978:207).

---

<sup>6</sup> More specifically, he defines *idiom* as “[a]ny Y, in any occurrence in which it is not a constituent of a larger Y”, where Y is “any grammatical form the meaning of which is not deducible from its structure” (Hockett 1958:172). In other words, any Saussurean sign is an idiom, except when it occurs as a constituent of another idiom.

In other words, idioms are licensed metaphorically, and only metaphorical uses of lexical items constitute idioms. Under this analysis, *nylon* in the sense of ‘stocking’ is idiomatic because it makes reference to the material that (some) stockings are made of.<sup>7</sup> Like Fraser, Nunberg also distinguishes different classes of idioms, although he subsumes monomorphemic and polymorphemic words in a single class (*lexical idioms*) distinct from the class of *phrasal idioms* (e.g. *pay lip service*).

Here, I assume the position from Distributed Morphology that both (apparently) monomorphemic words and compounds can be idioms (Halle & Marantz 1993; Halle & Marantz 1994; Harley 2014a). In Distributed Morphology, words are composed of acategorial roots combined with (often null) categorizing functional heads (e.g. [<sub>NP</sub> Ø [<sub>VP</sub> cat]]) (Marantz 1995). Thus, even words which appear to be monomorphemic are actually phrasal. Furthermore, because there is no separate morphological component, and all word-formation is syntactic, there is no principled difference between compounds and phrases. For researchers who exclude monomorphemic and polymorphemic words from the domain of idioms, the choice is typically one of stipulation rather than one of principle, and I see no compelling case for it.

---

<sup>7</sup> It is not entirely clear how this differs from polysemy, or whether metaphor-induced idiomaticity is intended to be a subset of polysemy in Nunberg’s system. (See discussion of polysemy in Section 1.2.6, as well as a more extensive discussion of Nunberg’s proposal in Section 4.2.5.2.)

### ***1.2.5 Should the class of idioms include larger expressions such as proverbs or clichés?***

The question of whether larger (sentential) expressions with memorized meanings, such as proverbs or clichés, should be considered idioms is also explicitly addressed by some researchers in the literature. Those who view idiomaticity as pervasive in language (e.g. Hockett 1958; Makkai 1972) tend to include these types of expressions, while others do not. Makkai, for instance, classifies as idioms many different kinds of expressions whose conventionalized uses are pragmatically odd when considered literally, such as “institutionalized politeness” (e.g. asking someone “how are you?” as a formal greeting rather than as a contentful question that requires a response). For Makkai, the domain of idioms also includes proverbs, familiar quotations, and some clichés. Jackendoff (1997) has a similar view of idiomaticity. His *Wheel of Fortune* corpus, upon which he bases his model of idiomaticity, is a collection of approximately 600 puzzles from the popular gameshow. The corpus consists of familiar, well-known words and phrases, including proper names, book and movie titles, clichés, and quotations, as well as phrases that are more prototypically idiomatic.

The issue here is really a more fundamental one: Do all conventionalized uses of language fall under the umbrella of idiomaticity, or are there principled distinctions to be made? I suggested in Section 1.2.4 above that the distinction between words and phrases is an arbitrary one, and there is thus no principled reason why words and compounds should not be considered idioms. Here I take the opposite position. Proverbs and clichés retain a level of semantic transparency that is uncharacteristic of idioms. Using non-

native speakers as a litmus test, one could imagine that a speaker hearing the expression *Too many cooks spoil the broth* for the first time could, with some deliberation, determine its intended meaning. Furthermore, proverbs can be interpreted more or less in isolation; this contrasts with idioms—particularly ambiguous ones—which require context to signal which interpretation should be activated.

This is not intended to be an exhaustive argument in favor of excluding proverbs and clichés from the domain of idioms, but rather an overview of the position held here. This position, while important for delimiting the basic category of *idiom* as understood in this work, is not essential for interpreting the remainder of the dissertation.

### ***1.2.6 How can, or should, idioms be distinguished from collocations, polysemy, light verb constructions, and metaphors?***

One larger consideration is how the notion of *idiom* fits with the related notions of *collocation*, *colligation*, *polysemy*, *light verb construction*, and *metaphor*. To varying degrees, these are all concepts at the intersection of figurative language and fixed expressions, and part of the work in this area has been focused on teasing them apart.

The notions of *collocation* and *idiom* are quite closely related and in fact are sometimes conflated in the literature. A collocation is a set of words whose composite meaning is semantically transparent and whose co-occurrence is highly probable. For instance, *foreseeable future* is a collocation because the word *future* has a relatively high probability of following the word *foreseeable*. An idiom, on the other hand, sometimes exhibits the probabilistic properties of a collocation, but always lacks its semantic

transparency. The idiom *bite the bullet* ‘accept something difficult or unpleasant’ shares few features with a collocation; there is very little predictability between the constituents, and the figurative meaning of the whole has little to do with its literal interpretation.

For Weinreich (1969), this is the distinguishing feature between expressions that contain meaningless elements and those that do not (see Section 1.2.3 above). Because a meaningless element appears in only one expression (e.g. *cockles* in *cockles of the heart*), it predicts with certain probability the appearance of the remaining elements of the expression, and (presumably) it can in principle be assigned whatever meaning is necessary to make the expression transparent. In contrast, idiomaticity occurs when the standard meanings of the constituents are not available, resulting in an opaque composite meaning. Weinreich summarizes the position this way: “What Mel’čuk calls ‘stability of collocation’ is thus a high degree of contextual restriction on the selection of a monosemous dictionary entry; what he calls ‘idiomaticity’ is a strong restriction on the selection of a subsense of a polysemous dictionary entry” (Weinreich 1969:44–45). Weinreich’s account seems problematic in that standard composition here would require *of the heart* to contribute its regular literal interpretation and then compose with the meaning assigned to *cockles* (whatever that may be), when in fact it is not clear how the meaning of the expression *cockles of the heart* ‘the core of one’s being’ includes the literal meaning of *heart* (Heidi Harley, p.c.); mutatis mutandis for *run* in *run the gamut*, *span* in *spic and span*, etc.

For Wood (1986), a collocation is “a composite of lexical items with a specialized, but not strictly unpredictable meaning” (Wood 1986:50), e.g. *hard work*,



*heavy drinking, political party*. Thus, under her definition, collocations are slightly more specialized. In highlighting the primary difference between idioms and collocations, she suggests that, “while idioms are completely unpredictable (non-compositional), collocations are at least roughly predictable, although specialized” (Wood 1986:52).

A closely related (and less frequently discussed) phenomenon is *colligation*. A colligation is “a class of collocations with a common word-class for one [or both – MAS] element” (Mitchell 1971:29). In other words, rather than one (or more) words predicting another with a strong probability, as in a collocation, a colligation is less restrictive on the particular elements involved, allowing free substitution within a given word class. Wood (1986) provides the example of {motive verb} + {‘directional’ particle}, which gives us such expressions as *tear across, lope along, race up*, etc. According to Wood, colligations also tend to be more predictable in meaning than collocations.

Perhaps a more difficult distinction to make is that between polysemy and idiomaticity. Weinreich (1969), in fact, proposes (but does not wholeheartedly endorse) a view of idiomaticity whereby it is subsumed under the umbrella of polysemy. A polysemous word is one with multiple (related) subsenses, and Weinreich observes that it is possible for a polysemous meaning to be available only in a specific context (e.g. for *blind*, the meaning ‘companion I’ve never met’ is only available in the context *blind date*). Idioms, then, represent a two-way contextual dependency; each element depends on the other to receive the context-dependent meaning (e.g. for *by heart*, *by* means ‘from’ in the context of *by heart*, and *heart* means ‘memory’ in the context of *by heart*, giving

the overall interpretation ‘from memory’ for the phrase *by heart*).<sup>8</sup> This is quite similar to the approach taken in this dissertation (see also Marantz 2001; Stone 2009; Harley & Stone 2013), whereby idiomaticity is conceived of as a conspiracy of context-dependent special meanings assigned to the content words in an expression.<sup>9</sup>

Gazdar et al. (1985) also appeal to polysemy, but only for a subset of idioms (see Section 4.2.5.2 for a more detailed discussion). For “decomposable” idioms with meanings that intuitively can be distributed among the elements of the expression (e.g. *spill the beans* ‘divulge a secret’, where *spill* roughly corresponds to ‘divulge’ and *the beans* roughly corresponds to ‘a secret’), Gazdar et al.’s partial function approach assigns the meaning ‘divulge’ to the lexical item **spill**” and the meaning ‘secret’ to **beans**”. Partial functions prevent these lexical items from combining with other predicates (e.g. *\*spill the information*). Nunberg, Sag & Wasow (1994) have a similar notion of assigning meaning to the elements of an idiomatic expression. However, unlike Gazdar et al., they do not limit the distribution of idiomatic elements via partial functions or some other formal mechanism but simply rule them out on pragmatic grounds (see also Pulman 1993).

---

<sup>8</sup> Weinreich brings up three issues with the polysemy account of idioms. First, we might want to prohibit bidirectional selections in principle (although they seem to be required in other domains, like phonology). Second, the polysemous “senses” assigned to the words in an idiom are often not semantically related in a way that is desirable (or even required) for polysemy. Finally, idioms are not equivalent to their semantic interpretations—it is impossible, for example, to say *from heart* to mean *by heart*—yet that is essentially how they are represented under the polysemy account.

<sup>9</sup> Marantz (2013) is careful to distinguish this from his cases of semantically-conditioned allosemy, which are subject to a stricter set of locality constraints than idiomaticity.

Importantly for our discussion below, Makkai (1966; 1972) explicitly argues against equating idiomaticity with polysemy, taking issue with Weinreich (1969)’s examples: “Most English adjective + noun lexemic idioms...cannot be assigned such subsenses as one may find in the case of *red herring* where the paraphrase, too, has an adjective followed by a noun and where, therefore, the temptation arises to seek a subsense ‘phony’ for *red* and a subsense ‘issue’ for *herring*” (Makkai 1972:48–49). Makkai uses as a counterexample *white elephant* ‘a possession unwanted by the owner but difficult to dispose of’ and notes that, “If we assign the meaning ‘difficult to get rid of, unwanted’ to the lexeme *white* and ‘property’ to the lexeme *elephant*, we are merely begging the question. We will have assigned arbitrary subsenses to the adjective and the noun in question by taking the ‘subsense’ out of the synchronically arbitrary meaning of the compound *white elephant*” (Makkai 1972:47). Makkai’s criticism amounts to saying that a polysemy account of idiomaticity is merely formal trickery. Assigning an arbitrary chunk of the idiom’s meaning to one element and another arbitrary chunk to another element ignores the intuition that the semantics of these expressions is holistic in some significant way.

Although Weinreich’s account falls prey to this criticism, some others are less susceptible. Accounts like that of Gazdar et al. (1985) and Nunberg, Sag & Wasow (1994), for instance, use a polysemy account only to explain decomposable idioms. In these cases, they claim, there is a “natural” sense in which the meaning of the whole is distributed among the parts. This contrasts with nondecomposable idioms, which have

holistic meanings that cannot be distributed in such a way; presumably they would include *white elephant* among this latter group.

The structural requirement account of Harley & Stone (2013) handles decomposable idioms in a similar way. However, unlike Gazdar et al. (1985) and Nunberg, Sag & Wasow (1994), they extend this account to nondecomposable idioms as well. For such idioms, like *kick the bucket* ‘die’, where there is no natural way to divide the figurative meaning among the constituents of the literal phrase, Harley & Stone postulate that *kick* (optionally) receives the interpretation ‘die’ in the context of *the bucket*; and *the bucket* (optionally) receives a null interpretation in the context of *kick*. By postulating that the verbal element bears the entire meaning, while the nominal element is null, this approach sidesteps (to a certain degree) Makkai’s concern that idiomatic meanings might be arbitrarily distributed in polysemic accounts, although the charge of formal trickery may still apply. I review evidence in Section 1.3, however, that there are some reasons to suspect that the formal properties of the NP *the bucket* do contribute to the interpretation of the whole idiom, even if the lexical content does not.

It is at least possible, then, to conceive of idiomaticity as a subtype of polysemy. Light verb constructions, on the other hand, are better thought of as forming an intersecting class with idioms. Light verb constructions are composed of a semantically bleached verb and an object noun that carries the primary semantic content of the expression (e.g. *take a walk*). In such a case, the process of semantic composition does not work in precisely the usual way; or, put another way, the relationship between the meaning of the whole expression and the meanings of the constituent parts is not as one

would expect in a purely compositional expression which exploited the “normal” meaning of the light verb. Thus, light verb constructions can be seen to overlap with idioms in an interesting way: more transparent light verb constructions (e.g. Persian *kâr kardan* ‘to work’, literally ‘work doing’) are less idiomatic, whereas more opaque light verb constructions (e.g. Persian *be âtaash keshidan* ‘to ignite’, literally ‘to fire pulling’) are more idiomatic.<sup>10</sup> While some linguists (e.g. Wittenberg & Piñango 2011) are unwilling to classify any light verb constructions as idiomatic, others use the term without reticence, particularly for the more opaque expressions (e.g. Folli, Harley & Karimi 2005); and, regardless of their rhetorical stance, most researchers who work on light verb constructions agree that they share some idiosyncratic properties with idiomatic expressions.

Given that light verb constructions are comprised of a semantically bleached verb and a semantically transparent object noun, some classic examples of idioms from the literature are likely better categorized as light verb constructions, e.g. *take advantage of*, *pay heed to*, etc. In these examples, the verb provides the argument structure, but the object noun provides the primary semantic content of the event. Thus, *pay heed* could be paraphrased as ‘heed’, much like the light verb construction *take a bath* could be paraphrased as ‘bathe’.

---

<sup>10</sup> Persian data from Folli, Harley & Karimi (2005). The Persian facts are quite interesting. Ghomeshi & Massam (1994) argue that these light verb constructions are actually a form of noun incorporation. However, Simin Karimi (p.c.) points out that these are not good candidates for incorporation because the nonverbal element is not always nominal; it is often separated from the light verb by one or more other elements; and it can scramble for focus. She suggests that “pseudo-incorporation”, in the sense of Massam (2001), may apply to the more idiomatic light verb constructions in Persian, where a much tighter relationship holds between the verbal and nonverbal elements.

One final distinction that merits comment is that between metaphors and idioms. Idioms are commonly thought to be “frozen” metaphors, that is, metaphorical uses of language that have become conventionalized (see for example Nunberg, Sag & Wasow 1994). However, in observing that idioms differ from metaphors by requiring the presence of specific words (e.g. *spill the beans* vs. *spill the peas*), Svenonius (2005:228) makes the opposite claim: idioms can be used as the basis for metaphors. That is, metaphors are in some sense more general than idioms, and therefore metaphorical extensions can be made from idiomatic expressions (e.g. *beard the lion in his den* ‘confront an adversary on his own territory’ can be metaphorically extended to uses such as *I’ll beard the editor in his den*). Thus, while metaphors surely do sometimes give rise to idioms, the relationship between the two phenomena is not as straightforward as commonly assumed. Makkai (1972:138–139), too, notes that the notions of *metaphor* and *idiom* are distinct. He uses the example of *downfall*, which is metaphorical, involving the ‘falling down’ of abstract things, but is in no way idiomatic, since both *down* and *fall* are interpreted in the usual way. Metaphors and idioms certainly share the property of being figurative; however, to consider them interchangeable is to do both a disservice.

These five categories—*collocation*, *colligation*, *polysemy*, *light verb construction*, and *metaphor*—all relate in varying degrees to the nexus where figurative language and fixed expressions meet. They represent distinct yet sometimes overlapping phenomena. As such, it is important to consider factors that may affect one or all of these categories when interpreting the data in subsequent chapters.

### **1.2.7 Summary**

It is important to establish the empirical domain of idioms before explicating a theoretical model of idiomaticity. In this section, I have provided an overview of many of the ways in which researchers disagree about what “counts” as an idiom. This dissertation side-steps many of these disagreements—e.g. whether the domain of idioms should encompass words (Section 1.2.4) or proverbs (Section 1.2.5)—by taking as its empirical focus phrasal verb-object idioms, which form what most consider to be the core of the class of idioms. However, even these core cases come under scrutiny when considering issues such as ambiguity (Section 1.2.2), meaningless elements (Section 1.2.3), and related verbal phenomena like light verb constructions (Section 1.2.6). In later chapters, particularly Chapter 3, it will be important to keep in mind that these factors may bear on the experimental and corpus results.

## **1.3 What constrains idiomaticity?**

The previous section reviewed some key issues around identifying exactly what types of expressions should be included in or excluded from the domain of idioms. The present section deals with a separate but related line of inquiry, namely whether constraints are imposed on idiomaticity by the grammar, and if so, what those constraints are. Based on empirical observations about general tendencies within the class of idioms, scholars in the field have made various proposals about what kinds of expressions idioms are or can be.

Perlmutter (1970) sets the stage for all subsequent work on idioms in the generative tradition. He conceives of idiomaticity as akin to strong selectional restrictions: “The restriction on the occurrence of these nouns [*recourse*, *heed*, and *headway* - MAS] can be stated as follows: in deep structure they occur *only* in the fixed phrases *have recourse (to)*, *pay heed (to)*, and *make headway*” (Perlmutter 1970:109–110).<sup>11,12</sup> Thus Perlmutter claims that, in spite of sometimes appearing in different configurations, idioms must always start the derivation as contiguous phrasal units. And, because these expressions can be passivized (e.g. *recourse was had to illegal methods*, where the idiomatic verb and its object NP are separated), as pointed out by Chomsky (1970:219–220, fn. 29), this provides “extremely strong evidence that there is a passive transformation in English which takes deep structure objects and makes them into subjects in surface structure.”

This basic premise became the foundation for a fairly substantial body of work, which made arguments in favor of transformations based on the syntactic behavior of idioms and the assumption that they must be contiguous at deep structure. These transformations include raising (Perlmutter 1970:109–110, fn. 7&8; Postal 1974:Chapter 12); *tough*-movement (Berman 1973; Lasnik & Fiengo 1974); and relative clause

---

<sup>11</sup> The implementation in HPSG is quite similar. See, for example, Sag, Wasow & Bender (2003:347–350) and the references cited therein.

<sup>12</sup> The idea that selection plays a major role in idiomaticity is still prevalent in contemporary theories. However, it is worth noting that Perlmutter only discusses idioms that contain cran-morphs. In these cases, it is (relatively) easy to see how selection plays a role: The distribution of the NP object qua cran-morph is restricted to the environment(s) where it is selected by the verb. The situation is more complicated for idioms like *kick the bucket*, where all of the elements have freer distributions; the question of how the idiomatic meaning arises in these cases was left for future development.



formation (Brame 1968, cited in McCawley 1981; Schachter 1973; Vergnaud 1974). Out of the same work came a related idea that became known as the Chomskyan Thesis, the hypothesis that some idioms are allowed only in “untransformed” (i.e. canonical, or deep structure) configurations, and some are allowed in both “untransformed” and “transformed” (or surface structure) configurations (i.e. variations), but no idioms should be possible in only “transformed” configurations to the exclusion of “untransformed” configurations. In other words, the Chomskyan Thesis states that idioms cannot undergo obligatory transformations; they must necessarily surface in the contiguous form found in deep structure.

Several counterexamples to both Perlmutter’s initial condition on idiomaticity and the subsequent Chomskyan Thesis have been presented in the literature. Here I discuss two of the most prominent examples, raised by Higgins (1974; 1981) and McCawley (1981), both drawing on configurations involving relative clauses.

Higgins (1974; 1981) introduces “proxy clauses” as a specific type of relative clause that serves as a proxy for some element contained within the clause itself (what he calls the “proxied” element). Consider, for example, (3) below:

- (3) “Yes, my father can seldom be prevailed on to give the waters what I think a fair trial...” [J. Austen, *Northanger Abbey*, 138, qtd. in Higgins 1981:1]

Here, the clause *what I think a fair trial* serves as a proxy for the NP *a fair trial*. Higgins notes that idioms readily participate in such constructions:

- (4) John certainly isn’t making what I would regard as headway. (Higgins 1974:3)

In such cases, “the object noun phrase associated with the idiom appears within a proxy clause, separated from its verb” (Higgins 1974:3). As such, it cannot be contiguous with the verb at any point in the derivation.

A similar objection has been raised for certain types of relative clauses, most notably by McCawley (1981). Consider the difference between (5) and (6) below.

(5) I’m well aware of the strings that Parky pulled to get me my job. (adapted from McCawley 1981:135)

(6) Parky pulled the strings that got me my job. (adapted from McCawley 1981:137)

In (5), the verb that forms an idiom with the head noun is located within the relative clause, so it is possible to suppose that the two were contiguous at deep structure (see Brame 1968). In (6), however, the idiomatic verb is in the main clause. Here, there was no analysis in the theory of the day under which the idiomatic noun and verb could be said to be contiguous at deep structure, although they are obviously adjacent at surface structure.

The most common proposal for dealing with the proxy clause counterexamples involves expanding the idea of what it means to be contiguous at deep structure. Higgins (1974:11), for instance, suggests (rather informally) that, “[f]rom proxy clauses one can see that a noun phrase like *headway* merely has to be equated to an anaphor which is the object of the corresponding verb; it need not be in the object position itself.” Similarly,

Vergnaud (1985) provides a more formal solution, showing that the noun need only form a chain with the relative pronoun, which can stand in the canonical object position.

To accommodate his own counterexamples, McCawley (1981) proposes two substantially different derivations for relative clauses. These are now known as the “raising” and “matching” structures, and they are widely accepted in more recent work on relativization (see for example Carlson 1977; Sauerland 2003, and the references cited therein).

Even stronger protests have been lodged against the Chomskyan Thesis, the claim that an idiom is allowed in surface structure forms only if it is also allowed in its deep structure form. Ruwet (1991:Chapter 5), in particular, takes Chomsky to task on this claim, providing a systematic dismantling of the Chomskyan Thesis within the framework of the time.

He begins by pointing out that the original formulation of the Chomskyan Thesis is, in fact, incoherent. Chomsky writes:

Thus idioms in general have the formal properties of non-idiomatic structures, and appear either in D-structure or S-structure form, but not only in S-structure or LF-form. D-structure, not S-structure or LF, appears to be the natural place for the operation of idiom rules, since it is only at D-structure that idioms are uniformly not “scattered” and it is only the D-structure forms that always exist for the idiom (with marked exceptions), S-structure sometimes being inaccessible to idiomatic interpretation. Thus at D-structure, idioms can be distinguished as subject or not subject to Move- $\alpha$ , determining the asymmetry just noted. (Chomsky 1981:146, fn. 94)

Ruwet takes issue with Chomsky’s use of “D-structure form”, “S-structure form”, and “LF-form” as proxies for *surface forms* that closely resemble the posited D-structure, S-structure, and LF representations. Furthermore, Ruwet points out that Chomsky’s

parenthetical “with marked exceptions” is so equivocating as to render the remainder of the passage virtually meaningless.

Aside from criticizing the wording, the bulk of Ruwet’s chapter is dedicated to amassing a laundry list of counterexamples to Chomsky’s substantive claim. These counterexamples come primarily in the form of idioms that are acceptable only in a “transformed” configuration but not in an “untransformed” one, and Ruwet makes a point of providing evidence from a variety of transformations. Three of his arguments are outlined below; the reader is referred to his lengthy discussion for a more comprehensive critique.

One piece of evidence comes from unaccusatives. In the analysis of the day, a rule of *there*-insertion would derive sentences like (7a) and (7c) from their counterparts in (7b) and (7d).

(7) Unaccusatives (examples from Ruwet 1991:200)

- (a) There is a fly in the ointment.
- (b) \*A fly is in the ointment.
- (c) There is no smoke without fire.
- (d) \*No smoke is without fire.

Given the data in (7), such an analysis is clearly inconsistent with the Chomskyan Thesis; the (b) and (d) sentences would be “underived”, yet they are unacceptable with the intended idiomatic interpretations. Under more recent analyses (e.g. Burzio 1986), the verb *be* (and others) is considered unaccusative, and its subject is base-generated within

the verb phrase. As such, the (a) and (c) sentences are in fact closer to the underlying D-structure representations, and the Chomskyan Thesis is temporarily rescued.

Ruwet does provide one counterexample to this alternative analysis, given below in (8a)-(8b); the French idiom *un ange passe* appears in the simple active sentence but not in *there*-insertion.

(8) An unaccusative counterexample (from Ruwet 1991:200)

(a) Un ange passe.

An angel passes (said after an uncomfortable silence in a conversation).

(b) \*Il passe un ange.

There passes an angel.

However, he provides no evidence that the French verb *passe* is unaccusative in this use, and this is his only counterexample, so the argument is rather weak. Ruwet concludes that, no matter which analysis of unaccusative verbs is adopted, the Chomskyan Thesis cannot be maintained, but the conclusion from these data alone seems tentative at best.

French prepositional postposing provides a more compelling case. Here, Ruwet shows that, whether one assumes V PP NP or V NP PP as the underlying constituent order in French, it is impossible to explain the behavior of idioms in accordance with the Chomskyan Thesis.

(9) Prepositional postposing (examples from Ruwet 1991:197-198)

- (a) Max a fait d'une pierre deux coups.

Max made with one stone two strokes (i.e. Max killed two birds with one stone).

- (b) \*Max a fait deux coups d'une pierre.

Max made two strokes with one stone.

- (c) Pierre fait flèche de tout bois.

Pierre makes arrow from all wood (i.e. Pierre takes advantage of all opportunities).

- (d) \*Pierre fait de tout bois flèche.

Pierre makes from all wood arrow.

The examples in (9) demonstrate both behaviors. Examples (9a-b) show an idiom that appears only in the V PP NP order but is illicit in the V NP PP order; examples (9c-d) show an idiom with the opposite pattern. Although the V NP PP order is much more common (as in c), several examples of V PP NP order (as in a) can also be found. In order to preserve the Chomskyan Thesis, one must either conclude that these two configurations are not derivationally related, or abandon the Chomskyan Thesis. More recent theoretical developments (e.g. Harley 2003) suggest that distinct underlying configurations may indeed give rise to the V NP PP and V PP NP surface orders, at least in English.

Ruwet also provides examples from relativization, as in (10) below.

(10) Finite relative clauses (examples from Ruwet 1991:214)

- (a) The straw that broke the camel's back.
- (b) \*The straw broke the camel's back.

The fact that (10a) is possible while (10b) is not suggests, again, that the Chomskyan Thesis must be wrong, at least insofar as the finite relative clause has its source in the simple active clause. As with the other counterexamples, however, developments in the theory of relativization provide models for these sentences where the relativized sentence is not derivationally related to the simple active (e.g. Sauerland 2003, among many others).

Ultimately, Ruwet (1991:233) concludes that “[t]he behavior of idiom chunks thus proves nothing, neither for nor against the existence of movement rules; the Chomskyan Thesis must be abandoned.” However, in light of more recent work, we might just as easily conclude that the concept of direct derivational relationships between active and “transformed” sentences must be abandoned—and, indeed, many of them have.

The Chomskyan Thesis figures less prominently in later work<sup>13</sup>, where the focus shifts to refining the notion of a possible idiom. One contentious claim about possible idioms is the No-Agent Idiom hypothesis of Marantz (1984; 1997). Marantz (1984) observes several asymmetries between subjects and objects which lead him to conclude

---

<sup>13</sup> This is perhaps because it loses some of its predictive power in models of the grammar where sentences are not derivationally related in the same way as linguists thought they were in the days of transformational grammar.

that idioms cannot have fixed external arguments with free internal arguments. This position is later refined (Marantz 1997) to exclude idioms with fixed agents, because “the root material in the agent phrase would be getting a special meaning (from the Encyclopedia) in the context of some structure or material below the head that projects the position for the agent—i.e., across a locality barrier for determination of special meaning” (Marantz 1997:209). That is, the agent-projecting head is a boundary across which idiomaticity cannot penetrate.

On the other side of this debate, Nunberg, Sag & Wasow (1994), O’Grady (1998), Horvath & Siloni (2002), and many others have provided what they believe to be counterexamples to Marantz’s generalization. Harley & Stone (2013) address many of these potential counterexamples, showing how they do not hold up under close scrutiny. Many of the counterexamples fall into one of two categories: The subject is not an Agent but rather an Experiencer (or some other role); or the idiom is not a true subject-verb idiom, but instead an idiomatic NP that frequently occurs with a certain verb but freely combines with others. Nonetheless, it is still an open question whether the No Agent Idioms hypothesis is just a tendency or an inviolable constraint.

Building on Marantz’s work, and recalling Perlmutter (1970)’s notion that “idioms tend to correspond to syntactic constituents” (Svenonius 2005:231), Svenonius proposes that phases set limits on what can be listed in the lexicon; in other words, he claims that idioms are phase-bound. As a simple case, he reiterates Marantz’s observation that there are verb phrase idioms and sentential idioms but no subject-verb idioms (that exclude the object). Furthermore, we find idioms with free Goal objects in a



double object structure (*read X the riot act*) and open possessor positions (*cramp X's style*), but none with free Theme objects in a double object structure or possessed nouns. Additionally, verb-particle constructions can be idiomatic (*give up*, *live down*), but verb-preposition combinations cannot (*play with X*, *mess with X* vs. *rely on X*, *count on X*).<sup>14</sup> All of this suggests that there are strict syntactic domains for idioms. There are T-domain idioms (*had better*, *have been*, *as X would have it*) and V-domain idioms (*kick the bucket*), but no idioms that span the two domains.

Svenonius runs into an issue with DPs, however. If the DP is a phase, then idioms like *kick the bucket* appear to span two domains, namely D and V. Svenonius allows for the possibility that DPs are not phases, but also provides an alternative analysis. If “the choice of *the* or *a* is predictable on the basis of other facts about the idiom” (Svenonius 2005:243), then the idiomatic object can be analyzed as a noun rather than a DP. Thus, these idiomatic expressions are stored in the lexicon as structures smaller than a single phase, simple V+N combinations, and the D is merged later. This saves the phase-bound hypothesis but forces Svenonius to introduce a new mechanism, multidominance. When the D is merged as the mother of N, N is already dominated by V. In other words, N will have two mothers. Multidominance has far-reaching implications for linearization, binding, and many other aspects of the grammar.

---

<sup>14</sup> For Svenonius, verb-preposition combinations are not idiomatic but rather the result of syntactic selection. The distinction is a subtle one, but Svenonius appears to be arguing that selection is a (slightly) more general case: “When a verb appears with a particular preposition, then verbs with similar meanings generally also appear with the same preposition” (Svenonius 2005:236); thus *toy with DP*, *play with DP*, and *mess with DP* vs. *depend on DP*, *rely on DP*, and *count on DP*.

Other accounts, less concerned with the identity of DP as a phasal boundary, hearken back to Perlmutter (1970)'s selectional account, whereby the idiomatic object NP must be selected by the verb and, consequently, these two elements must be contiguous at the start of the derivation. The approach proposed by O'Grady (1998) and Bruening (2010), for instance, is similar to Perlmutter's but more developed.

O'Grady formulates the following two constraints on idioms:

- (11) *The Continuity Constraint*: An idiom's component parts must form a chain, where a chain consists of licensing relations between heads. (from O'Grady 1998:284)
- (12) *The Hierarchy Constraint*: Any arguments that are part of a verbal idiom must be lower on the hierarchy than arguments that are not part of the idiom. (from O'Grady 1998:293)<sup>15</sup>

According to the first constraint, an idiom is comprised of units in a selectional relationship, or a series of selectional relationships. For example, verb-object idioms are licit because verbs select their complements. This allows for discontinuous idioms, as unselected material may intervene between the selected components.

---

<sup>15</sup> Bruening (2010) has a constraint similar to (11), but he replaces (12) with the following:

- i. *Constraint on Idiomatic Interpretation*: If X selects a lexical category Y, and X and Y are interpreted idiomatically, all of the selected arguments of Y must be interpreted as part of the idiom that includes X and Y.
- ii. Lexical categories are V, N, A, Adv. (from Bruening 2010:532)

Both constraints are intended to capture the same empirical generalizations.

The latter constraint holds that certain arguments may or may not be part of a verbal idiom, according to a predetermined hierarchy, namely Agent > Theme > Goal/Location. That is, the Theme argument may only be part of an idiom that contains the Goal/Location argument, and the Agent argument may only be idiomatic if the other two arguments are idiomatic as well. This provides a way of capturing the observed subject-object asymmetry in idioms discussed above (Marantz 1984; Marantz 1997; Harley & Stone 2013). O'Grady also claims that the Hierarchy Constraint is strong but violable, allowing for a small number of exceptions.

In contrast to the selectional approach, Koopman & Sportiche (1991) propose a constituency-based constraint on idiomaticity. They propose that idioms are subject to the following structural condition:

(13) Constraint on idiomaticity (Koopman & Sportiche 1991:224)

If X is the minimal constituent containing all the idiomatic material, the head of X is part of the idiom.

Condition (13) says that the head of the smallest constituent that contains all of the idiomatic material must also be idiomatic. Like Svenonius (2005), Koopman & Sportiche propose this constraint to account for the observation that the possible open positions in an idiom are not unbounded. For instance, while possessor positions can be free (e.g. *lose NP's cool*), it is not possible to have an idiom consisting of an adjective and an adverb, with free NP and V positions (e.g. *pale NP slowly V-ed*, to borrow their example). Although they attempt to account for the same empirical data, Koopman &

Sportiche do so with an account that appeals to idiom constituency rather than selectional requirements.<sup>16</sup>

Finally, one of the most salient observations about idiom behavior—and one of the most agreed-upon—is that the functional components of idioms contribute their semantic content to the interpretation of the whole, even when the whole expression is interpreted figuratively. For instance, function words like *the* always operate in the same way as their counterparts in literal expressions. Furthermore, as originally observed by Nunberg (1978) and then developed by McGinnis (2002), an idiom's aspectual class matches that of its nonidiomatic phrasal counterpart. For instance, *kick the bucket* is an achievement (cf. Vendler 1957)—the event is telic, or has an endpoint, and happens instantaneously. Nunberg observes that its idiomatic counterpart is also an achievement. This observation about parallel event structure between figurative and literal readings holds for all verb phrase idioms, a surprising fact if idioms are simply listed lexical items.<sup>17</sup> Earlier work on aspectual composition (e.g. Krifka 1998) showed that a great deal of aspectual categorization in fact depends on the compositional contributions of function words in sentences. Taking that development as a starting point, McGinnis further observes that there are idioms from all four of Vendler's aspectual classes, and

---

<sup>16</sup> For Koopman & Sportiche, subjects must be internal to the verb-phrase to account for sentential idioms like *the shit hit the fan*, where tense is variable (e.g. *the shit will hit the fan*). If the subject were external to the VP, then the minimal constituent containing all idiomatic material would be S (or CP), and tense should be fixed. However, if the subject is VP-internal, then VP is the minimal constituent containing all idiomatic material, and tense can vary freely.

<sup>17</sup> Ruwet (1991) makes a similar argument for French. See also Hamblin & Gibbs Jr. (1999) for experimental evidence that supports this claim. For an opposing viewpoint, see Glasbey (2003; 2007).

that aspect in idioms is compositional in the same way that it is in literal expressions. For example, combining a verb with a count noun results in a telic idiomatic predicate (e.g. *hang a left*), while combining the same verb with a mass noun results in an atelic idiomatic predicate (e.g. *hang fire*), even though the literal meaning of the verb is not accessed. These parallels can be explained if the functional elements make the same semantic contribution, and hence affect event structure composition, in the same way in both literal and figurative interpretations. This observation challenged the prevailing conceptualization of noncompositionality in idioms, showing that they are compositional at least in regard to their aspectual class.

I argue below that this piece of evidence connects to the debate concerning the polysemy of lexical items within idioms (permitting regular composition of irregular lexical meanings; see discussion in Section 1.2.6 above) vs. lexical-item-like views of idioms, where a whole idiomatic string is simply lexically listed with a particular interpretation. The apparent evidence that the aspectual contributions of functional items are computed equally in literal and figurative readings militates in favor of the compositional polysemy view.

Having now provided some perspective on the overall picture of idiom identification and categorization, we focus in on the specific issues that are addressed in this work, in particular the question of idiom flexibility addressed briefly in Section 1.2.1 above, and the question of idiom compositionality, discussed in Section 1.2.6 (on polysemy) and Section 1.3 (on the aspectual interpretation of idioms).

## 1.4 What are some characteristics of idiom behavior?

Idiom flexibility is a blanket term for variations from an idiom’s canonical form. Some idioms are more flexible than others, meaning that they allow more syntactic variations<sup>18</sup>. Other idioms are completely inflexible. This section highlights many of the possible structural variations in which idioms can occur, as well as some prominent methods for investigating these phenomena. Section 1.4.1 focuses on the range of idiom variations. In Section 1.4.2, I discuss experimental investigations of idiom flexibility. Finally, Section 1.4.3 covers corpus-based approaches. The overview of these approaches sets the stage for the remaining chapters, which detail my own experimental and corpus investigations of idiom flexibility.

### 1.4.1 *The range of idiom flexibility*

Because this dissertation focuses on verb-object idioms, the list below concentrates on syntactic operations that can affect the (morpho)syntactic relationships in verb-object structures. Note that not all idioms allow all of these variations; indeed, the reason why some idioms allow certain variations and others do not is the main focus of the remaining chapters.

---

<sup>18</sup> “Variation” is intended to be a theory-neutral term referring to any non-canonical form that idioms—specifically here, verb-object idioms—might occur in (e.g. passivization, object incorporation). Other similar (but theory-laden) terms include “transformation” and “construction”.

One class of variations involves changes to the object NP, including pronominalization of the object (14), changing the number of the object (15), quantification of the object (16), and modification of the object (17).

With certain idioms, a pronoun later in the sentence can refer back to a noun inside an idiomatic expression. In (14) below, we see the pronoun *they* referring back to the idiomatic NP *beans*.

(14) Pronominalization of an idiomatic object NP

Ramya spilled the beans, and it wasn't long before they spread throughout the entire neighborhood. (*spill the beans*)

Another possible variation involves changing the number of the object, as in (15) below, where the canonically singular object NP appears as a plural. (It is also possible for canonically plural NPs to be singular under certain conditions.)

(15) Changing the number of the object NP<sup>19</sup>

Let's go to the club tonight and watch Kenneth bust some moves. (*bust a move*)

---

<sup>19</sup> It seems like changing the number of the object alone is insufficient; in order to license this variation, a quantifier must also be added. This is true even with idioms that have *the* instead of *a*:

- i. We dropped several balls leading up to the party, so it's no surprise that it was a disaster.
- ii. # We dropped the balls leading up to the party.

Sometimes including another variation, in addition to pluralization, is also sufficient to permit the latter, such as passivization in (iii):

- iii. Balls were dropped leading up to the party.

Quantification of the object NP is relatively common, and is often coupled with a change in the number of the object (see footnote 19). In (16), the idiomatic NP is quantified by *many*, though the full range of quantifiers is available.

(16) Quantification of the object NP

Alec turned many heads as he walked down Broadway Street. (*turn heads*)

Adjectival modification of the object is the most complex of these variations. As discussed by Ernst (1981), several types of modification are possible. These types are exemplified in (17a-d).

(17) Adjectival modification of the object NP

- (a) Donna kicked the proverbial bucket. (*kick the bucket*)
- (b) On the first day of class, no one wanted to break the conversational ice, so we sat in silence until the teacher arrived. (*break the ice*)
- (c) People were all too eager to jump on the horse-drawn Reagan bandwagon. (*jump on the bandwagon*; from Ernst 1981:58)
- (d) In spite of the treatment the other refugees received from the rescue party in the desert, he bit his thirst-swollen tongue and kept to himself. (*bite NP's tongue*; from Ernst 1981:59)

---

Heidi Harley (p.c.) suggests that this might have to do with focus. A change to the canonical form of an idiom is only licensed when the change receives sufficient focus, and pluralization may not be salient enough to receive that effect on its own.



External modification (17a-b) is common in idioms but less common in literal expressions. Unlike more standard types of modification, the adjective modifies the entire expression rather than modifying just the noun phrase. In (17a), *proverbial* refers not to a specific kind of bucket but rather signals that the entire expression is being used figuratively. Similarly in (17b), *conversational* describes the context in which the expression *break the ice* should be interpreted; it does not specify a particular kind of ice. Ernst calls these adjectives “domain delimiters”.

Internal modification is similar to standard adjectival modification, although the adjective can be interpreted either literally or figuratively. The clever example from Ernst in (17c) has both types of adjectives: *horse-drawn* modifies the literal meaning of *bandwagon* and must undergo metaphorization in order to be interpreted, while *Reagan* modifies the figurative sense of *bandwagon* evoked by the expression, namely ‘cause or movement’. In both cases, the idiomatic expression is interpreted figuratively, and the adjective contributes to the figurative interpretation.

In contrast, in conjunctive modification, the expression is interpreted both literally and figuratively, and the adjective is taken to modify the literal meaning of the object NP. In (17d), *thirst-swollen* modifies *tongue*, and we understand this to mean that the subject literally has a thirst-swollen tongue; at the same time, we interpret the expression figuratively, with the understanding that the subject is avoiding saying something that might be inappropriate. Of particular interest here is the fact that the expression can simultaneously contribute to both the literal and figurative meaning of the sentence.

Other variations involve changes in the syntactic configuration of elements.

These include passivization (18), subject-to-subject raising (19), tough-movement (20), relativization (21), proxy clause formation (22), cleft sentences (23), topicalization (24), object incorporation (25), and VP-ellipsis (26).

Some idioms retain their idiomatic interpretation in the passive, as in (18) below.

This is one of the most commonly discussed variations; see, for example, Chomsky (1970), Nunberg (1978), Nunberg, Sag & Wasow (1994), Folli & Harley (2007), among many others.

(18) Passivization

The ice was broken as soon as Abby introduced herself. (*break the ice*)

Subject-to-subject raising is another common variation; a representative example is provided in (19). Abeillé (1995) notes that it applies more freely than many other variations. Any idiomatic subject—including a passivized object—can undergo subject-to-subject raising. See also discussion in Perlmutter (1970:109–110, fn. 7&8) and Postal (1974:Chapter 12).

(19) Subject-to-subject raising

The beans seem to have been spilled all over the neighborhood. (*spill the beans*)

Unlike subject-to-subject raising, tough-movement is less freely applicable: some idioms are interpretable in the tough-movement variation, as in (20) below, while some are not

(e.g. #*The ice was easy to break*). More extensive discussion of tough-movement in idioms can be found in Berman (1973), Lasnik & Fiengo (1974), and Ruwet (1991).

(20) Tough-movement

A can of worms is often easier to open than one expects. (*open a can of worms*;  
Ruwet 1991:186)

As with tough-movement, relativization applies only to certain idioms. Furthermore, there are two possibilities for relativization, as discussed in McCawley (1981) and exemplified in (21) below. In (21a), the idiomatic verb is in the relative clause; in (21b), it is in the main clause. These phenomena are discussed in detail in Brame (1968), Schachter (1973), and McCawley (1981); see also Section 1.3 above.

(21) Relativization

(a) I'm well aware of the strings that Parky pulled to get me my job. (*pull strings*; adapted from McCawley 1981:135)

(b) Parky pulled the strings that got me my job. (*pull strings*; adapted from McCawley 1981:137)

Less commonly discussed in the literature are proxy clauses. Proxy clauses are a specific type of relative clause where the clause stands in as a proxy for something else, typically a noun within the clause itself (see discussion in Section 1.3 above). In example (22), the clause *what I would regard as headway* is serving as a proxy for the noun *headway*.

Proxy clauses are discussed extensively by Higgins (1974; 1981) and also addressed by Vergnaud (1985). This variation is only applicable to a small handful of idioms.

(22) Proxy clause formation

John certainly isn't making what I would regard as headway. (*make headway*;  
from Higgins 1981)

Clefting is another variation that is relatively underrepresented in the literature, particularly for English. It is claimed by some to be disallowed for English idioms (e.g. Fraser 1970), though see the related discussion of proxy clauses above, which share similar features with clefting. This variation is much more common for French idioms and is discussed extensively by Ruwet (1991) and Abeillé (1995). Like many variations in English, some French idioms allow clefting while others do not.

(23) Cleft sentences

C'est des monts et (des) merveilles que nous a promis le président. (*promettre monts et merveilles* 'promise the moon' [French]; from Ruwet 1991:192)

It is mountains and marvels that the president has promised us.

Similar to clefting in French, some but not all English idioms can participate in topicalization. In (24) below, the idiomatic object NP is shifted to the topic position at the beginning of the sentence. Further discussion and examples can be found in Wasow, Sag & Nunberg (1984), Gazdar et al. (1985), and Nunberg, Sag & Wasow (1994).

## (24) Topicalization

Those beans not even Tiffany would spill. (*spill the beans*)

Unlike the variations discussed to this point, object incorporation (which could also be called “synthetic compound formation”) is almost entirely unrepresented in the literature.<sup>20</sup> However, it is fairly robust in its application to English idioms. This variation involves at least two changes from the canonical structure: the bare NP object appears in pre-verbal position (without its determiner); and the expression is adjectival rather than verbal (with the verb bearing participial *-ing* morphology). The example in (25) below shows object incorporation with *break the ice*, which appears as the adjectival expression *ice-breaking*.

## (25) Object incorporation

Today’s ice-breaking activities include standard introductions and a game of Two Truths and a Lie. (*break the ice*)

Finally, we see in (26) that part of an idiom can sometimes be elided through VP-ellipsis. As with subject-to-subject raising, this only applies to expressions with idiomatic subjects, or to idiomatic objects that have undergone passivization. Idiomatic VP-ellipsis is discussed by Wasow, Sag & Nunberg (1984), Gazdar et al. (1985), and Nunberg, Sag & Wasow (1994).

---

<sup>20</sup> Stone (to appear) and the references cited therein provide the only discussion that I am aware of.

## (26) VP-ellipsis

There were so many little details required to make our project a success that we worried the ball would be dropped at some point, and sure enough it was. (*drop the ball*)

Two further types of idiomatic variations are discussed by Egan (2008): figurative modification (27) and extendibility (28). Figurative modification is similar to internal modification as in (17c) above, except that the modifier need not be in prenominal position; thus, there is much more flexibility with this type of modification.

## (27) Figurative modification

The strings we've been pulling to keep you out of prison are fraying badly.  
(*pull strings*; from Egan 2008:391)

While *fraying badly* literally modifies *strings* in the above example, it is interpreted figuratively when the idiomatic meaning is activated. If *pull strings* means roughly 'exert influence', then *fraying badly* here indicates that the influence is not having the desired effect. These cases are discussed extensively by Egan (2008).

Also discussed by Egan (2008) is extendibility. Here, the idiomatic interpretation is extended to other parts of the discourse.

## (28) Extendibility

- (a) Brazil is a team with more rabbits in its hat than most. (*pull a rabbit out of X's hat*; Egan 2008:392)
- (b) *Speaker A*: I hear Mr. Jones kicked the bucket.  
*Speaker B*: Yeah. He almost connected yesterday; today he really put the boot on it. (*kick the bucket*; Egan 2008:393)

For example, in (28b), once Speaker A has invoked the idiom *kick the bucket* to mean that Mr. Jones has died, Speaker B uses related terms (*connected*, *boot*) to invoke the same semantic space of death.<sup>21</sup>

Given this abundant variation, it is impossible to sustain a notion of idioms as purely frozen expressions. However, systematic investigations of idiom flexibility have been sparse. One reason for this is that introspection about idioms is quite difficult. The importance of distinguishing literal from idiomatic readings, and semantic from syntactic ill-formedness, has often meant that introspective grammaticality judgments, conducted by experienced linguists who understand these nuances well, have formed the primary source of data about idiom behavior, yet, as noted above, little consensus has emerged. In what follows, I discuss experimental and corpus-based approaches to identifying empirical patterns of idiom flexibility. These approaches avoid many of the difficulties

---

<sup>21</sup> The line between extendibility (and perhaps even figurative modification) and wordplay is tenuous at best. Egan (2008) argues that no such line should be drawn, but I find the examples in (28) to be significantly different from those in (14)-(26).

associated with introspection in this domain, while introducing some new difficulties of their own.

#### ***1.4.2 Experimental approaches to idiom flexibility***

Significant experimental work has been done on idioms, much of it focused on idiom flexibility. The aim of this line of work has generally been to establish a clearer picture of the relationship between idioms and the lexicon.

When compared to literal phrases, or literal interpretations of idiomatic phrases, idioms are generally processed more quickly (Swinney & Cutler 1979; Gibbs 1980; Gibbs Jr. & Gonzales 1985; Gibbs Jr. 1986; Cacciari & Tabossi 1988; McGlone, Glucksberg & Cacciari 1994; Titone & Connine 1994a; Tabossi, Fanari & Wolf 2008). This is a robust finding and has led many researchers to hypothesize that idioms are stored whole, much like large words (Bobrow & Bell 1973; Swinney & Cutler 1979). They conclude that idioms must be accessed in chunks, rather than processed piece by piece, in order to account for the faster processing times. Further evidence for this “words-with-spaces” view comes from the fact that literal uses of idioms are more memorable than figurative uses, suggesting that literal processing requires more effort (Gibbs 1980).

Nonetheless, several experimental findings have challenged the “words-with-spaces” view. For instance, Cacciari & Tabossi (1988) and Titone &



Connine (1994a) found that less predictable idioms<sup>22</sup> are processed more quickly in literal contexts than in figurative ones. That is, the behavior of these idioms is the exact opposite of what is predicted based on the findings reported above. Furthermore, idioms show syntactic priming effects (Peterson et al. 2001). And, in production contexts, subjects are more likely to produce idiom blends<sup>23</sup> when the two idioms share similar syntactic or semantic properties (Cutting & Bock 1997). These results suggest that the “words-with-spaces” hypothesis is too simplistic: there is evidence that, during processing and production, speakers have access to internal syntactic information about the idiom, as well as some aspects of both the literal and figurative meanings. Cacciari & Tabossi (1988) propose the configuration hypothesis, whereby idioms have distributed lexical representations. As such, literal processing can proceed until the string has been identified as idiomatic, at which point the literal parse is abandoned for the figurative one. For highly predictable idioms, this point is early in the idiom; for idioms with low predictability, it is much later. They call this special point in idiom processing the *idiom key*.

A third alternative to both the “words-with-spaces” and the configuration hypotheses is the idiom decomposition hypothesis, sketched above in Section 1.2.6 above and expounded in great detail in Section 4.2.5.2 (Nunberg 1978; Nunberg, Sag & Wasow

---

<sup>22</sup> Predictability in idioms is defined by Titone & Connine (1994b:252) as “the probability of completing an incomplete phrase idiomatically.” *Hit the sack* is an example of an idiom with low predictability. Titone & Connine (1994b) completed norming studies for predictability and several other dimensions of idiomaticity, including familiarity, compositionality, and literality.

<sup>23</sup> An idiom blend is a production error where a word from one idiom replaces a word in another idiom (e.g. *kick the maker* from *kick the bucket* + *meet your maker*; Cutting & Bock 1997).

1994). The idiom decomposition hypothesis explores the idea that some idioms are more flexible than others. A series of studies by Gibbs and his colleagues found that subjects can reliably classify idioms as normally decomposable, abnormally decomposable, and nondecomposable (Gibbs & Nayak 1989; Gibbs et al. 1989; cf. Nunberg 1978). Based on those groupings, they found that normally and abnormally decomposable idioms are processed more quickly than nondecomposable idioms (Gibbs Jr., Nayak & Cutting 1989); normally and abnormally decomposable idioms permit more lexical substitutions than nondecomposable idioms (Gibbs et al. 1989); and normally decomposable idioms are more amenable to syntactic variations than abnormally decomposable or nondecomposable idioms (Gibbs & Nayak 1989). However, subsequent work seeking to replicate those results found no effect of decomposability. Tabossi, Fanari & Wolf (2008) found that subjects were no better than chance at classifying idioms according to decomposability. Furthermore, decomposability was shown to have no effect on semantic priming (Peterson et al. 2001) or blending in idiom production (Cutting & Bock 1997). Finally, Tabossi, Fanari & Wolf (2008) found no effect of decomposability on processing speed.

It is unclear how to interpret these contradictory results, but one possibility is that some of the earlier work used a biased or small sample of idioms, or that the authors conflated decomposability with other factors. For instance, it has been shown that idiom flexibility is correlated with familiarity (McGlone, Glucksberg & Cacciari 1994; Reagan 1987), transparency (Gibbs & Nayak 1989), and (lack of) literality (Gibbs & Nayak 1989). There are also contradictory results regarding Fraser's (1970) Frozenness

Hierarchy, another measure of idiom flexibility (see Section 4.2.3). Swinney & Cutler (1979) found that flexibility did not correlate with an idiom's level of frozenness. However, Reagan (1987) found the opposite result: Fraser's Frozenness Hierarchy corresponded closely with his empirical results.

While there is still much work to be done, these results bring the picture of idiom flexibility a bit more into focus. Flexibility is affected by a complex web of factors, including the speaker's familiarity with the idiom, the transparency of the idiom, and whether or not the idiom has a meaningful literal interpretation. It should be clear, however, that a simple "words-with-spaces" view is insufficient to account for idiom behavior, and that the idiom decomposition hypothesis is also called into question. Regardless of an idiom's decomposability, speakers and hearers have access to information about the idiom's internal syntactic structure, as well as the literal meanings of its constituents.

### ***1.4.3 Corpus work on idiom flexibility***

Laboratory experiments provide one way of approaching the problem of idiom flexibility; another approach involves corpus investigations. While the experiments discussed in the previous section measure subjects' responses to idiomatic stimuli in a controlled environment, corpus studies examine idiomatic usage "in the wild". That is, these investigations look at idioms as they are used in real life.

Compared to the robust experimental work, there is considerably less corpus work on idioms. Nonetheless, the few comprehensive studies that exist have converged

on the fact that, by and large, idioms are more flexible than they are commonly perceived to be (Abeillé 1995; Moon 1998; Fellbaum 2007). The attested flexibility includes lexical substitutions as well as syntactic variations. Furthermore, in her corpus work on French, Abeillé (1995) found no evidence for the idiom decomposition hypothesis. There was no correlation between decomposability and any of the idiom variations she investigated: passivization, NP extraction (clefting and topicalization), quantifier insertion, or relative clause insertion.

As with laboratory experiments, corpus investigations provide useful insights into idiom flexibility. The acknowledgment that idioms are, on the whole, more flexible than previously thought provides good reason for continued investigation into their behavior, and Abeillé's results regarding the idiom decomposition hypothesis are in line with more recent experimental results (e.g. Tabossi, Fanari & Wolf 2008). However, because corpus results can only provide positive evidence for an idiom's participation in a particular structural variation, not many have attempted to use corpora to investigate idiom flexibility. One of the contributions of this dissertation is to show that corpus data can be utilized effectively to inform hypotheses about these and related phenomena.

## **1.5 Conclusion and overview of the dissertation**

The purpose of this dissertation is to explore the role of idioms in a theory of grammar. The model of idiomaticity advocated for here shares some key components with the earliest models in the generative tradition. It preserves the basic intuition of Perlmutter (1970) that idiomaticity is about selectional restrictions among the lexical items that

comprise the idiom. It also embraces the notion from Weinreich (1969) that idiomaticity is akin to polysemy, where the specific use of a lexical item in an idiom is a subsense of its literal use. For idioms where the figurative meaning can be intuitively mapped onto the literal elements of the expression, this is modeled in a way similar to that of the idiom decomposition hypothesis (Nunberg, Sag & Wasow 1994, etc.). For idioms where no such mapping can be done, I suggest that the meaning of the entire expression is associated with the verbal element, and the nominal element has a null subsense in the specific idiomatic context. (See Harley & Stone (2013), Harley (2014a), and Harley (2014b) for a similar view.) In this way, compositionality can be maintained as a component not just of certain “decomposable” idioms but of all verb-object idioms, preserving the important contribution of McGinnis (2002) and others regarding the aspectual composition of functional elements in idioms.

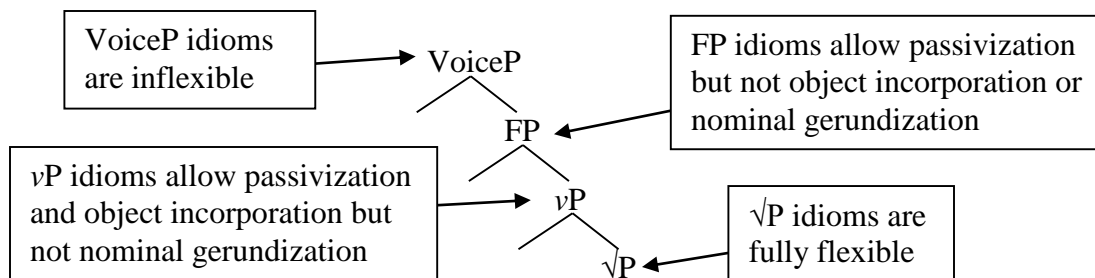
This model of idiomaticity is based on the results of several novel empirical studies, which are reported here. In Chapter 2, I discuss two experiments designed to investigate syntactic flexibility in idioms. The first was a judgment task where subjects were asked to rate the acceptability of idioms in their canonical form (active) as well as two different syntactic variations (passive and nominal gerund). In this study, subjects reported a fundamental misunderstanding of the task, failing to distinguish between syntactic and semantic ill-formedness. Seeking to eliminate the need for subjects to make conscious judgments about the target stimuli, I conducted a second study, a self-paced reading task where reading times were compared for idioms and literal expressions in the active and passive. For both experiments, active sentences were judged better or read

faster than sentences in other syntactic variations. However, beyond that, there was no indication that idioms could be grouped into flexible and inflexible classes—or even passivizable and nonpassivizable classes—on the basis of these experiments.

Chapter 3 turns to corpora as another source of data regarding idiom flexibility. The results of two corpus investigations are reported, one using Google Books and the other using the Corpus of Contemporary American English (Davies 2008-). These investigations examined the usage of idioms in six syntactic variations (passive, topicalization, nominal gerundization, object incorporation, object modification, and object quantification). Although more robust in Google Books, both corpora revealed subset relationships among three of the variations: if an idiom occurred in nominal gerundization, it occurred in object incorporation; and if it occurred in object incorporation, it occurred in passivization.

These subset relationships among idiom variations gives rise to a four-way typology of idioms, introduced in (1) above and repeated here in (29).

(29) Idiom types with functional structure



Chapter 4 details several models of idiomaticity and shows how they fare when faced with this four-way typology. Of the three basic approaches to idiom flexibility—lexical

specification, idiom decomposition, and structural requirement—the structural requirement approach alone can account for this kind of systematic subset behavior.

The lexical specification approach (e.g. Katz 1973) holds that idiom flexibility is a matter of stipulation that must be specified directly in the lexical entry for each individual idiom. This approach predicts that idiom flexibility should be random, not systematic. According to the idiom decomposition approach (e.g. Nunberg, Sag & Wasow 1994), flexible idioms are those for which elements of the figurative meaning can intuitively be mapped onto elements of the literal expression, whereas inflexible idioms are those for which no such mapping holds. As such, idiom decomposition predicts two classes of idioms, flexible and inflexible, with little room for variation within the class of flexible idioms. Finally, the structural requirement approach (e.g. Harley & Stone 2013) proposes that idioms can vary in the amount of functional structure necessary to license an idiom's figurative interpretation. This specification in turn determines which syntactic variations that idiom will participate in. Because syntactic variation is tied to independently motivated functional structure in a hierarchical verb phrase, this account predicts the existence of patterns exactly like the subset behavior observed in Chapter 3. I conclude that the structural requirement approach provides the most viable model of idiom flexibility given the empirical facts.

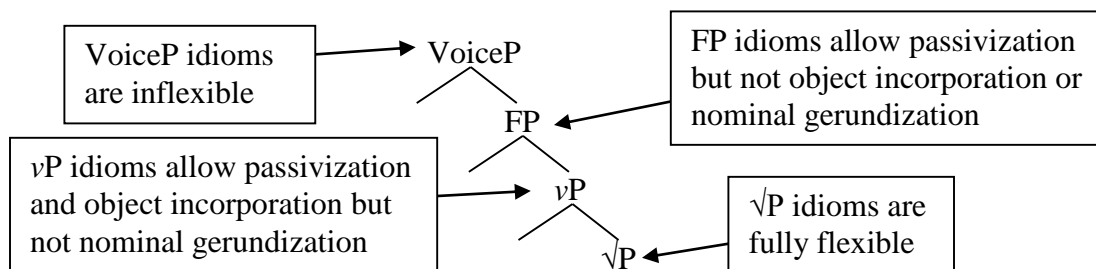
## CHAPTER 2: EXPERIMENTAL INVESTIGATIONS OF IDIOM FLEXIBILITY

### 2.1 Introduction

As mentioned in Chapter 1, the goal of this dissertation is to investigate and account for the range of flexibility exhibited by verb-object idioms in English. The previous chapter provided a detailed exploration of the notion of *idiom* and examined some characteristic idiom flexibility that has been discussed in the literature.

The chapter also provided an overview of the structural requirement model of idiom flexibility (Folli & Harley 2007). Recall that this model posits structural differences among idioms that display different behaviors with regard to flexibility. These structural differences correspond to patterns of flexibility involving three variations—passivization, object incorporation, and nominal gerundization—as shown in the tree in (30) below.

(30) Idiom types with functional structure





Leaving aside object incorporation and its associated projection, FP, for the moment<sup>24</sup>, let us consider some of the ramifications of such a proposal.

This model posits a difference in the functional structure associated with each idiom type. Thus, while *kick the bucket* ‘die’, *steal the show* ‘be the center of attention’, and *turn the tables* ‘reverse a situation’ have identical surface syntax, the structural requirement model contends that they are not uniform beneath the surface. Inflexible *kick the bucket*, a VoiceP idiom, requires a specific (active) Voice head, Voice<sub>DO</sub>, to license its figurative meaning. This particular Voice head is crucially unavailable in the passive and gerundization variations; in the passive, a different Voice head is used, while in the nominal gerund, Voice is absent entirely (e.g. Kratzer 1996). This accounts for the fact that #*The bucket was kicked by John* and #*John’s kicking of the bucket* can both receive only a literal interpretation. In contrast, semi-flexible *steal the show*, a *v*P idiom, only requires the *v* head to license its figurative interpretation, allowing for idiomatic readings in both the active and the passive, since Voice is not a licenser. However, #*Karen’s stealing of the show* is illicit because the nominal gerund does not have a *v*P; it has an *n*P instead. Finally, flexible *turn the tables*, a *√*P idiom, places no restrictions on the functional structure required to license its idiomatic interpretation. It is thus licit in both the passive and the nominal gerund, as well as the active.

Whether an idiom is specified as a VoiceP idiom, a *v*P idiom, or a *√*P idiom is an idiosyncratic property of the individual item. The idiom’s behavior with respect to

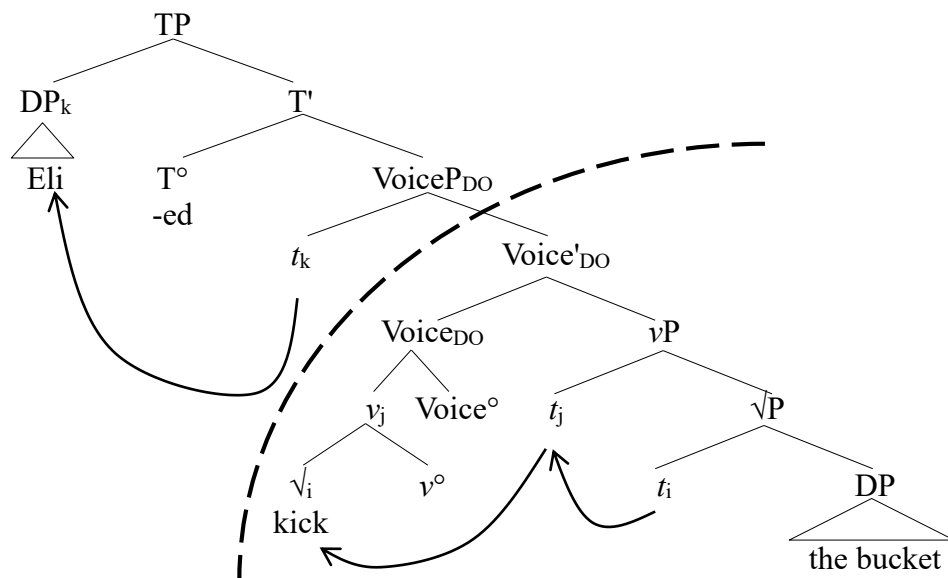
---

<sup>24</sup> Object incorporation is the least-studied of the variations under investigation here, and its relation to verbal functional structure is the most tenuous. I return to this point in some detail in Section 4.3.2.

passivization, object incorporation, and gerundization then falls out from this specification.

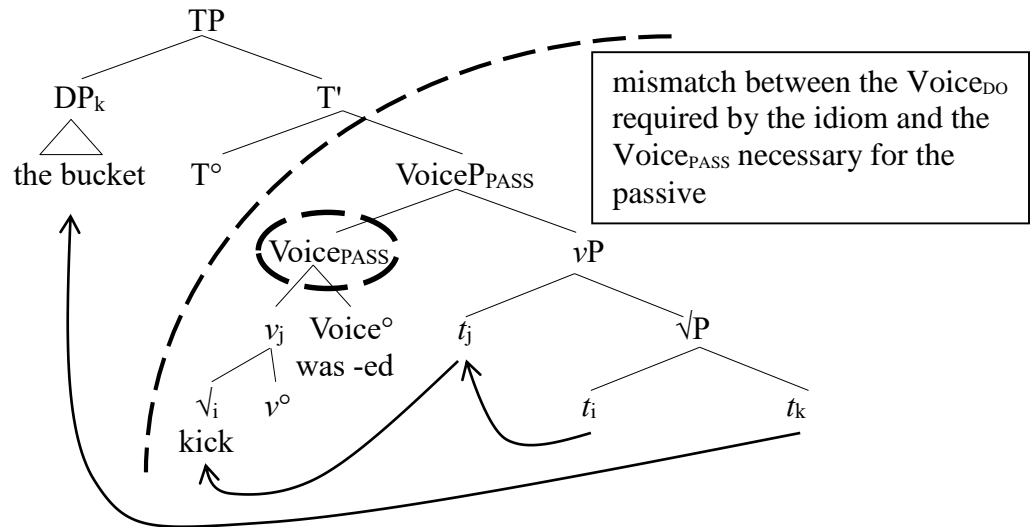
The trees below show the VoiceP idiom *kick the bucket* in the active (31), passive (32), and nominal gerund (33).

(31) VoiceP idiom – *Eli kicked the bucket*



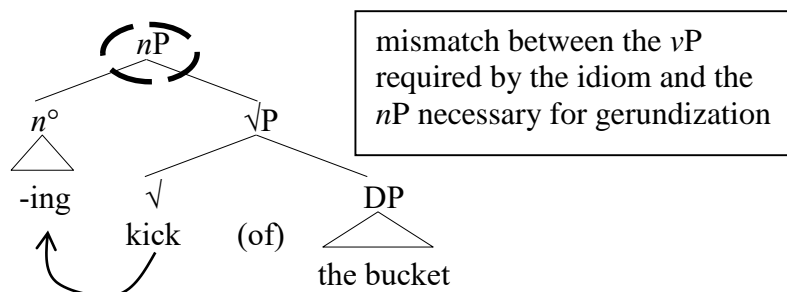
The dashed line in (31) shows the upper boundary of the idiomatic domain for *kick the bucket*; all of the functional structure (and lexical content) below this line must be present in order to license the figurative interpretation.

In the passive variation, VoiceP<sub>DO</sub> is not present; it is replaced with the passive Voice projection, VoiceP<sub>PASS</sub>, as in (32) below. Consequently, the passive structure does not contain the functional material necessary for the figurative interpretation of *kick the bucket*, and the structure can only be interpreted literally.

(32) VoiceP idiom – #*The bucket was kicked*

While *kick the bucket* does not allow passivization, *steal the show* and *turn the tables* both do. This is because both idiom types—vP idioms and √P idioms, respectively—are associated with functional structure that is *lower* in the tree than VoiceP, allowing for free variation at the Voice level.

As a vP idiom, *steal the show* disallows gerundization because it is fixed at the vP level; that means no variation in the categorizing functional head is allowed. As a VoiceP idiom, *kick the bucket* disallows gerundization for a slightly different reason: because this idiom is fixed at the VoiceP level, the functional heads lower in the tree must also be fixed. Thus, vP is fixed simply because VoiceP is fixed. As nominal gerunds are nominal rather than verbal in nature, they have *nP* as their categorizing phrase rather than vP, so the nominal gerundization variation is incompatible with idioms which require vP. This is shown for *kick the bucket* in (33) below.

(33) VoiceP idiom – *#kicking of the bucket*

Because *turn the tables* is a simple root-plus-object idiom, it allows free variation at the level of categorizer, which means that it can receive its idiomatic interpretation in the nominal gerund.

Given that the model posits these structural differences, this chapter explores the predictions of such a model and seeks to test those predictions in a series of experiments. Because the model posits a structural contrast between the variations, it predicts that the distinction between passivizable and nonpassivizable idioms, as well as the distinction between gerundizable and non-gerundizable idioms, should be categorical. However, data on this phenomenon is difficult to acquire through the standard means of grammaticality judgments. First of all, implementing judgment methodology with naïve speakers, while ideal from a scientific perspective, is difficult because it requires subjects to make subtle distinctions, such as differentiating grammatical acceptability from wordplay or coercion. Second, even with trained speakers, judgments do not fall into two neat groups with respect to passivizability or gerundizability. This raises the question of whether individual speakers' uncertainty conceals an essentially bimodal distribution—that is, whether this uncertainty is masking a small but real difference between idioms

that do allow these variations and those that do not—or whether the distinction between these two groups of idioms is a gradient rather than a categorical one.

The experiments discussed in this chapter were designed to test the structural contrast prediction of this model, that is, whether there is a categorical difference between the two groups of idioms. The details of a judgment task, including experimental methodology and results, are outlined in Section 2.2. This experiment failed to yield meaningful results regarding the nature of idiom passivizability and gerundizability due to issues with experimental design. However, it did reveal some interesting patterns of acceptability among the active, passive, and nominal gerund variations more broadly. A follow-up experiment using the self-paced reading task methodology is detailed in Section 2.3. This experiment, too, suffered from design issues that resulted in a lack of meaningful results. Section 2.4 provides a contextualized discussion of the results of both experiments. Ultimately, I conclude that the results of these experiments do not directly bear on the hypothesis under investigation, and I turn in Chapter 3 to corpus investigations as an alternate means of exploring these phenomena.

## **2.2 Experiment One**

In this experiment, subjects provided acceptability judgments for sentences containing idiomatic verb-object phrases in active, passive, and nominal gerund forms. Stimuli consisted of two-sentence scenarios followed by the target idiomatic sentence, which was rated on a 1-7 scale. Results were compared across the three conditions. I begin by

outlining the methodology used in this experiment (Section 2.2.1), followed by the experiment's predictions (Section 2.2.2) and results (Section 2.2.3).

### ***2.2.1 Methodology***

This section describes the experiment's participants (Section 2.2.1.1), materials (Section 2.2.1.2), and procedure (Section 2.2.1.3).

#### *2.2.1.1 Participants*

One hundred forty-two University of Arizona undergraduates from the linguistics experiment pool participated in this experiment for extra credit. All subjects were over 18 years of age; the median age was 20 years. Eighty-six subjects were females and fifty-six were males, and all were native speakers of English.<sup>25</sup>

#### *2.2.1.2 Materials*

The stimuli used in this experiment consisted of 81 English verb-object idioms. These idioms were selected based on their familiarity to the target population of subjects. In order to determine familiarity, an informal idiom familiarity survey was completed by a separate group of 17 native English-speaking University of Arizona undergraduates, none of whom participated in the main experiment. Participants were asked to rate their familiarity with 150 English verb-object idioms on a scale of 1-7, where 1 meant

---

<sup>25</sup> Fifty-seven additional subjects were run after the experiment was complete, allowing them to get course credit. Their data is not included in the results reported here.

“encountered rarely” and 7 meant “encountered often”. All of the idioms used in this experiment had a familiarity score higher than 3. None of the subjects participated in both the familiarity survey and the acceptability experiment detailed below. All idioms consisted of either verb and object or verb, adjective, and object.

Each idiom was presented in three conditions—active, passive, and (nominal) gerund. The study was counterbalanced such that each subject saw an equal number of actives (27), passives (27), and gerunds (27), but each subject saw each idiom exactly once, in only one of the three conditions. A sample stimulus paradigm is provided in (34) below. Each subject saw one and only one of these three stimuli.

(34) Sample stimulus paradigm<sup>26</sup>

**Active:** I heard that Larry cleared the air.

**Passive:** I heard that the air was cleared by Larry.

**Nominal:** I heard about Larry’s clearing of the air.

There were 47 participants in Group A, 48 in Group B, and 47 in Group C. Three pseudorandom orders were used for stimuli presentation.

---

<sup>26</sup> Stimuli include the container phrase “I heard...” because the nominal gerund cannot form a complete sentence on its own: *\*Larry’s clearing of the air*. Embedding idioms from all conditions under “I heard...” ensures that they are all read similarly and at the same place in the stimulus sentence.

### 2.2.1.3 Procedure

The experiment was designed and delivered online using DatStat Illume Version

4.7.1.8639. Potential participants were directed to one of three URLs where they could access the survey. The experiment began with the instructions in (35):

(35) Experiment instructions<sup>27</sup>

The purpose of this experiment is to find out whether some sentences of English sound odd in certain contexts. For each question, you will read a brief paragraph, followed by a sentence in **bold**, which someone might use to describe the situation outlined in the paragraph. You will be asked to rate whether the meaning of the **bold** sentence *immediately* makes sense, given the context provided in the paragraph. You will use a scale of 1 - 7, where 1 means "very odd in this context" and 7 means "perfectly natural in this context".

Consider the following example:

Maria was sitting at home one day. She was bored, so she started looking at random internet pages.

**I heard that Maria surfed the web.**

This example is *perfectly natural*, since you could easily use the sentence in bold to describe the situation in the paragraph.

Now consider the example below:

Maria was sitting at home one day. She was bored, so she started looking at random internet pages.

**I heard that Maria ate a sandwich.**

Even though the sentence in bold is a fine sentence of English, this example is *very odd* because the bold sentence does not describe the paragraph above it.

---

<sup>27</sup> Schütze (1996) provides a good introduction to issues in experimental syntax. For a discussion of the relative merits of scalar versus magnitude estimation methodologies, see Sprouse (2007; 2008), and Weskott & Fanselow (2008), among others.



The subjects were then presented with six practice items, two in each condition (active, passive, or gerund), followed by 81 survey items. A sample survey item is given in (36) below. A complete list of items can be found in Appendix A.

(36) Sample survey item

Joshua was a competitive NCAA athlete. When he was exposed for taking steroids, he had to deal with all of the resulting negative media coverage.

**I heard that the music was faced by Joshua.**

After the 81 survey items, subjects gave a brief summary of how they determined whether the bold sentence was “odd in this context”.

The amount of time it took subjects to complete the 81 survey items was recorded. The minimum time was 2.2 minutes, and the maximum time was 77.9 minutes; the mean time was 19.5 minutes. Five subjects whose completion time was greater than three standard deviations from the mean (greater than 55 minutes) were considered to be outliers and were excluded from further analysis. In addition, ten subjects who took less than seven minutes were also excluded because they did not have sufficient time to read and respond to all 81 survey items. A total of 15 subjects (11%) were excluded in this way, six from Group A, five from Group B, and four from Group C. Following the survey, subjects were asked to identify whether or not they were familiar with each of the idioms. All responses to items where subjects did not know the idiom were discarded (1105 items; 11%).

### 2.2.2 *Predictions*

If the difference between passivizable and nonpassivizable idioms is a structural one, we expect there to be a sharp distinction in judgments between the two. That is, subjects should either rate passivized idioms as good or bad, but nowhere in between. Therefore, the primary prediction of the hypothesis is that there will be a bimodal distribution in subjects' responses to idioms presented in the passive variation.<sup>28</sup>

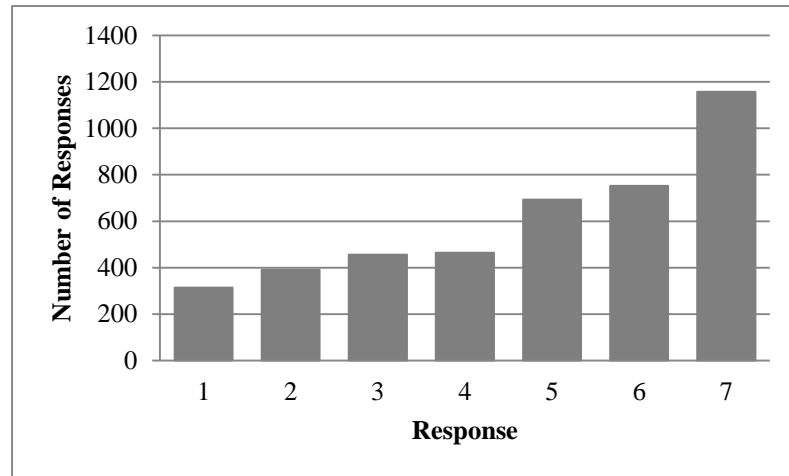
### 2.2.3 *Results*

Recall that the hypothesis predicts a bimodal distribution in the responses to items in the passive condition—that is, some items should be good (the items that, by hypothesis, are not specified for Voice), and other items should be bad (the ones that, by hypothesis, *are* specified for Voice). Looking at the histogram of passive responses in Figure 2.1, there is not an obvious bimodal distribution of the type predicted by the hypothesis in question. That is, there are not two clear modes that would signify two different classes of idioms, idioms which are “good” and “bad” in the passive.<sup>29</sup>

---

<sup>28</sup> The same prediction is true for gerundizable vs. non-gerundizable idioms. Although they are not the focus of the discussion below, the gerund results pattern closely with the passive results.

<sup>29</sup> To confirm this impression, I used Hartigan's Dip Test of Unimodality (Hartigan & Hartigan 1985) to test for a unimodal distribution in responses to passive items. (This measure is best suited for these data because it does not assume a normal distribution.) Somewhat surprisingly, the test returned significance,  $\text{dip} = 0.08893094$ ,  $n = 4228$ ,  $p < .001$ , indicating that there is more than one mode. However, it is not clear where the multimodality in this sample is coming from; certainly it is not the straightforward bimodality predicted by the structural requirement model.



**Figure 2.1: Histogram of responses to Passive items**

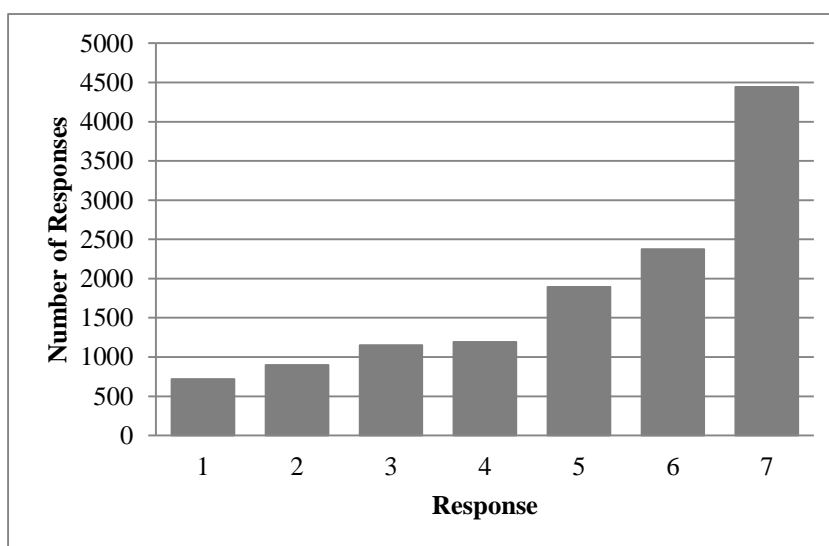
I return to this issue in greater detail in the next section.

#### **2.2.4 Discussion**

As noted, the histogram in Figure 2.1 shows that the expected good/bad bimodal distribution in the passive condition does not hold across these data. Instead, there is a non-normal distribution skewed heavily toward the upper (“good”) range of the scale. Taken at face value, these results contradict the structural requirement hypothesis. However, there are several possible explanations for the discrepancy between the expected and actual results, only some of which directly challenge this hypothesis. For instance, subjects might not have been using the full 7-point scale, or might have been using the scale differently from each other, making it harder for a clear bimodal distribution to emerge. These kinds of subject-specific issues are related to experimental design but have no bearing on the question of which hypothesis best accounts for idiom

flexibility. Another possible explanation for these data is that some idioms provide sharper contrasts between acceptable and unacceptable forms than others. This kind of result would be more difficult for the structural requirement hypothesis to account for. Future work in this domain could help distinguish between these alternatives.

However, there is one possible explanation that I now consider in more depth. The skewing that was seen among the passive items also holds over responses to all three conditions—actives, passives, and nominal gerunds. The histogram in Figure 2.2 shows the total number of each response (1-7) across all three categories (actives, passives, and gerunds). It should be clear that the skewing toward higher scores holds not just for passives, but also when actives and gerunds are included as well.



**Figure 2.2: Histogram of responses to all items**

This tendency for a positive-skewed distribution of responses requires explanation. One possibility is provided by many subjects' responses to the debriefing

question in the experiment. Recall that, after completing the rating portion of the experiment, subjects were asked to provide a brief description of how they decided whether the sentence in bold was “odd” in the given context. It is clear from these responses that subjects were not uniform in how they rated “oddness”. In particular, a large subset of subjects reported ignoring differences in syntactic frame and focusing on the semantic interpretation of the idiom with respect to the context provided:

(37) Selected subject responses to debriefing question

Although every sentence sounded a little odd because it rearranged the wording of the figures of speech i tried to determine whether the figure of speech matched the situation. (subject 35)

I was a little confused as to whether or not I should have been focusing on the syntax or just if the phrase made sense in context, i mostly did the latter but the phrasing was so weird sometimes that it threw me off. (subject 48)

I would try to make sense of the sentence in bold and if it was relevant in any way to the paragraph then I would usually give it a 5 or higher. If I didn't think it was relevant at all then I would give it an “odd”. (subject 55)

I recognized most of the phrases, but some of them didn't fit because they were too extreme. The one I remember the best was when the woman passed away in her sleep peacefully, and the sentence said she kicked the bucket. That seems a little too intense. So, I basically just thought of how much they really contributed to the sentence. (subject 74)

I noticed that all or most of the sentences were forms of expression and rather than basing the validity of the sentence based on its structure which sounded incorrect sometimes I based it on how relevant the expression was to the main sentence. (subject 112)

Although the wording was odd in almost all of them, I focused more on the situation and if it would be social appropriate to say the frase (subject 145)

...with some phrases like “kick the bucket” used when the woman passed away peacefully, I thought that that particular idiom was understandable but not fully

applicable--certain connotations meant that it wasn't a great choice. (subject 152)

Some subjects reported explicitly ignoring structures that sounded weird in favor of rating semantic mismatches (e.g. subjects 35, 48, and 112); some even provided examples of specific semantic mismatches (e.g. subjects 74 and 152).

While the types of responses provided by subjects overall are certainly more varied than this, this sample represents a large subset of all responses. And, insofar as subjects' evaluations of their own responses can be trusted, this can provide some insight into the non-normal distribution of responses seen in the histogram in Figure 2.2.

Because all items were intentionally designed for the idiomatic stimuli to match the context paragraphs, the strong skewing toward the “good” end of the scale is readily explained if subjects were, by and large, simply judging whether or not the idiom was a good fit for the context. Put another way, we can consider these results to show that subjects were essentially rating the experiment design—how well did the experiment designer match the context paragraphs with the idiomatic stimuli? Since this was the actual intention of the designer, the responses unsurprisingly fell largely at the “good” end of the scale.

Although it is impossible to determine precisely why subjects made these kinds of judgments, it is likely that the instructions were a significant contributing factor. The target sentences in the instructions were all literal expressions in the simple active voice. This was an intentional decision meant to avoid priming subjects for idiomatic stimuli or noncanonical variations. However, it seems that it had an unintended consequence,

leading subjects to focus on perceived semantic anomalies while ignoring any syntactic ones.

Based on this evidence, I tentatively conclude that the results of this experiment do not bear directly on the hypothesis under consideration in this chapter. Nonetheless, there is an observable difference between active responses on the one hand and passive and gerund responses on the other (shown in Figure 2.3 below) that seems to support Greenbaum (1976; 1977)'s finding that subjects generally prefer active sentences to passive ones, all other things being equal. I now explore this idea in a bit more detail.

Greenbaum (1976; 1977) had subjects rate closely related active/passive pairs like those in (38) on a 5-tiered acceptability scale from “completely unacceptable” to “perfectly OK”. On a separate occasion, he had the same subjects rate the same sentences on a 5-tiered frequency scale from “very rare” to “very frequent”.

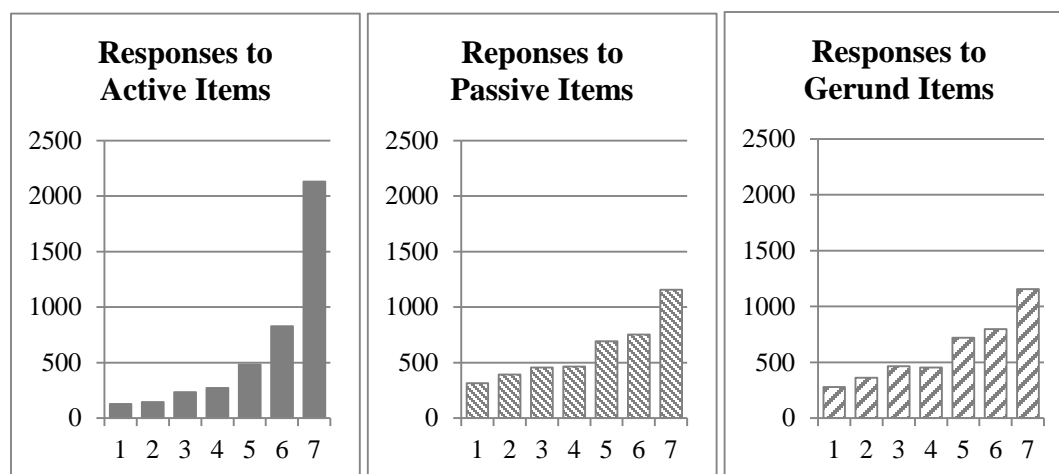
(38) Sample active/passive pair from Greenbaum (1977)

Marvin saw Susan.

Susan was seen by Marvin.

He found that subjects rated active sentences both more acceptable and more frequent than passive sentences, and concluded that acceptability ratings are closely tied to frequency.

Histograms for active, passive, and nominal gerund responses from the current experiment are provided in Figure 2.3. Visually, these results appear to confirm those of Greenbaum (1976; 1977) for actives vs. passives and extend them to actives vs. gerunds.



**Figure 2.3: Histograms of responses to Active, Passive, and Gerund items**

What we see visually is partially confirmed statistically. A Type II ANOVA shows a significant effect of syntax  $F(2, 926) = 5.1034, p < .01$ . Using a pairwise t-test with Holm adjustment to further analyze the differences between the three syntactic configurations, we can see that there is a significant difference between actives and passives ( $p < .01$ ). However, the difference between actives and gerunds is not significant ( $p = .19$ ), nor is that between passives and gerunds ( $p = .19$ ). The means for the three groups are provided in Table 2.1 below.

	Active	Passive	Gerund
Mean	3.70	3.19	3.46

**Table 2.1: Means of responses by Syntax**

These results corroborate Greenbaum (1976; 1977)'s findings concerning the relative acceptability of passives as compared to actives—that is, actives are generally



more acceptable than passives. Interestingly, however, gerunds are somewhere in between such that they are not significantly different from either passives or actives. Based on my own intuitions, I would have expected gerunds to have been categorically less acceptable than passives; the results reported above contradict that expectation and in fact suggest that nominals are *more* acceptable, albeit not to a statistically significant degree. Because frequency was not examined in this study, it remains to be seen whether the constructional frequencies of passives and nominal gerunds are comparable to their relative acceptability, as Greenbaum would predict, or whether some other factor must account for this somewhat surprising result.

One other significant difference between this experiment and that conducted by Greenbaum (1976; 1977) is the linguistic domain of application. Greenbaum's stimuli are all taken from the literal domain, while those used in the present study are idiomatic. The fact that actives and passives are rated similarly in the two domains is not particularly surprising; nonetheless, this experiment provides confirmation for that perception. Additionally, the fact that idiomatic expressions parallel literal ones in this respect provides oblique evidence for the idea that the syntax of idioms is similar to the syntax of literal expressions, contrary to the "words-with-spaces" view (see Section 1.4.2).

In this section, I have shown that the results of this experiment most likely do not bear directly on the hypothesis under consideration, given the odd distribution of results and subjects' responses to the debriefing question. In the next section, I discuss a follow-

up experiment designed to test the same predictions while avoiding some of the pitfalls of the current experiment.

## 2.3 Experiment Two

This experiment compared subjects' reading times on sentences containing idiomatic phrases with literal sentences using a self-paced reading task (Thibadeau, Just & Carpenter 1982; Just, Carpenter & Woolley 1982). The stimuli were biclausal sentences designed to have a garden path effect: the initial clause contained an idiom in either the active or passive, followed in the second clause by a disambiguating context that favored the idiomatic interpretation (cf. Blais & Titone 2008).<sup>30</sup> Subjects are expected to read more quickly when the two clauses are semantically congruent—that is, they make sense together—and more slowly when the two clauses are semantically incongruent. Both passivizable idioms and nonpassivizable idioms were presented in the passive, but we expect that only passivizable idioms should receive an idiomatic interpretation in this variation. The prediction, then, is that nonpassivizable idioms should trigger a reading time slowdown in the passive due to semantic incongruity between the accessed literal interpretation of the first clause and the idiom-requiring context of the second clause. Additional details regarding the methodology used in this experiment (Section 2.3.1), the predictions (Section 2.3.2), and the results (Section 2.3.3) are provided below.

---

<sup>30</sup> Although nominal gerunds were a factor in Experiment 1, they were not included in Experiment 2. This decision was made for two reasons: First, it is difficult to design balanced stimuli that allow actives, passives, and nominal gerunds. Second, the failure of Experiment 1 to yield meaningful results bearing on the hypothesis regarding idiom flexibility led to caution and simplification in the design of Experiment 2.

### ***2.3.1 Methodology***

This section lays out the methodology used in this experiment, starting with a description of the participants (Section 2.3.1.1), followed by a discussion of the materials used (Section 2.3.1.2), and ending with the procedure (Section 2.3.1.3).

#### ***2.3.1.1 Participants***

Seventy-two subjects participated in this experiment. They were recruited from the same linguistics experiment pool at the University of Arizona as those in the first experiment and also received course credit for their participation. None of the subjects participated in the acceptability experiment or the familiarity survey. Seventeen subjects were excluded from the analysis due to computer malfunctions during the experiment procedure. Eight additional subjects were excluded as non-native speakers of English, and ten were excluded for achieving less than 80% accuracy on the comprehension questions included to ensure that subjects were paying attention to the content of the experiment. The issue of accuracy was crucial, since it was essential that subjects were accessing the idiomatic interpretations of phrases where available. Results from the remaining 37 subjects are analyzed in the following sections.

#### ***2.3.1.2 Materials***

The stimuli for this experiment consisted of 64 English verb-object idioms, which are a subset of those used in the previous experiment. Each appeared in four different conditions for a total of 256 unique items. One idiom was discarded due to a

counterbalancing error, leaving 63 idioms and 252 unique items for analysis. The four conditions resulted from crossing two factors, each with two levels, in a 2x2 factorial design. All target items were embedded in a biclausal sentence, where the target item appeared as the predicate of the initial clause with a proper name as the subject (or in a by-phrase for passives), and the second clause consisted of a context for the idiomatic interpretation.

The key issue of direct interest in this experiment is whether the canonical grammaticality judgments about idiom passivizability correlate with performance on a self-paced reading task. More specifically, the hypothesis predicts that idioms which are interpretable in the passive will be read more quickly than idioms which are not. For this reason, one of the factors is Syntax, with two levels: Active and Passive. Each idiom appeared in both the active and the passive voice, as demonstrated for *bite the dust* ‘die’ in (39) below.

(39) Active and passive semantically congruent items for *bite the dust*

- (a) Lucille *bit the dust*, and her family and friends were shocked and upset.
- (b) *The dust was bitten* by Lucille, and her family and friends were shocked and upset.

Note that the second clause is identical across both conditions. This is a crucial feature of the experimental design which is discussed in more detail in Section 2.3.2 when I address the predictions of the experiment.

Idioms containing verbs which are independently known not to passivize were excluded from the experiment. For instance, the verb *have* does not passivize, even when nonidiomatic (#*The book was had by John*), so idioms like *have the munchies* ‘be hungry for a snack’ and *have the last say* ‘have the authority to make decisions’ were excluded.

The second factor controlled in this experiment is semantic Congruence, and the two levels are Congruent and Incongruent. Semantically Congruent stimuli are those where the meaning of the second clause is semantically plausible given some conceptually possible idiomatic interpretation of the first. Semantically Incongruent stimuli, on the other hand, are those where the meaning of the second clause is not plausible given any possible interpretation of the first.

The semantically Incongruent items serve as controls for the Congruent items, providing reading times for implausible items against which plausible items can be measured. The implausible active and implausible passive items for the target stimulus *bite the dust* are provided in (40) below.

(40) Congruent and Incongruent active items for *bite the dust*

- (a) Lucille bit the dust, and her family and friends were shocked and upset.
- (b) Lucille shook the rug, and her family and friends were shocked and upset.

The sentence in (40a) is semantically Congruent because *her family and friends were shocked and upset* is a plausible follow-up to the idiomatic interpretation of *Lucille bit the dust*, where *bite the dust* means ‘die’. Example (40b), in contrast, is implausible,

because *her family and friends were shocked and upset* is an implausible follow-up to the literal clause *Lucille shook the rug*.

Notice that, in order to maintain an identical second clause across all four conditions, enabling direct comparison of reading times for this region of the sentence, the initial clause in the Incongruent condition does not contain the target idiom; instead, it contains a syntactically parallel literal phrase, which is semantically incongruent with the second clause. Again, the justification for this experimental design is discussed in more detail in Section 2.3.2.

When creating implausible (literal) counterparts to idiomatic items, the surface morphosyntactic properties of the clause were matched in every dimension: determiner type (definite, indefinite, or no determiner), number of object (singular, plural, or mass), and presence or absence of a pre-nominal adjective. In addition, the proper name used as the subject or in the passive by-phrase was identical across all four conditions. A representative sample of each of these pairings is given in Table 2.2. All of the stimuli were reviewed by several trained linguists to ensure that they were appropriately congruent or incongruent.

<b>Idiomatic Clause</b>	<b>Incongruent (Literal) Counterpart</b>	<b>Matched Dimension</b>
Victor drew <u>a</u> blank...	Victor stole <u>a</u> hammer...	Indefinite determiner
Jimmy caught <u>fire</u> ...	Jimmy froze <u>water</u> ...	Mass noun object
Jesse drove a <u>hard</u> bargain...	Jesse sliced a <u>ripe</u> pear...	Presence of pre-nominal adjective

**Table 2.2: Idiomatic/literal clause pairings**

A total of 40 fillers were used in the experiment. These paralleled the stimuli just discussed except that they were all semantically congruent items with literal initial clauses. Eight of the fillers had active initial clauses; the remaining 32 were passive. The bias toward passive fillers was intended to offset the likelihood that many of the items in the Passive ~ Congruent condition would actually be semantically incongruent due to the nonpassivizable nature of many idioms. This issue is discussed in more detail under “Predictions” in Section 2.3.2. A complete list of stimuli and fillers is provided in Appendix B.

#### *2.3.1.3 Procedure*

The items were divided into four counterbalanced lists, where each list contained each item exactly one time in one of the four conditions: Active ~ Congruent, Active ~ Incongruent, Passive ~ Congruent, or Passive ~ Incongruent. The order of items in the experiment was pseudo-randomized uniquely for each subject by the experiment software, DMDX (Forster & Forster 2003), the only constraint on randomization being that no more than three items in the same condition could appear consecutively. Each subject was randomly assigned to one of the four lists, such that each subject saw each item exactly once, in one of the four conditions. Each subject also saw an equal number of items in each condition. Fillers were identical across all four lists. Data collection took place in a single session that lasted between 30 and 45 minutes.

Participants were tested in small rooms on the University of Arizona campus. Each subject was seated alone in a room at a desk containing a computer screen and a

standard keyboard. The stimuli were presented visually on the screen, and responses were gathered using the keyboard, which logged both correct and incorrect responses, as well as reaction times, using the DMDX software package.

The experiment was a self-paced reading task where subjects were presented with the biclausal target sentences one clause at a time. Approximately one fifth of the items were followed by a comprehension question to ensure that subjects were reading for content and not just speed.

Participants were instructed, both orally and in writing, to read the sentences as quickly as possible while still understanding the content. For each item, subjects saw two rows of dashes, with spaces and punctuation included, representing the complete sentence. At their own pace, they pressed the right shift key, causing the first clause to replace the top row of dashes. When they finished reading, pressing the right shift key again would bring back dashes in place of the initial clause and bring up the text of the second clause. When finished reading that portion of the sentence, subjects were instructed to press the right shift key a final time. If a comprehension question appeared on the screen, they were instructed to use the left and right shift keys to choose the answer corresponding to the leftmost or rightmost answer on the screen, respectively. Onscreen feedback was provided on the accuracy of their responses so subjects could adjust their reading speed if necessary. Subjects then saw a blank screen where they could briefly pause between items, pressing the space bar when they were ready to begin reading the next item. Each clause or question would remain on the screen for a



maximum of 4000ms, after which the experiment would automatically advance to the next clause or item.

A sample trial block for *bite the dust* in the Active ~ Congruent condition is provided in (41) below, using a monospaced font similar to the one used for the experiment. The hashtags were not in the original presentation but have been used below to indicate subsequent screens, that is, to show where subjects would be required to press a button before progressing to the next screen.

(41) Trial block for Active ~ Congruent *bite the dust*

```

----- -- -- -- --,
-- -- ----- -- -- ----- -- -- ----- .

###

Lucille bit the dust,
-- -- ----- -- -- ----- -- -- ----- .

###

----- -- -- -- --,

and her family and friends were shocked and upset.

###

Who was upset?

Lucille.          Lucille's family.
```

The experiment began with a practice block of 10 items, followed by the experimental block of 104 items (64 experimental items plus 40 fillers). Reading times for each clause were recorded, as were responses to comprehension questions. After the experimental block, subjects completed the familiarity block, where they were asked whether they were familiar with the idioms they had just encountered. Their responses were entered using the left and right shift keys and were also recorded by DMDX.

Individual items were discarded if the subject was unfamiliar with the idiom (5.6%); if the subject timed out at 4000ms (4.2%) or responded in less than 250ms (.59%); or if the reading time was more than two standard deviations outside the subject's mean for that condition (3.0%). The analyses reported in Section 2.3.3 below were run on the remaining data points.

### **2.3.2 Predictions**

Recall that the stimuli for this experiment are biclausal sentences presented one clause at a time. The basic motivation for the experiment is to test readers' underlying grammatical knowledge by measuring reading times in order to avoid some of the pitfalls of more standard grammaticality judgments. The idea is that subjects will read a potentially idiomatic expression in the initial clause and will form an interpretation based on their grammatical competence. When the second clause is presented, they will try to reconcile the (unambiguous) meaning of the new material with the interpretation already established for the initial clause. If the two meanings are semantically congruent, subjects will read the second clause quickly. If there is a mismatch in interpretation,

however, subjects will read the second clause more slowly, being forced to slow down and try to reconcile the meanings of the two clauses. The semantic incongruity effect is well-attested in the literature (Boland 1997; Braze et al. 2002; De Vincenzi et al. 2003).

Given the nature of the semantic incongruity effect, the differences in reading time are expected to appear in the second clause, where subjects are actively (if subconsciously) reconciling the meanings of the two clauses. Because of this, it is important that the second clause be identical across the four conditions to ensure that any measured effect is not actually an effect of clause length, word frequency, or some other unanticipated factor.<sup>31</sup> This also means that all of the reported reading times below are for the second clause.

In this experiment, Congruence serves as a control to ensure that the semantic incongruence effect holds even when idiomatic material is involved. Due to the semantic incongruence effect, semantically congruent items should be read significantly faster than semantically incongruent ones. This condition also provides a baseline for “slow” reading times against which the items in the Passive ~ Congruent condition can be compared.

The Syntax condition is used to test whether there is a difference in processing time between active and passive items. The literature on the processing of actives and passives reveals an asymmetry between online and offline measures. While offline measures such as grammatical well-formedness judgment tasks tend to suggest that

---

<sup>31</sup> Note that this does introduce a potential confounding factor, since the Congruent and Incongruent items also systematically differ on a second dimension, namely whether or not the item contains an idiom in the first clause: All of the Congruent items are idiomatic, while all of the Incongruent items are not.

passives are less acceptable than actives (e.g. McMahon 1963; Mehler 1963; Gough 1966; Gough 1965; Slobin 1966; Forster & Olbrei 1972), passives actually perform better than actives in online measures such as self-paced reading tasks (e.g. Bever & Townsend 1979; Carrithers & Bever 1984; Frazier et al. 1984; Carrithers 1986). For this self-paced reading task, then, passive items should be read more quickly than active ones.

However, the introduction of idiomatic material raises an interesting question: Do idioms behave the same as their literal counterparts with respect to speed of processing in different syntactic frames? In point of fact, the structural requirement hypothesis predicts that many of the items in the Passive ~ Congruent condition will actually be semantically infelicitous due to the unavailability of the idiomatic reading. If this is the case, there will be an interaction between the two conditions, Syntax and Congruence. Unless semantic incongruity completely washes out the effect of syntax, the Incongruent items—because they are all literal—should behave as predicted by previous research, with the Passive items being read more quickly than the Active ones. In the Congruent items, on the other hand, the effect of semantic infelicity caused by the presumed nonpassivizability of some of the idioms should counterbalance the effect of Syntax, and the average reading times for the Passive items will be the same as or slightly slower than the average for Active items.

Finally, recall that the theoretical account posits a structural distinction between passivizable and nonpassivizable idioms. The null hypothesis is that speakers treat idioms uniformly with respect to passivizability; that is, each idiom is either passivizable or nonpassivizable for all speakers. If this distinction actually holds, we expect to see a

categorical difference between the two classes of idioms, which should manifest itself in the reading times for the Passive ~ Congruent condition. It is important to keep in mind that the Congruent condition provides a second clause that is semantically plausible given the *idiomatic* interpretation of the first clause. Thus, passivizable idioms should have quick reading times in the Passive ~ Congruent condition, since the idiomatic interpretation is available in the passive variation. In contrast, nonpassivizable idioms should have slower reading times (resembling those of items in the Passive ~ Incongruent condition), since the idiomatic interpretation is unavailable in the passive variation for these idioms, and the second clause is semantically consistent only with the idiomatic reading. Although the idioms used in this experiment have not been categorized a priori as either passivizable or nonpassivizable, we expect to see a bimodal distribution in the reading times of items in the Passive ~ Congruent condition, where the “fast” items represent passivizable idioms and the “slow” items represent nonpassivizable idioms, respectively. These results can be checked against native speaker intuitions after the fact.

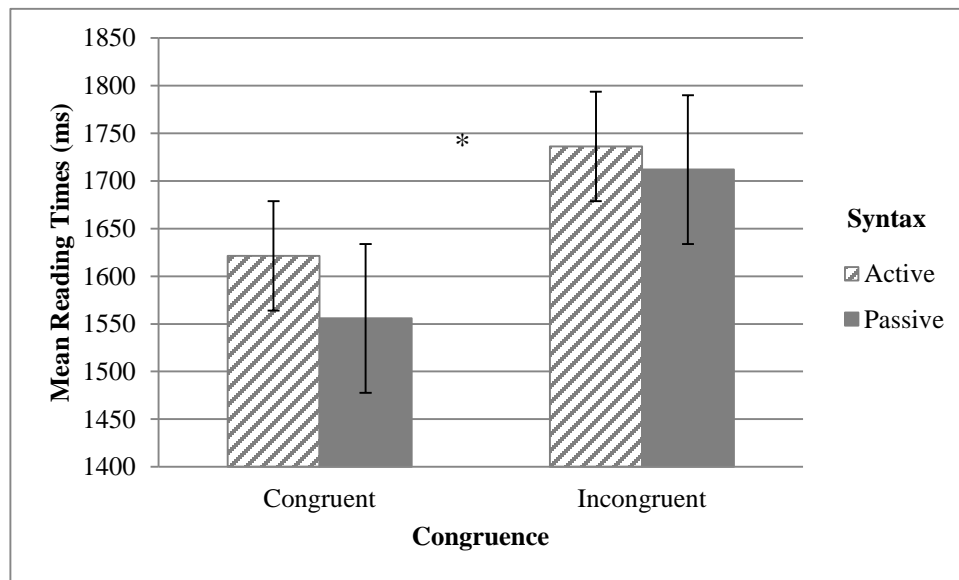
### 2.3.3 *Results*

A between-subjects ANOVA was run with Congruence (2 levels) and Syntax (2 levels) as independent factors and reading time for the second clause as the dependent measure. The means for each condition are given in Table 2.3; these data are represented graphically in Figure 2.4. There was a significant main effect of Congruence both by subjects ( $F_1(1,36)=12.753, p<.05$ ) and by items ( $F_2(1,62)=10.020, p<.05$ ). The main effect of Syntax was not significant, either by subjects ( $F_1(1,36)=1.484, p=.231$ ) or by

items ( $F_2(1,62)=.249$ ,  $p=.619$ ). There was no significant interaction between Congruence and Syntax, either by subjects ( $F_1(1,36)=.393$ ,  $p=.535$ ) or by items ( $F_2(1,62)=.0217$ ,  $p=.883$ ). Participants read the second clause more quickly when its meaning was consistent with the meaning of the initial clause. Active and passive clauses were read with statistically equal speed.

	<b>Congruent</b>	<b>Incongruent</b>
<b>Active</b>	1621.35	1736.16
<b>Passive</b>	1555.84	1711.97

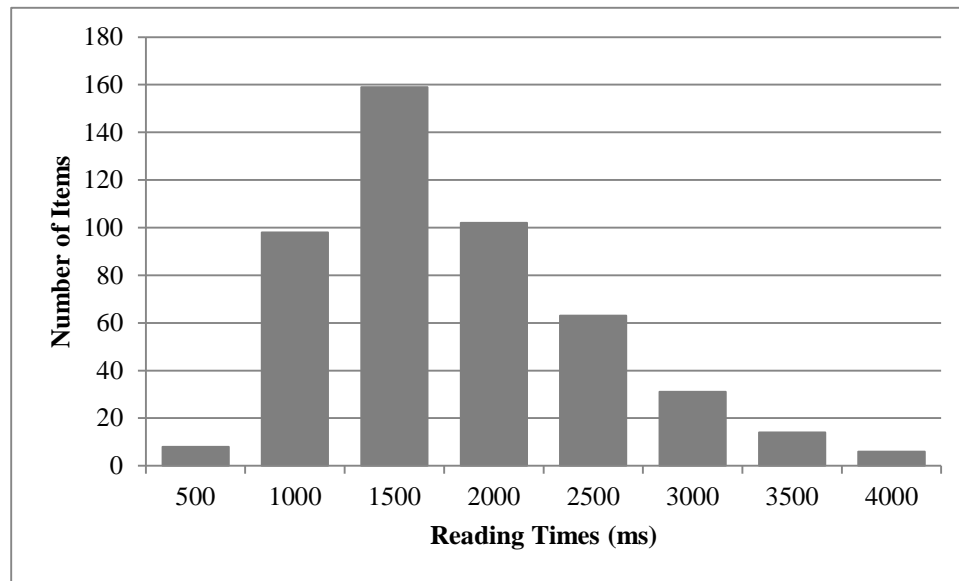
**Table 2.3: Mean reading times for all items in milliseconds**



**Figure 2.4: Mean reading times for Congruent and Incongruent items**

Here, the expected bimodal distribution should be found among the Passive ~ Congruent items; however, as with Experiment 1, no such distribution is present in the

data. The histogram in Figure 2.5 below, which shows the reading times for the Passive ~ Congruent items, reveals a normal curve with a slightly positive skew, showing no indication of bimodality.<sup>32</sup>



**Figure 2.5: Histogram of responses to Passive ~ Congruent items**

The following section provides a more detailed discussion of these results.

### **2.3.4 Discussion**

The results from the previous section yield several points of discussion. Overall, subjects took longer to read semantically incongruent items than congruent ones, but they read active and passive sentences with about equal speed. Furthermore, the distribution of items in the Passive ~ Congruent condition was clearly unimodal, suggesting that there is

---

<sup>32</sup> I ran Hartigan's Dip Test for Unimodality (Hartigan & Hartigan 1985) on the Passive ~ Congruent items, and the results of the test confirm that the distribution is unimodal ( $D=.0098$ ,  $p=.9932$ ).

no clear distinction on this task between passivizable and nonpassivizable idioms. Each of these findings is elaborated in this section, with specific reference to the predictions outlined earlier.

The only statistically significant finding reported above is the main effect of Congruence on reading time. This indicates that the present experimental design is effectively capturing the well-attested effect of semantic incongruity in reading times, where semantically Congruent sentences are read more quickly than Incongruent ones. The novel aspect of this finding is that the difference holds across idiomatic and literal items. Because all of the Congruent items were also idiomatic, while all of the Incongruent items were literal, it was initially unclear whether that would affect the semantic congruity effect. These findings show that the effect of meaning is likely still playing a role in these items. Because idioms are typically processed more quickly than literal speech (Gibbs 1980; McGlone, Glucksberg & Cacciari 1994; Ortony et al. 1978; Swinney & Cutler 1979), it is also possible that the reading time of the first clause is having a carry-over effect on the reading time of the second clause.

Somewhat surprising was the lack of main effect for Syntax. Although the mean reading times for Active items were faster than Passive items in both the Congruent and Incongruent conditions, this effect was not even trending toward significance, as indicated by the reported *p*-values. This is especially interesting in the Incongruent condition, where previous research suggests that passives are read more quickly than actives in online measures. Two possible explanations come to mind for the lack of effect. One is that the slowing effect of semantic incongruity is strong enough to cover



up any facilitation effect of syntax; that is, the negative semantic effect might wash out the syntactic benefit. Another possible explanation is that the effect of syntax takes place in the initial clause, which actually contains the syntactic difference, rather than in the second clause.

Within the Congruent condition, the prediction was much less clear with respect to whether Syntax would be significant, since the anticipated bimodal distribution within the Passive ~ Congruent items was expected to affect the mean reading time in that condition. The fact that no bimodal distribution was found, coupled with the apparent parallelism between items in the Congruent and Incongruent conditions, suggests that whatever mechanism accounts for the lack of significant effect of Syntax in the Incongruent items is likely also responsible for the lack of effect in the Congruent items.

Perhaps the most surprising finding is the lack of bimodal distribution in the Passive ~ Congruent condition. This result contradicts the structural requirement hypothesis of Folli & Harley (2007). However, the result also runs contrary to native speaker intuitions that some idioms are clearly passivizable while others are not. The distribution of items in the Passive ~ Congruent condition suggests that all idioms were treated equally in this experiment with respect to passivizability. As a case in point, a more fine-grained analysis reveals that there is no statistical difference in the mean reading times for the canonically nonpassivizable idiom *kick the bucket* (two-sample  $t(7.797)=0.507$ ,  $p=.626$ ) or the canonically passivizable idiom *spill the beans* (two-sample  $t(10.926)=-.257$ ,  $p=.802$ ) across all four conditions. The mean RTs for each condition are shown in Table 2.4 below. If faster reading times correlated with the

established acceptability judgments, passive *kick the bucket* should be slower than active *kick the bucket* due to the reported unacceptability of *#the bucket was kicked*. For *spill the beans*, the picture is slightly less clear; the two might be expected to be roughly equal, or the passive might be faster, given the known tendency for passives to be processed more quickly in online measures.

	<i>kick the bucket</i>	<i>spill the beans</i>
<b>Active ~ Congruent</b>	1689.95	1873.54
<b>Passive ~ Congruent</b>	1510.32	1979.18

**Table 2.4: Mean reading times for two individual items in milliseconds**

A new question now arises, which is how to reconcile the data presented here with the offline grammaticality judgments of idiom passivizability. It does not seem wise to discard either; on the one hand, the canonical judgments are well-attested in the literature (e.g. Darnell 1972; Palmer 1981; Huddleston 1984; McGinnis 2002), while on the other hand the Congruence effect shows that subjects are at least performing the basic experimental task appropriately. Instead, it seems reasonable to conclude that the reading times reported here are measuring some other kind of acceptability rather than the grammaticality of idiomatic readings. One possible explanation is that subjects are doing a kind of low-level syntactic and semantic processing during the online task, and passivized idioms are somehow able to “escape” that initial pass without being recognized as ungrammatical (Townsend & Bever 2001). This would suggest that their syntactic and semantic form is acceptable at some basic level, allowing these phrases through the first pass. They are only ruled out at a later stage of interpretation where

more fine-grained checking of the derivation takes place. This is consistent with Townsend and Bever's model of the grammar, where "we 'understand' everything twice" (153).

Another possible explanation is that there is a satiation effect (Kounios, Kotz & Holcomb 2000; Snyder 2000). It is possible that subjects are getting enough exposure to idiomatic items, or to the passive variation, or to both, that items become felicitous as the experiment progresses. Testing this account would require considering the order of presentation of items as a factor in the regression model; if order of presentation significantly affects subjects' reaction times, it would be reasonable to conclude that subjects were experiencing a satiation effect due to repeated exposure to similar items.

## **2.4 Conclusion**

The experiments reported here sought to test the structural requirement hypothesis of Folli & Harley (2007) regarding idiom passivizability. Experiment 1 was an acceptability judgment task that asked speakers to evaluate idioms in the active, passive, and nominal gerund on a 1-7 scale. Experiment 2 was a self-paced reading task designed to examine speakers' underlying grammatical knowledge by measuring their reading times on the second clause of biclausal sentences containing active and passive idioms.

Taken at face value, the results of both experiments fail to confirm the structural requirement model of idiom flexibility (Folli & Harley 2007), but a closer look revealed that these experiments might not be bearing directly on the question at hand. In Experiment 1, subjects reported ignoring the syntactic frame and responding to semantic

felicity when providing judgments. In Experiment 2, the lack of reading time differences for canonically nonpassivizable idioms suggests that the task might not be providing access to the stage of grammatical processing that determines idiomaticity.

In the next chapter, I turn to a corpus study examining actual usage of idioms in noncanonical forms, such as passives and nominal gerunds. This corpus investigation is able to expand the range of variations under examination and to provide another perspective on the issue of idiom flexibility.

## **CHAPTER 3: CORPUS INVESTIGATIONS OF IDIOM FLEXIBILITY**

### **3.1 Introduction**

In this chapter, I discuss the results of two corpus investigations designed to provide insight into the types of variations that idioms exhibit in naturally occurring speech. The studies, which examined the behavior of verb-object idioms in Google Books and the Corpus of Contemporary American English (COCA), found patterns of systematic variation in the flexibility of these idioms. These results support the structural requirement hypothesis of Folli & Harley (2007), sketched in the previous two chapters, which predicts that idioms can vary in the types of variations they allow, but that this variability should be systematic in precisely the way found here.

These corpus studies build on the experiments reported in the previous chapter by contributing a new kind of data for consideration. Corpora provide access to naturally occurring language—whether in the form of spontaneous speech or, in the case of Google Books and other written corpora, more carefully crafted written text. Using corpora as a means of gaining insight into linguistic patterns is an especially sensible step when other types of data are more difficult to access. With idioms, this seems to be the case. Consulting one's own intuitions often results in more confusion than certainty; this is definitely true for me and many speakers I have consulted during the course of this research. And, as evidenced by the results reported in Chapter 2, even carefully designed experimental protocols are not an easy or reliable way either to elicit acceptability judgments or to gain indirect access to native speakers' underlying grammatical

knowledge of idiomatic phenomena. Given these considerations, corpora are a useful alternative, bypassing the elicitation of judgments and focusing on “real world” usage.

Corpora have been criticized as a data source in some linguistic traditions, both for their lack of comprehensiveness and for the possible inclusion of ungrammatical examples. This is most notable in the tradition following Chomsky (1965), which highlights the difference between *competence*, a speaker’s tacit knowledge of language, and *performance*, the speech acts resulting from a speaker’s *competence*; for Chomsky, it is the former which is the object of study for linguists, and corpora, he contends, tell us only about the latter. Furthermore, targeting a specific population of speakers is difficult without a limited corpus, since it is impossible to eliminate examples produced by non-native speakers or speakers of certain dialects. This can sometimes result in nonstandard or even ungrammatical sentences in the corpus. Furthermore, and perhaps more importantly, the absence of an exemplar of a given phenomenon in a corpus does not necessarily mean that that phenomenon is ungrammatical; it simply means that the phenomenon is unattested in that corpus. It could, instead, be too infrequent for the size of the corpus, be uncommon in the particular register of the corpus, or simply be an accidental gap.

However, a corpus search can provide several benefits over traditional acceptability judgments, particularly in situations where such judgments have been shown to be difficult to elicit. First, acceptability judgments often force individual speakers to consider very unnatural or awkward sentences, whereas a corpus contains exemplars of actual production. Second, acceptability judgments are often given out of

context, or with contrived contexts. Eliciting judgments of this type requires trained native speakers who are familiar with the task. In contrast, corpora contain natural contexts where the phenomena under consideration actually occur. Finally, acceptability judgments are often limited in number, being constrained by resources such as a finite set of participants and a finite amount of time; conversely, corpora can be quite large, with billions of words of potential examples provided by a very large and (potentially) diverse pool of speakers.

Using corpus data to *supplement*, rather than replace, acceptability judgments and linguistic intuitions represents the best of both worlds. It allows the researcher to make generalizations based on what is attested in the corpus, and then test those generalizations against native speaker intuitions. Conversely, the corpus can be used to support hypotheses by producing examples of predicted grammatical structures, or to contradict hypotheses by providing examples of unpredicted ones. Furthermore, with a bit of manual work, ungrammatical or non-native examples can often be recognized as such and eliminated rather than being accepted as part of the dataset.<sup>33</sup>

The present study highlights the interplay between corpus investigations and theoretical modelling: Using the Google Books and COCA corpora, I test the predictions of the structural requirement model of idiom flexibility. The purpose of this investigation

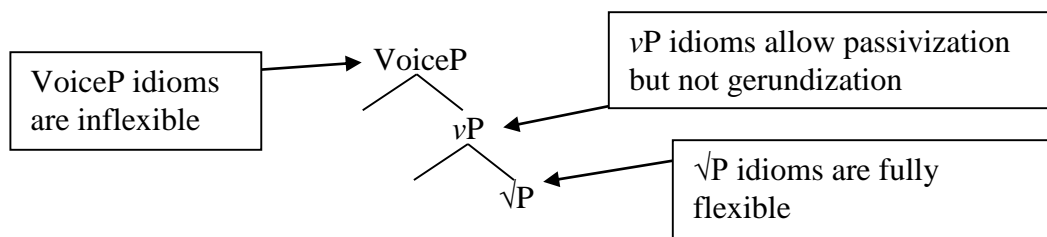
---

<sup>33</sup> Mike Hammond (p.c.) points out that elicited data and corpus data may differ for theoretical reasons that are related to the medium; thus, it is not always fair to draw conclusions about the phenomenon under investigation based solely on one type of data or the other. In fact, the point of this dissertation is to explore the same phenomenon from multiple perspectives—theoretical, experimental, and corpus-based—and look for points of convergence. Such convergence suggests that the findings reported here are not mere coincidence or an artefact of the chosen media but are truly of theoretical significance, though a more thorough investigation would be welcome.

was to determine how observed patterns of idiom flexibility match up against the patterns predicted by this hypothesis. Below, I briefly review those predictions before discussing the study in detail.

As discussed in Chapter 2, the structural requirement model of Folli & Harley (2007) posits structural differences among idioms to account for their different behaviors. As such, we expect to see several idiom types based on these structural differences. Furthermore, because the idiom types are associated with established functional projections arranged hierarchically in the verb phrase, as shown in (1) below, we can speculate about what kinds of patterns might arise.

(42) Idiom types with functional structure



Noncanonical syntactic variations are associated with certain functional projections (e.g. passivization is associated with VoiceP). In order for a variation to apply to a given idiom, the idiom must admit modifications to the projection associated with that variation. Thus, an idiom that is fixed with respect to a certain functional projection will *not* allow the variation associated with that projection. Given that projections lower in the tree are “nested” below higher projections, this model then makes the following prediction: Idioms that allow variations at a lower level will necessarily allow variations at all higher levels as well. For instance, an idiom that



allows nominal gerundization, which is associated with  $\nu$ P, will necessarily allow passivization, associated with VoiceP.<sup>34</sup>

This predicts at least three types of idioms, shown in the tree in (42) and elaborated in (43) below.

(43) Predicted idiom types

- (a) **VoiceP idioms:** Fixed at the VoiceP level, these idioms allow neither passivization nor gerundization.
- (b)  **$\nu$ P idioms:** Fixed at the  $\nu$ P level, these idioms are passivizable but not gerundizable.
- (c)  **$\sqrt{\text{P}}$  idioms:** True root-plus-object idioms, these idioms allow both passivization and gerundization.

This should manifest itself as subset behavior in the corpus results: The set of idioms that occur in gerundization should be a subset of the set of idioms that occur in passivization. In contrast, if the set of gerundizable idioms has little overlap with the set of passivizable idioms, or if the set of gerundizable idioms is a superset of the set of passivizable idioms, those results would be surprising given this model.

These three predicted types are based on variations that are strongly associated with the verbal functional hierarchy (see Section 4.3 for a detailed discussion). Given the amount of research currently focused on the extended verbal domain, additional

---

<sup>34</sup> Some researchers have previously noted this subset behavior, e.g. Fraser (1970), Punske (2012), and Bruening (2014).

functional projections are likely to be identified as potential loci for idiomaticity by subsequent research.<sup>35</sup> Nonetheless, the corpus studies reported below also include several additional variations which are not (clearly) associated with verbal functional structure. The predictions of the structural requirement model with respect to these variations are less clear, but we do not expect to find the same subset relationships among variations that are not tied to nested functional projections.

The remainder of the chapter is structured as follows: Section 3.2 introduces previous corpus-based approaches to the study of idioms. I then present two corpus studies of my own—one using Google Books in Section 3.3 and one using COCA in Section 3.4. Each of these sections includes a breakdown of my methodology, including a discussion of the corpus used, variations investigated, and search procedure followed, as well as a summary of the results. Section 3.5 provides a detailed discussion of the results of both studies. The chapter ends with some brief concluding remarks in Section 3.6.

## **3.2 Previous work**

Surprisingly little corpus work has been done examining idiom flexibility. The following discussion highlights two such large-scale corpus-based investigations, which provide some context for the studies conducted for this dissertation.

The work of Moon (1998) is a corpus investigation of the flexibility of “Fixed Expressions and Idioms” (FEIs) in British English. Using the Oxford Hector Pilot

---

<sup>35</sup> See discussion in Section 4.3.2 for one such possibility.

Corpus with just over 18 million words, Moon searched for 6776 such FEIs, noting for each one its semantic, syntactic, and lexical features, including whether or not the idiom occurred in any variations (or “transformations”). Significantly, Moon notes that “Fixedness is a key property of FEIs, yet around 40% of database FEIs have lexical variations or strongly institutionalized transformations, and around 14% have two or more variations on their canonical forms” (120). In other words, Moon found that almost half of the idioms she investigated were flexible in some way; this is surprising given the fact that idioms are commonly thought to be “fixed” or “frozen”. (See discussion in Section 1.2.1.)

Christiane Fellbaum and her colleagues (Fellbaum 2002; Fellbaum et al. 2006; Fellbaum 2007) completed a similar large-scale corpus investigation of German idioms. Similar to the current study but much larger in scope, Fellbaum used the 1 billion-word *Digitales Woerterbuch der Deutschen Sprache* corpus to search for all (potential) occurrences of an idiomatic expression in any form, then catalogued the ways in which those expressions deviated from their canonical forms. This type of study is very comprehensive in that it captures—in principle—every occurrence of an idiom, rather than being limited by a specific search or the searcher’s imagination. Fellbaum’s project was largely automated, though humans still had to sort through the results to determine whether each was idiomatic or literal. Fellbaum’s findings are similar to those of Moon:

She found that, in particular discourse contexts, idioms behave more flexibly than the notion of “fixed expression” would suggest.<sup>36</sup>

These findings about the relative robustness of idiom flexibility lead to a question regarding the nature of this flexibility, a question which Moon and Fellbaum leave relatively unexplored: Are there patterns to idiom flexibility? The present investigation seeks to shed light on this question using the same kinds of corpus methodologies. Specifically, these corpus studies explore patterns of syntactic flexibility within English verb-object idioms.

### **3.3 Google Books study**

#### **3.3.1 Methodology**

This section highlights the methodology used to procure my data, including details of the corpus (Section 3.3.1.1) and a close examination of the structures under investigation (Section 3.3.1.2), as well as the search procedure used in this study (Section 3.3.1.4).

##### *3.3.1.1 Corpus*

For this study, I used the English-language portion of Google Books, an online collection of more than 4.5 million digitized books containing over 468 billion words (Lin et al.

---

<sup>36</sup> Considerably more work has been done in the related field of computational linguistics (see, e.g., Fazly, Cook & Stevenson 2009 and the references cited therein), where corpora are used in machine learning. Idioms pose a particular challenge in the areas of parsing and information retrieval. However, issues of idiom flexibility are relevant only inasmuch as they contribute to a machine’s ability to recognize and understand continuous speech or text when the parts of the idiom may be noncontiguous. This issue is only tangentially related to the concerns being addressed here and is far beyond the scope of this work.

2012). These volumes represent books published as early as 1520; the majority were published after 1800. The corpus contains books from all published genres.

The size of Google Books makes it a good candidate for studying idiomatic expressions. While idioms themselves are relatively frequent, their manipulated forms are decidedly less so (Moon 1998). Smaller corpora contain too few instances of idiom variations to be useful.

However, the sheer size of Google Books, combined with the ambiguity of returned “hits”, is also a limitation. It is impossible to know with certainty how many instances of a given string occur in the corpus, and it can likewise be impossible to determine whether each is idiomatic or literal. I return to this issue in detail when I discuss my procedure in Section 3.3.1.4 below.

### *3.3.1.2 Materials*

Thirty-eight verb-object idioms were tested in this corpus study. They were selected based on the results of the idiom familiarity survey outlined in Section 2.2.1.2 above. While there is substantial overlap between the idioms investigated in this study and those used for Experiments 1 and 2, the latter is not a subset of the former. This is due to the fact that some highly frequent idioms that were deliberately excluded from Experiments 1 and 2 were included in the corpus study.

Recall that idioms were excluded from Experiments 1 and 2 if their selectional requirements prevented them from fitting the syntactic template used in the experimental design (e.g. *hit the spot* does not take an animate human subject). Idioms containing

verbs that typically do not passivize were also excluded (e.g. *have the last say*), as those idioms' inability to passivize could be attributed to properties of the verb itself rather than the idiomatic expression as a whole. For this corpus study, no such restrictions were imposed on the idioms under consideration. While a controlled experimental environment requires uniformity among experimental items, neither selectional nor (presumed) syntactic restrictions need be considered in an investigation of actual recorded usage such as a corpus study. Instead, the focus is not on what speakers are *expected* to do, but rather on what they *actually* do. For this reason, idioms that were excluded from the experimental studies in Chapter 2 were included in the present corpus study. A complete list of the idioms investigated here can be found in Table 3.2 below.

### *3.3.1.3 Variations investigated*

For this study, six particular variations were investigated: object modification (44a), object quantification (44b), topicalization (44c), passivization (44d), nominal gerundization (44e), and object incorporation (44f).

(44) Variations investigated<sup>37</sup>

- (a) You're gonna help murder this guy, you **draw the moral line** at extramarital sex? (Assignment, The 1997)
- (b) We **draw another line**. And they blow up an entire village... (Ind\_Limbaugh 1995)
- (c) **The line** I was hesitant to **draw**. (constructed example)
- (d) **The line was drawn**, and I think it was a turning point for the president... (Newsweek 2000)
- (e) I -- I'm seeking some -- a **drawing of the line** by the companies themselves... (CBS\_FaceNation 1995)
- (f) Cases that fall between the poles may present hard **line-drawing** questions... (MichLawReview 2001)

The variations were selected based on several criteria. Passivization and nominal gerundization are predicted by the structural requirement model to participate in the subset behavior outlined above. Furthermore, topicalization, object quantification, and object modification (as well as passivization) are claimed by Nunberg, Sag & Wasow (1994) to behave differently for different idiom types in their model (see Section 4.2.5.2). The final variation, object incorporation, was brought to my attention by Colin Gorrie (p.c.). Although introspection suggests that it applies rather robustly to idioms, it has to

---

<sup>37</sup> Unless otherwise noted, these illustrative examples are taken from the Corpus of Contemporary American English, which was used in the second study reported in Section 3.4 below.

my knowledge received no attention in the literature on idiom flexibility, making it an interesting candidate for investigation here. Generally, each of these variations is notable for its compatibility with some but not all idioms, making this selection a good testing-ground for idiomatic flexibility.<sup>38</sup>

These variations fall into two natural groups. Object modification and object quantification both involve inserting one or more intervening words between the verb and object of the idiom. I call these **modification-type variations**, as they involve adding a pre-object modifier. Crucially, no syntactic movement of the idiom's constituents is required. Passivization, nominal gerundization, and object incorporation, on the other hand, all involve some movement of the elements of the idiom; that is, the idiom's components do not remain in their canonical order and/or maintain their canonical form. In the passive (44d) above, for instance, the object of the idiom *draw the line* appears in subject position, and the elements of the idiom are noncontiguous. In nominal gerundization (44e), the elements retain the same order, but the verb takes on standard nominalizing morphology (i.e. the *-ing* suffix), and case-marking *of* intervenes between the verb and object. The object incorporation variation in (44f) inverts the verb and object, dropping any determiners, though the remaining elements of the idiom remain adjacent in the sentence. Because passive, nominal gerundization and object

---

<sup>38</sup> Topicalization was very difficult to search for, as any search string that might yield results would inevitably produce a wide range of false positives. For this reason, as well as the relative infrequency of topicalization in English (Speyer 2005), very few topicalized idiomatic structures were found. Therefore, topicalization has been excluded from further discussion.



incorporation all affect the argument structure of the idiom, I call them **argument-structure-type variations**.

Modification-Type Variations	Argument-Structure-Type Variations
Object modification	Passivization
Object quantification	Nominal gerundization
	Object incorporation

**Table 3.1: Variation types**

Below, I point out that these two variation types have distinct syntactic characteristics, based on the different effects each type has on the idiomatic constituents. This is elaborated in the discussion of the results (Section 3.3.2). In the next section, I discuss the procedure used for the corpus search.

#### *3.3.1.4 Procedure*

I searched the Google Books corpus (<https://books.google.com>) for each of the 38 idioms outlined in Section 3.3.1.2 in each of the variations discussed above, for a total of  $38 \times 6$  or 228 searches. I used the Google Books user interface to search for each string, using the Boolean operator \* to represent optional or unknown material. A representative paradigm of the six search strings used for the six variations is provided in (45) below.

## (45) Google Books search strings

- (a) “break the \* ice” – object modification
- (b) “break \* ice” – object quantification“
- (c) “ice \* broke” – topicalization
- (e) “ice \* broken” – passivization
- (f) “breaking of the ice” – nominal gerundization
- (g) “ice-breaking” – object incorporation<sup>39</sup>

The tense varies across the six variations, based on which felt most natural to me.

However, the tense of each variation was held constant for each idiom; hence, past tense instances of object modification and object quantification were omitted, as were present tense instances of topicalization, across all idioms. In the passive, the tensed auxiliary verb was represented with the Boolean operator \*, allowing for all tenses to be included.

There is no tense specified in the nominal gerundization and object incorporation structures.<sup>40</sup>

---

<sup>39</sup> Google Books is apparently insensitive to punctuation in search strings, even those in quotation marks. Thus, searching for “ice-breaking” returned results such as “ice breaking”, “ice. Breaking”, and “ice, breaking”, among others. In the case of “ice breaking”, this was a positive result, as the idiomatic object-incorporated structure is often written both with and without a hyphen. In the other cases, however, it simply produced more noise.

<sup>40</sup> The search methodology also limited me in some cases to searching for a single variation at a time, although it is logically possible for an idiom to undergo multiple variations simultaneously. For instance, searching for “breaking of the ice” would omit “breaking of the proverbial ice”. Notice, however, that the passivization and topicalization search strings allow for the co-occurrence of modification and quantification, since the object NP occurs at the beginning of the search string. This is simply an artefact of the search interface.

Because the purpose of this investigation was to determine whether speakers of English use idiomatic expressions in (specific) noncanonical structures, idiom/variation pairs were coded as “1” or “0”, where “1” means that at least one example of the relevant idiomatic variation occurred and “0” means no relevant examples occurred. After executing a search, the results were inspected to check for idiomatic strings. As soon as one idiomatic use of the string was found, the idiom was coded as “1” for that variation. If I examined the first 100 hits and found no idiomatic uses of the string, the idiom was coded as “0” for that variation.

I decided to use binary coding for these data in large part because getting a specific count of idiomatic examples for each string was problematic. First, for some search strings, over one million results were returned. Due to limited time and resources, it was simply impossible to manually determine whether each was being used idiomatically or literally.<sup>41</sup> A way of approximating the counts would be to examine a fixed number of results for each search string, e.g. 100, and then extrapolate the number of idiomatic items found in that sample over the total number of results returned; this would produce an estimated count of the total number of idiomatic items. However, there is a broad issue with Google’s search and return algorithm. The number of “hits” is itself an estimate that changes from one search to the next, even when identical search strings are entered. The estimate can be off by many hundreds or even thousands, making an extrapolation based on that number meaningless. Another factor confounding

---

<sup>41</sup> Distinguishing idiomatic from literal uses of strings is the focus of much work in distributed computational semantics (see, e.g., Salton, Ross & Kelleher 2016 and references cited therein). However, it is far beyond the scope of my expertise, as well as this dissertation.

counts in Google Books is duplicate materials. For example, a single quote from the same author could appear many times in one search, either quoted in several different books or reprinted in multiple editions of the same book. Although this accounts for only a small number of results, counting these as separate occurrences is problematic when the total number of hits is very low.<sup>42</sup> However, excluding these duplicates is impossible when the total number of results is too large to manually search. For these reasons, it was determined that a categorical yes/no characterization of the data was most appropriate.

Some search strings yielded examples that did not fit the desired structural pattern but were nonetheless interesting examples of idiomatic usage in noncanonical forms. These examples, when found, were recorded in a database.

Once the data had been gathered, I ran several analyses on the results. These are discussed in detail in the following section.

### 3.3.2 *Results*

Table 3.2 shows the results of this corpus investigation, where “1” indicates that the given idiom occurred in the specified variation in the corpus at least one time, and “0” indicates that the idiom did not occur in that variation in the corpus.

<b>Idiom</b>	<b>Argument-Structure Type</b>			<b>Modification Type</b>	
	<b>Pass.</b>	<b>Inc.</b>	<b>Ger.</b>	<b>Mod.</b>	<b>Quant.</b>
break the ice	1	1	1	1	1
break the news	1	1	1	1	1
bust a move	0	1	0	0	0

---

<sup>42</sup> In principle, COCA could have the same problem; however, the carefully curated nature of the COCA corpus implies that such overlap in printed materials is less likely to occur.

Idiom	Argument-Structure Type			Modification Type	
	Pass.	Inc.	Ger.	Mod.	Quant.
cross paths	1	1	1	1	0
draw a blank	0	0	0	1	0
draw the line	1	1	1	0	0
drop a bomb	1	1	1	1	0
give a hand	0	0	0	1	0
go the distance	0	0	0	1	1
hang a left	0	0	0	0	0
hang a right	0	0	0	0	0
hang a U-ey	0	0	0	0	0
have a go	0	0	0	1	0
have the last say	0	0	0	0	0
hit home	0	0	0	0	0
hit the hay	1	1	1	0	0
hit the road	0	0	0	1	0
hit the sack	0	0	0	0	0
hit the spot	0	1	0	1	0
jump the gun	1	1	1	0	0
lend a hand	1	1	1	1	0
make a killing	1	0	0	1	0
make amends	1	1	1	1	1
make way	1	0	0	0	0
pay attention	1	1	1	1	1
pop the question	1	1	1	1	0
pull strings	1	1	1	1	1
rock the boat	1	1	1	1	0
see the light	0	0	0	1	0
spill the beans	1	1	1	0	1
start a family	1	1	1	1	0
take a back seat	1	0	0	0	1
take a chance	1	1	1	1	1
take a chill pill	0	0	0	0	0
take a leak	0	0	0	0	0
take a raincheck	0	0	0	0	0
talk trash	1	1	1	1	1

Idiom	Argument-Structure Type			Modification Type	
	Pass.	Inc.	Ger.	Mod.	Quant.
turn heads	1	0	0	0	1

Pass. = Passivization, Inc. = Object Incorporation, Ger. = Nominal Gerundization,  
Mod. = Object Modification, Quant. = Object Quantification

**Table 3.2: Results of Google Books corpus search**

The data appear to confirm the primary prediction of the structural requirement approach: There is a subset relationship between the gerundization and passivization variations such that all of the idioms that occurred in gerundization also occurred in passivization. Furthermore, the set of idioms that occurred in object incorporation is a subset of the idioms that occurred in passivization (with two exceptions) and a superset of the idioms that occurred in gerundization. Thus, the argument-structure-type variations are all involved in subset relationships of the kind predicted by the structural requirement model.

There are two exceptions to this generalization: Both *hit the spot* and *bust a move* occurred in the object incorporation variation but not the passive. In the case of *hit the spot*, there was interference in all of my searches from the unrelated idiomatic NP *sweet spot*, which might have affected these results; for instance, it might have been an abundance of irrelevant hits, rather than an utter lack of relevant ones, which prevented me from finding passive instances of *hit the spot*. For *bust a move*, it is less clear what might be causing this inconsistency, although I speculate that this might be some sort of light verb construction rather than a truly idiomatic form (see discussion in Section 1.2.6).

In spite of these two exceptions, this relationship among the argument-structure-type variations appears to be robust. In contrast, the modification-type variations show a slightly weaker subset relationship: All but three of the idioms that allow quantification also allow modification. The modification-type and argument-structure-type variations appear to be relatively independent from one another.

Several statistical analyses were conducted on the basis of these observations to test these impressionistic assessments, and the results are reported in the following subsections.

#### *3.3.2.1 Conditional probabilities*

The subset relationships identified among the variations are not perfect; that is, there are exceptions, at least in some cases. One way to gauge the strength of these relationships is using conditional probabilities, which measure the likelihood that one event will occur, given the fact that another event has already occurred.

Table 3.3 below shows the conditional probabilities for each of the variation pairs.

<b>B \ A</b>	<b>Quant.</b>	<b>Mod.</b>	<b>Pass.</b>	<b>Inc.</b>	<b>Ger.</b>
<b>Quant.</b>		0.73	0.91	0.73	0.73
<b>Mod.</b>	0.38		0.67	0.67	0.62
<b>Pass.</b>	0.48	0.67		0.81	0.81
<b>Inc.</b>	0.42	0.74	0.89		0.89
<b>Ger.</b>	0.47	0.76	1.00	1.00	

**Table 3.3: Conditional probabilities  $P(A|B)$  for Google Books**

This table provides  $P(A|B)$ , the probability of variation A, given variation B, where A is represented by the row headers and B is represented by the column headers. Thus,  $P(\text{Quantification}|\text{Modification})$ , the probability that an idiom occurred in modification given that it also occurred in quantification, is 0.38 or 38%. However,  $P(\text{Modification}|\text{Quantification})$ , the probability that an idiom occurred in quantification given that it also occurred in modification, is 0.73 or 73%.

The cells in the table with bold outlines are those focused on in the discussion above. A conditional probability of 1.00 indicates a perfect subset relationship; that is, there is a 100% chance that an idiom occurred in passivization, given that it also occurred in gerundization. The same is true for incorporation and gerundization. The third subset relationship identified among the argument-structure-type variations was between incorporation and passivization. We can see from the table that there was an 89% chance that an idiom that occurred in incorporation also occurred in passivization. As for the modification-type variations, there was a 73% chance that an idiom that occurred in quantification also occurred in modification.



Of interest are the cells with dotted outlines. These were not discussed above but also have very high conditional probabilities. The lower-rightmost of the two,  $P(\text{Gerundization}|\text{Incorporation})$ , is significant because the sets of idioms that occurred in these two variations overlapped substantially. In other words, while gerundization was a perfect predictor of incorporation, the opposite was also nearly true because most idioms that occurred in one of the two variations occurred in both.

The other dotted cell,  $P(\text{Passivization}|\text{Quantification})$ , with a probability of 0.91, is more surprising, spanning the argument-structure-type and modification-type divide. However, some degree of correlation between passivization and determiner type is expected, given that passivization involves the promotion of the object from a canonically focused, new information, nonspecific position (i.e. the object position) to a canonically topical, old information, specific position (i.e. the subject position). This has consequences for the types of determiners that occur in the two locations, and hence we expect passivization and quantification to interact with modifiable idioms.

While conditional probability provides a nice way of quantifying the observed subset relationships, it is not a statistical test per se. That is, it does not test the statistical significance of these subset relationships. The following sections explore some potential ways to do just that.

### *3.3.2.2 Dependences between variations*

One way to find statistical corroboration for these subset relationships is to run tests for dependence between each of the variations. Although dependence is a two-way

relationship while subset relationships are unidirectional, it is possible that this kind of test can provide insight into the data. The most well-known test of this type is the chi-squared test for independence. The null hypothesis in such tests is that the rows and columns in a 2x2 contingency table are independent. A significant result, then, suggests dependence between the factors under consideration.

In order to make use of the simple test for independence, I transformed the results from Table 3.2 into 2x2 contingency tables, one for each pair of variations, where the rows distinguish idioms that have a “0” versus those that have a “1” for the first variation, and the columns distinguish idioms that have a “0” versus those that have a “1” for the second variation. As an example, the contingency table comparing the passive and object incorporation variations is provided in Table 3.4 below.

		<b>Object Incorporation</b>	
		<b>0</b>	<b>1</b>
<b>Passivization</b>	<b>0</b>	15	2
	<b>1</b>	4	17

**Table 3.4: Contingency table for Google Books data**

In this table, the top left cell tells us that 15 idioms were assigned a “0” for both passivization and object incorporation, i.e. 15 idioms occurred in neither the passive nor the object incorporation variation. The top right cell says that two idioms occurred in object incorporation but not the passive. According to the bottom left cell, four idioms occurred in passivization but not object incorporation. Finally, 17 idioms occurred in both passivization and object incorporation, as shown in the bottom right cell.

Contingency tables identical to the one in Table 3.4 were created for each of the ten logically possible pairings of the five variations investigated in this study. Then Fisher's Exact Test was used to test the independence of each pair of variations. Because there were many cells in the contingency tables with small or zero values, I used Fisher's Exact Test in place of the chi-squared test, though the two tests work in similar ways. The results of the test are provided in Table 3.5.

	<b>Modification</b>	<b>Passivization</b>	<b>Object Incorporation</b>	<b>Nominal Gerundization</b>
<b>Quantification</b>	$p=0.0062$	$p=0.0099$	$p=0.15$	$p=0.037$
<b>Modification</b>		$p=0.19$	$p=0.049$	$p=0.025$
<b>Passivization</b>			<b><math>p=4.74\text{e-}05^*</math></b>	<b><math>p=2.20\text{e-}07^*</math></b>
<b>Object Incorporation</b>				<b><math>p=1.89\text{e-}08^*</math></b>

**Table 3.5: Fisher's Exact Test results for Google Books**

Because this is a multiple-comparison procedure where a statistical test is applied multiple times to the same dataset, Bonferroni correction is necessary to reduce the risk of type I error, or the possibility of a "false positive". Applying the necessary correction, the accepted significance level has been lowered to  $p<0.005$ , rather than the standard  $p<0.05$ .<sup>43</sup> Bearing this mind, the table shows highly significant results (indicated with stars and shown in bold) for the comparisons between the argument-structure-type

<sup>43</sup> Bonferroni correction is calculated using the formula  $\alpha' = \alpha / k$ , where  $\alpha$  is the original significance level,  $k$  is the number of statistical tests performed, and  $\alpha'$  is the new significance level.

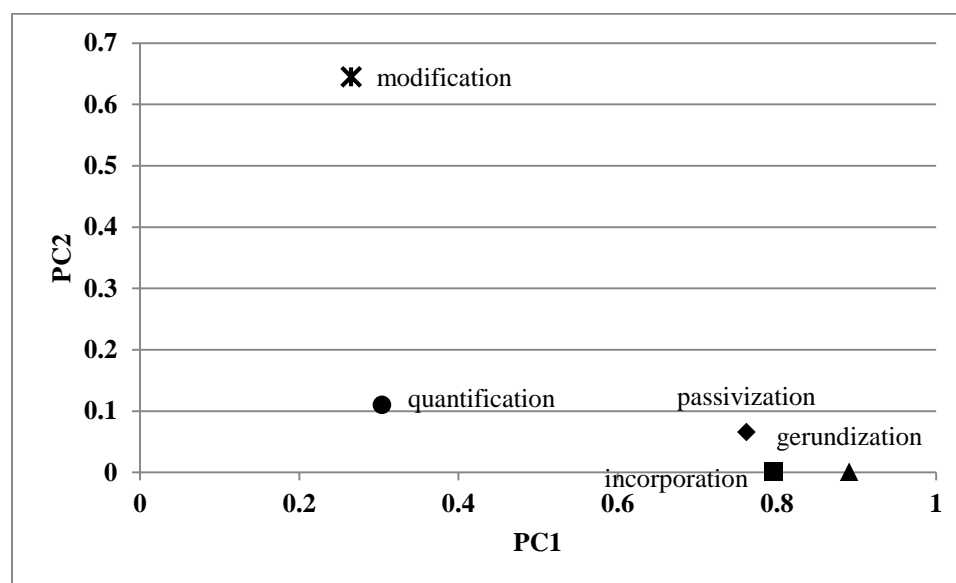
variations, namely passivization, object incorporation, and nominal gerundization, suggesting that these three variations are not independent. While some of the other comparisons approach significance—e.g. quantification and modification—none reach the minimum threshold of  $p < 0.005$ .

To be clear, Fisher’s Exact Test does *not* test for subset relationships, which are unidirectional; rather, it tests for a bidirectional dependence between the variations. However, we can see by comparing Table 3.5 with Table 3.3 that the variations that are highly correlated are the same ones that are involved in subset relationships.

### 3.3.2.3 *Multiple Correspondence Analysis*

Multiple Correspondence Analysis (MCA) offers another way of analyzing these data. MCA is a form of dimensionality reduction that aims to account for the largest amount of variance among a set of variables using the smallest number of dimensions. Its more familiar counterpart, Principal Component Analysis (PCA), performs a similar dimensionality reduction with continuous variables, while MCA is used with categorical variables. In both cases, this reduction in complexity is accomplished by “transforming to a new set of variables, the principal components (PCs), which are uncorrelated, and which are ordered so that the first *few* retain most of the variation present in *all* of the original variables” (Jolliffe 2002:1). For the present purposes, MCA can be used to identify relationships among the five variations investigated in the corpus study by determining how much of the variance among them can be accounted for using a reduced number of dimensions.

Figure 3.1 below provides a graphical representation of the first two principal components resulting from a Multiple Correspondence Analysis of the five variations investigated in Google Books.



**Figure 3.1: Multiple Correspondence Analysis (MCA) for Google Books**

In this figure, the five variations are graphed according to their correlation coefficients with respect to the first and second principal components identified by the analysis. More specifically, the horizontal axis shows the degree of correlation with the first principal component (PC1), and the vertical axis shows the degree of correlation with the second principal component (PC2).

All five variations correlate significantly with PC1, although to different degrees.<sup>44</sup> Gerundization ( $R^2=0.89$ ,  $p<.001$ ), incorporation ( $R^2=0.80$ ,  $p<.001$ ), and

<sup>44</sup> This is not surprising given that these factors are the very data used to determine the principal components in the first place. In other words, this exercise can be somewhat circular. However, it is still useful in determining which factors contribute more or less to each component.

passivization ( $R^2=0.76$ ,  $p<.001$ ) are correlated most highly with this dimension. Quantification ( $R^2=0.30$ ,  $p<.001$ ) and modification ( $R^2=0.27$ ,  $p<.001$ ) also have statistically significant correlations with respect to PC1, although their correlation coefficients, indicated by the value of  $R^2$ , are much smaller. In contrast, only modification ( $R^2=0.65$ ,  $p<.001$ ) and quantification ( $R^2=0.11$ ,  $p<.05$ ) correlate with PC2, and the correlation coefficient for quantification is substantially smaller.

These statistical differences manifest in three distinct groups in Figure 3.1: The three argument-structure-type variations cluster together in the lower right, showing that they are more highly correlated with PC1, but not correlated with PC2. The two modification-type variations, however, stand apart from both the cluster of argument-structure-type variations as well as each other. More specifically, quantification and modification are less strongly correlated with PC1 than the argument-structure-type variations (although the correlations are still statistically significant). However, modification is strongly correlated with PC2, while the correlation between quantification and PC2 is much weaker.

Although the nature of the principal components cannot be identified precisely, it is clear that the behavior of the argument-structure-type variations is different from that of the modification-type variations, particularly with respect to the first principal component. The numerical data in Table 3.6 support this conclusion.

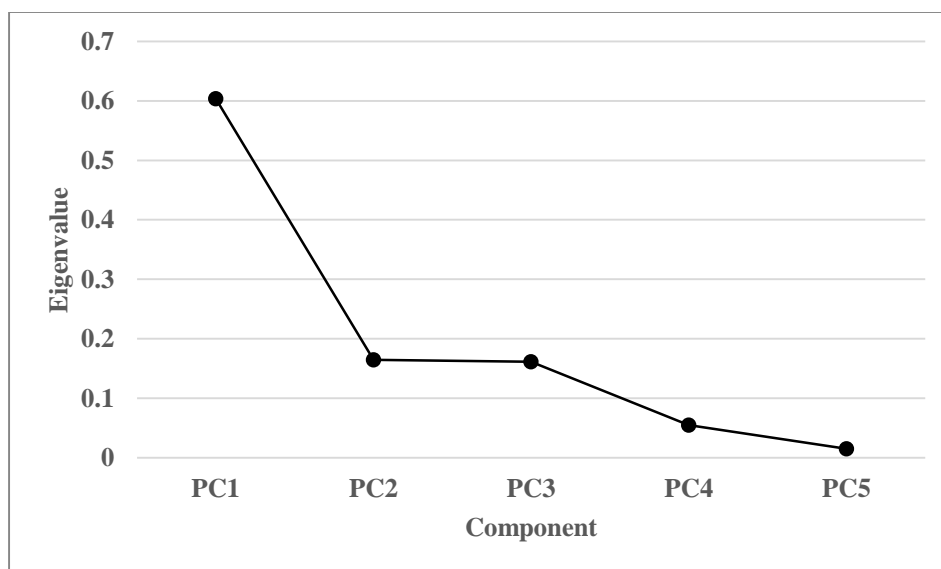
	<b>PC1</b>	<b>PC2</b>	<b>PC3</b>	<b>PC4</b>	<b>PC5</b>
<b>Eigenvalue</b>	0.60	0.16	0.16	0.05	0.02
<b>Percentage of Variance</b>	60.36	16.46	16.17	5.50	1.51
<b>Cumulative Percentage</b>	60.36	76.82	92.99	98.49	100.00

**Table 3.6: Multiple Correspondence Analysis (MCA) for Google Books**

These data show that PC1 accounts for 60% of the total variance in the data, while PC2 accounts for an additional 16%. PC3, which is not represented in the two-dimensional depiction in Figure 3.1, is also playing a notable role, accounting for just over 16% of the variance. But the remaining two dimensions, PC4 and PC5, account for a combined 7%, suggesting that at most three dimensions are needed to explain the variance seen among these five variations. In fact, PC1 accounts for over half the variance, suggesting that it is by far the most significant dimension. The scree plot in Figure 3.2 depicts the substantial difference between the contribution of PC1 and the remaining principal components. It is not unreasonable to exclude the remaining dimensions from further discussion.<sup>45</sup>

---

<sup>45</sup> Typically, the “elbow” in a scree plot is considered the cutoff point for including or excluding components from analysis: only components to the left of the “elbow” are retained. In Figure 3.2, PC1 alone is to the left of the “elbow”. Another small drop-off occurs between PC3 and PC4, so it might also be helpful to consider the roles of these two components in further analyses.



**Figure 3.2: Scree Plot for Google Books MCA**

Because PC1 draws a clear separation between the argument-structure-type variations, which strongly correlate with this dimension, and the modification-type variations, which only weakly correlate, it seems plausible that this dimension represents a way of differentiating variations that affect argument structure from those that do not. The structural implications of the three argument-structure-type variations are the focus of Section 4.3 below.

#### *3.3.2.4 Summary*

The results of this study suggest that there is a subset relationship among the argument-structure-type variations, while the modification-type variations are only weakly related to each other and unrelated to the argument-structure-type variations. Within the argument-structure-type variations, gerundizable idioms are a subset of the idioms that



participate in incorporation, which are, in turn, a subset of passivizable idioms. That is, if an idiom occurred in the nominal gerundization variation, it also occurred in the object incorporation variation; and if it occurred in the incorporation variation, it also occurred in the passivization variation. Conditional probabilities, Fisher's Exact Test, and Multiple Correspondence Analysis all provide some evidence for this empirical observation. Specifically, these three subset relationships all have very high probabilities. Fisher's Exact Test corroborates these relationships by revealing significant levels of dependence among all of the argument-structure-type variations. Finally, the MCA analysis found that PC1, which separates the argument-structure-type variations from the modification-type variations, accounts for over half the total variance, suggesting that this distinction is a significant one.

These results are in line with the predictions of the structural requirement model of idiom flexibility. The subset relationship identified in the Google Books corpus between passivizability and gerundizability is precisely the one predicted by the model—gerundizable idioms are a subset of passivizable ones, not the other way around. The fact that object incorporation also participates in these subset relationships is somewhat surprising; I return to this in Section 4.3.2.

Section 3.4 below details a follow-up study I completed using the Corpus of Contemporary American English. The new study, while parallel to the Google Books study in many ways, was designed to improve upon several shortcomings, particularly as they pertain to the corpus itself, having to do primarily with the search interface and the

inability to get count data (rather than simple binary data). These differences are discussed in detail.

### **3.4 COCA study**

#### ***3.4.1 Methodology***

The methodology used in this study was similar to that used in the Google Books study outlined above. However, some technical differences arose due to disparities between the two corpora; there were also some improvements made to the procedure. This section highlights these differences, beginning with a description of the COCA corpus.

##### ***3.4.1.1 Corpus***

The Corpus of Contemporary American English (COCA; Davies 2008-) consists of over 450 million words from 189,431 texts of American English, balanced for time period and genre. The corpus contains approximately 20 million words from each year from 1990-2012, and the texts are drawn roughly evenly from five genres: spoken (95 million words), fiction (90 million words), popular magazines (95 million words), newspapers (92 million words), and academic journals (91 million words).

COCA differs drastically from Google Books, both in composition and size, and these differences have concomitant advantages and drawbacks. The fact that COCA is a balanced corpus—covering five genres (including spoken English)—that spans only 22 recent years in the history of English, rather than a collection of books spanning the entire history of written English, is likely to affect the results of a study investigating idiomatic

flexibility, although it is not entirely clear how. Nonetheless, it is advantageous to be able to code and sort results by genre, allowing for the investigation of questions such as whether speakers are more likely to use idiomatic variations in one genre or another. Though not the focus of this study, such investigations are a rich area for future research.

Considerations of size have both advantages and disadvantages. On the one hand, a fixed corpus with a definite and predetermined size, coupled with a user-friendly search interface, eliminates many of the problems with Google Books outlined in Section 3.3.1.4 above. For instance, there is no longer an ambiguity with the term “hits”, where the search algorithm retrieves an estimated number of results. Instead, the number of results is absolute. Furthermore, because the corpus is considerably smaller, it is easier<sup>46</sup> to manually search through each result.

In spite of these benefits, the size of COCA is also a major concern. Although 450 million words is considered to be a large corpus for many applications, it is three orders of magnitude smaller than the 468 billion-word Google Books corpus. It is possible, therefore, that paucity of data will be an issue here.

### 3.4.1.2 Materials

For the COCA study, 40 verb-object idioms were used. Thirty-two were selected from those used in the Google Books study; six were excluded due to various factors, discussed immediately below, and eight new idioms were added to bring the total to 40. The idioms *go the distance*, *hit the spot*, and *see the light* were excluded from the COCA

---

<sup>46</sup> Although not always possible, an issue I address in Section 3.4.1.4 below.

study due to interference. Searches for *hit the spot* often returned instances of the related idiomatic NP *sweet spot*, used in conjunction with the literal verb *hit* (see discussion of this idiom in Section 3.3.2 above). Idiomatic uses of *see the light* were nearly overwhelmed by metaphorical religious uses (i.e. *see the light of God*). Similarly, there was no clear division between idiomatic and literal or metaphorical uses of *go the distance*. The remaining three idioms, *cross paths*, *give a hand*, and *lend a hand* were excluded due to difficulties with argument structure.<sup>47</sup> Specifically, *cross paths* requires either a plural subject (e.g. *Lisa and Lauren crossed paths*) or a post-verbal prepositional phrase headed by *with* (e.g. *Lisa crossed paths with Lauren*); *give a hand* and *lend a hand* require either an indirect object (e.g. *Lisa gave Lauren a hand*) or a dative *to*-phrase (e.g. *Lisa gave a hand to Lauren*).<sup>48</sup> In these cases, the restrictive argument structure requirements resulted in unexpected limitations on the search results; for example, the search string used to find active instances of the idiom *give a hand* omitted all cases with an intervening indirect object.

Eight new idioms were added to the COCA search. *Clear the air*, *drop the ball*, *kick the bucket*, *steal the show*, and *turn the tables* were selected based on the results of the original idiom familiarity survey (see Section 2.2.1). *Have the final say* was selected as a minimal pair with *have the last say*; both idioms have the same meaning and similar structures, and I wanted to investigate whether their behaviors were reflected similarly in

---

<sup>47</sup> Recall that these idioms were excluded from the experiments in Chapter 2 for the same reason.

<sup>48</sup> Note that the recipient may sometimes be omitted, particularly with *lend a hand*, as in *Eli is lending a hand for this project*.

the corpus. Finally, *drop trou* and *hang ten* were added as (relatively) newer innovations, to see whether that might affect their manipulability. For a full list of idioms used in this study, see Table 3.7.

### 3.4.1.3 *Variations investigated*

The same six variations were investigated in COCA as in Google Books (see Section 3.3.1.3).

### 3.4.1.4 *Procedure*

Due to the intricacies of the COCA search interface, the procedure was modified slightly for this investigation from that used for Google Books. First, in addition to searching for each of the 40 idioms in all six variations outlined above, searches were also conducted for each idiom in the active form. This was done as a way of determining the idiom's overall frequency in the corpus.<sup>49</sup> Second, the advanced search interface allowed for more sophisticated search techniques, involving lemmatized versions of verbs, part-of-speech specifications for both verbs and objects, and distance-mediated requirements on searches like topicalization. A paradigm of search strings used for the idiom *break the ice*, paralleling that used in Google Books presented in (45), is given in (46) below. (A discussion of the notational conventions used here follows.)

---

<sup>49</sup> This was not done in the Google Books study for practical reasons: Recall that it was impossible to count the exact number of “hits” for idioms in each variation, and the same was true to an even larger extent for idioms in their canonical forms, for all of the reasons discussed in Section 3.3.1.4 above.

(46) COCA search strings<sup>50</sup>

- (a) [j\*][n\*] ice 4/0 [break] – object modification
- (b) [break] [quantifiers] ice – object quantification
- (c) . the ice 0/9 [break] – topicalization<sup>51</sup>
- (d) [be] [break] 3/0 ice – passivization
- (e) breaking of the ice – nominal gerundization
- (f) ice-breaking / ice breaking / icebreaking – object incorporation<sup>52</sup>
- (g) [break] the ice – active

Note that the COCA search interface does not require quotation marks for an exact string match. A word in square brackets matches all lemmas of that word (e.g. a search for [break] would return all instances of *break*, *breaks*, *breaking*, *broke*, etc.). The bracketed notations [j\*] and [n\*] refer to the class of words tagged as adjectives and nouns, respectively. [quantifiers] refers to the following custom-defined list of quantifiers: *all*, *another*, *any*, *both*, *each*, *enough*, *every*, *few*, *little*, *many*, *most*, *much*, *no*, *several*, *some*. Numeric notations refer to uses of the COCA interface's collocate finder. This feature

---

<sup>50</sup> Some slightly modified variants of these search strings were used for specific idioms when necessary to limit a word to a certain part of speech or eliminate collocations in order to reduce the number of items returned. For instance, for the idiom *hit home*, I used the active search string “[hit] home -[run]”, eliminating all instances of the highly frequent collocation “home run(s)”.

<sup>51</sup> The period preceding “the ice” in this search string is meant to capture the fact that topicalized elements typically appear at the beginning of sentences. The topicalized phrase and verb were allowed to occur as far apart as the search parameters permitted, because the intervening subject could be of an indeterminate size.

<sup>52</sup> Three separate search strings were used to find instances of object incorporation, as punctuation is inconsistent in these forms, and there is no way to capture all three orthographic conventions with a single search.

finds instances of a word (or lemma) within a certain distance of another word or phrase. In the case of object modification, for example, this search entry looks for a lemma of *break* up to four words to the left of an adjective or noun that is followed by *ice*. In the case of topicalization, on the other hand, the search looks the other way; 0/9 indicates that the search will look for a lemma of *break* up to nine words to the right of a period and space followed immediately by *the ice*.

The sophisticated search interface provided by COCA has several benefits over the Google Books search. First, lemmatized structures allow the results to encompass all tenses, rather than being limited to a single tense based on the verb form chosen. Second, part-of-speech tags mitigate interference from unwanted results that are generated by the Boolean operator \*; although some results still do not match the desired type, mismatches are greatly reduced. Finally, the collocate finder feature, which allows the user to limit the number of intervening words between two target words or phrases, is ideal for things like passives, which sometimes have intervening modals and sometimes do not.

After completing a search, the results were separated into idiomatic and nonidiomatic strings. Due to the smaller size of the corpus, the majority of searches yielded a small enough number of results that each could be examined individually, eliminating the need to extrapolate from a smaller sample, as would have been necessary with the Google Books results.<sup>53</sup> Unlike with the previous study, then, the exact number

---

<sup>53</sup> The following searches yielded too many items to be counted manually; in these cases, a random sample of 100 strings was examined, and the number of idiomatic items from that sample was extrapolated over the total number of results to determine the total number of idiomatic items for that search: *draw the line* – active, object modification; *make way* – active; *hit home* – active; *take a back seat* – active; *rock the boat* – active; *pay attention* – active, object quantification, object modification, passive; *break the news* – active,

of idiomatic results for each search string was recorded, rather than a categorical “yes” or “no”. Illustrative idiomatic examples from each search were saved into the COCA search interface for later access, as were examples that did not fit the desired search pattern but were nevertheless of interest.

After gathering the data, comparable analyses were conducted on these results to those done on the Google Books data. The results and analyses are discussed immediately below.

### 3.4.2 Results

The results of the COCA corpus study are presented in Table 3.7 below.<sup>54</sup> An asterisk indicates that a number was generated via extrapolation, as discussed in footnote 53, rather than by manually categorizing and counting each search result.

	Control	Argument-Structure-Type			Modification-Type	
Idiom	Act.	Pass.	Inc.	Ger.	Mod.	Quant.
break the ice	202	15	11	0	3	1
break the news	433*	8	5	0	127*	11
bust a move	21	0	1	0	2	0
clear the air	194	5	1	3	3	0
draw a blank	101	1	0	0	9	2
draw the line	935*	93	17	1	142*	11
drop a bomb	30	2	0	0	18	1
drop the ball	214*	4	0	0	5	0
drop trou	18	0	0	0	0	0

object modification; *take a chance* – active; *turn heads* – active; *start a family* – active; *hit the road* – active; *make amends* – active; *drop the ball* – active. Only 18 out of 360 searches conducted, or 5%, required extrapolation. These searches are marked with a \* in Table 3.7.

<sup>54</sup> As with the Google Books search, topicalization was vanishingly rare and, thus, has been omitted from further discussion.



	Control	Argument-Structure-Type			Modification-Type	
Idiom	Act.	Pass.	Inc.	Ger.	Mod.	Quant.
hang a left	32	0	0	0	3	1
hang a right	22	0	0	0	5	0
hang ten	15	0	0	0	1	0
have a go	126	0	0	0	7	22
have the final say	155	0	0	0	1	0
have the last say	18	0	0	0	0	0
hit home	472*	0	0	0	0	0
hit the hay	30	0	0	0	1	0
hit the road	678*	0	0	0	6	0
hit the sack	59	0	0	0	0	0
jump the gun	148	0	1	0	2	0
kick the bucket	38	0	0	0	1	0
make a killing	162	1	0	0	26	1
make amends	505*	6	2	1	15	8
make way	783*	0	0	0	0	0
pay attention	7254*	713*	0	0	2242*	2402*
pop the question	120	2	0	0	5	0
pull a u-ey <sup>55</sup>	2	0	0	0	0	0
pull strings	69	5	10	0	13	33
rock the boat	212*	0	3	0	8	0
spill the beans	123	0	0	1	10	3
start a family	402*	0	0	0	23	8
steal the show	241	1	7	0	6	1
take a back seat	364*	0	0	0	6	0
take a chance	903*	0	5	0	78	36
take a chill pill	8	0	0	0	1	0
take a leak	132	0	0	0	6	0
take a raincheck	6	0	0	0	0	0
talk trash	97	1	132	0	8	4

<sup>55</sup> In the Google Books search, *hang a u-ey* was used in place of *pull a u-ey*, the former being more common in my dialect. However, there were no results for *hang a u-ey* in COCA, so *pull a u-ey* was selected as a roughly equivalent substitute. The issue of idioms with similar forms and meanings has received some attention in the literature; see, e.g. the discussions of this phenomenon in Gazdar et al. (1985) and Moon (1998).

	Control	Argument-Structure-Type			Modification-Type	
Idiom	Act.	Pass.	Inc.	Ger.	Mod.	Quant.
turn heads	240*	2	52	2	14	29
turn the tables	349	89	4	6	3	0

Pass. = Passivization, Inc. = Object Incorporation, Ger. = Nominal Gerundization,

Mod. = Object Modification, Quant. = Object Quantification

\*This number is generated via extrapolation and is not an exact count; see footnote 53.

**Table 3.7: Results of COCA corpus search**

Three observations are readily apparent from looking at the data. First and most prominent is the relative paucity of manipulated idioms compared to the number of active idioms. While the active column contains numbers often in the tens or hundreds—and no zeroes at all—the variation columns are dominated by single-digit figures, with zero outnumbering the others. In fact, of the 200 cells (excluding the active column<sup>56</sup>), 114—or 57%—are occupied by zeroes. The Google Books data contain similar numbers: 101 out of 190 cells—just over 53%—are occupied by zeroes. In that case, however, the zeroes were less noticeable because the raw numbers were not included, and “1” provided less of a contrast with “0”; the active column was also missing for comparison.<sup>57</sup>

The second observation of note here is that the subset relationships identified in the Google Books data are not readily apparent in the COCA results, though there appears to be a similar trend here. If gerundizable idioms are a subset of the idioms that

<sup>56</sup> Actives are excluded here because they were not included in the Google Books search.

<sup>57</sup> Constructional frequency is one variable that is not controlled in this experiment but very well may be affecting these results. That is, it is possible that passivization, nominal gerundization, and the other variations under investigation here are not just uncommon for idioms but are rare overall. This possibility should certainly be explored in future work.

occurred in object incorporation, and those are a subset of passivizable idioms, then we need to address a larger number of exceptions than in the previous study.

*Jump the gun*, *rock the boat*, and *take a chance* occurred in object incorporation but not passivization; *spill the beans* occurred in nominal gerundization but neither object incorporation nor passivization. In the case of *jump the gun*, the object-incorporated structure has taken on an established meaning; *gun-jumping* is the act of trading securities on the basis of information that has not yet been disclosed to the public. It is not surprising, then, that this specialized meaning would be more frequent (and, thus, more likely to appear in a corpus that includes newspaper text) than the passive variation. As for *rock the boat*, the object incorporated *boat-rocking* is a frequent term in newsprint, particularly in texts about the business world. Again, where matters of frequency are of utmost concern, it is possible that this structure has a higher relative distribution than that of its passive counterpart. While *take a chance* does not necessarily fall into this category, I note in passing that it looks suspiciously like a light verb construction, an issue which was briefly addressed in Section 1.2.6. Finally, *spill the beans* almost certainly falls victim to the poverty of data here; this idiom is canonically passivizable, and several native speakers and trained linguists corroborate my intuition that it can undergo object incorporation, yet both of these structures are missing from the corpus. As a final comment, I note that the majority of these exceptions are of the same type as those found in the Google Books data; that is, the two exceptions in Google Books were also idioms that participated in object incorporation but not passivization. The role of

object incorporation is examined in the light of these exceptions in some detail in Section 4.3.2.

Considering these exceptions, it seems possible—perhaps even likely—that an overall lack of data is playing a role here, i.e. that these are accidental gaps. Recall that COCA is three orders of magnitude smaller than Google Books in word count. Given that, as well as the large number of single-digit figures in the cells in Table 3.7, it is a real possibility that the subset relationships would be more robust if there were more source material; the relevant examples might turn up to fill in the gaps.

The third—and perhaps most interesting—thing to note about the data in Table 3.7 is the fact that new subset relationships exist among these data, identifying modification as an *ubervariation* of sorts. No idiom occurred in any other variation if it did not also occur in the modification variation. This fact is quite surprising given that modification was not particularly common in the Google Books data. It is possible that the search string used to find instances of modification in Google Books—which contained the generic Boolean operator \* rather than part-of-speech specific tags, as in COCA—yielded many false positives, biasing the search toward empty results, while the COCA interface allowed for a very precise search string, which made finding instances of modification much easier in the latter. The disparity in results between Google Books and COCA is discussed in much more detail in Section 3.5.

Given the subset relationships identified in these data, similar statistical measures to those used to quantify the relationships found in Google Books can be applied to the COCA data as well, in an effort to see how closely the two datasets match. These same

analyses can also provide insights into the nature of the subset relationships that hold between modification and the other variations. The results of those analyses are presented immediately below.

#### 3.4.2.1 Conditional probabilities

Once again, conditional probability was used to quantify the subset relationships between each pair of variations. These probabilities are given in Table 3.8 below.

<b>B \ A</b>	<b>Quant.</b>	<b>Mod.</b>	<b>Pass.</b>	<b>Inc.</b>	<b>Ger.</b>
<b>Quant.</b>		1.00	0.71	0.53	0.24
<b>Mod.</b>	0.52		0.48	0.42	0.18
<b>Pass.</b>	0.75	1.00		0.63	0.31
<b>Inc.</b>	0.64	1.00	0.71		0.36
<b>Ger.</b>	0.67	1.00	0.83	0.83	

**Table 3.8: Conditional probabilities  $P(A|B)$  for COCA**

Not surprisingly, all of the cells in the modification column have a probability of 1.00. This is due to the perfect subset relationships identified above: If an idiom occurred in any of the other variations, it also occurred in modification.

The cells with bold outlining, those indicating the robust subset relationships among the argument-structure-type variations from Google Books, reveal weaker probabilities, which aligns with the fact that there are more exceptions in the COCA results than there were in Google Books. However, there is still an 83% chance that an idiom that occurred in gerundization also occurred in passivization, and an 83% chance that an idiom that occurred in gerundization also occurred in incorporation; these are the highest probabilities in the table outside the Modification column. The percentage is

lower—just 71%—for the probability that an idiom that occurred in incorporation also occurred in passivization. We can see that, as in Google Books,  $P(\text{Passivization}|\text{Incorporation})$  is the weakest of the three subset relationships, though all three are weaker overall.

Finally, two other cells have comparable probabilities to those discussed above; they have dotted outlines in Table 3.8.  $P(\text{Passivization}|\text{Quantification})$  is 71% and  $P(\text{Quantification}|\text{Passivization})$  is 75%. That is, passivization and quantification both predict each other with a fairly high probability. As discussed above (Section 3.3.2.1), the relationship between passivization and quantification is not surprising given the interaction between passivization and definiteness of the object, although this interaction is not specifically predicted by the structural requirement model.

#### *3.4.2.2 Dependences between variations*

In order to test the robustness of the relationships between variations, Fisher's Exact Test was used for the Google Books data. However, because the COCA results are not binary but continuous, a more appropriate test here is Kendall's Tau rank correlation, which is appropriate for data that lacks a normal distribution, as do those in Table 3.7. Rank correlations identify dependencies between sets of data by first ranking the items in each dataset and then identifying differences between the two rankings; the more differences between the rankings, the less likely there is to be a dependency between the two datasets. Items are ranked based on their values, so the most frequent idiom in a given

variation receives the highest rank.<sup>58</sup> Unlike the more familiar Spearman's Rho rank correlation, Kendall's Tau imposes a less severe penalty for data with extreme mismatches in rank (e.g. a single data point that is ranked very highly in one dataset but very low in the other), representing general concordance among datasets.<sup>59</sup>

As with Fisher's Exact Test, Kendall's Tau is not a test for unidirectional relationships like subsets but rather for bidirectional correlation; thus, it provides comparisons between two datasets, where a significant  $p$ -value indicates that those two datasets are correlated. The Kendall's Tau rank correlation results for the raw numbers from Table 3.7 are given in Table 3.9 below.

	<b>Modification</b>	<b>Passivization</b>	<b>Object Incorporation</b>	<b>Nominal Gerundization</b>
<b>Quantification</b>	$T_s(38)=0.66$ $p=1.20e-07^*$	$T_s(38)=0.43$ $p=0.0013^*$	$T_s(38)=0.38$ $p=0.0056$	$T_s(38)=0.19$ $p=0.17$
<b>Modification</b>		$T_s(38)=0.43$ $p=5.27e-4^*$	$T_s(38)=0.33$ $p=0.0099$	$T_s(38)=0.20$ $p=0.13$
<b>Passivization</b>			$T_s(38)=0.46$ $p=7.06e-4^*$	$T_s(38)=0.40$ $p=0.0049^*$
<b>Object Incorporation</b>				$T_s(38)=0.37$ $p=0.0096$

**Table 3.9: Kendall's Tau rank correlations for COCA**

<sup>58</sup> The large number of zeroes necessitates the use of Kendall's Tau-b to resolve ties in rankings. Ties are resolved by assigning each tied value the mean rank. For example, if the dataset were {1, 3, 3, 5}, the corresponding ranks would be {1, 2.5, 2.5, 4}, where 2.5 is the mean of 2 and 3.

<sup>59</sup> As mentioned in Section 3.3.2.2 above, the more familiar chi-squared test is also inappropriate here due to the large number of cells with small values or zeroes.

Again, we assume a significance level of  $p < 0.005$  due to Bonferroni correction.

Significant results are starred in the table, and bold cells indicate the subset relationships under discussion.

Whereas, in the Google Books results, significant correlation as identified by Fisher's Exact Test corresponded fairly closely to the subset relationships identified by conditional probability, that does not seem to be the case here. Of the four perfect subset relationships identified with modification (i.e. if an idiom occurred in any of the other four variations, it also occurred in object modification), only two are significant: quantification~modification and passivization~modification. Furthermore, two of the three imperfect subset relationships identified among the argument-structure-type modifications—incorporation~passivization and gerundization~passivization—are significant here; however, the third—gerundization~incorporation—is not. Finally, the passivization~quantification relationship is significant. This is not surprising given our previous discussion of this relationship, as well as its relatively high conditional probability from Table 3.8.

Overall, these results bring little to bear on the discussion of subset relationships, which is not surprising given that Kendall's Tau is a test of bidirectional correlation rather than unidirectional subsets. It is also worth noting that the relatively high Tau coefficients and correspondingly low  $p$ -values throughout Table 3.9 can be attributed in part to the large number of zeroes in the raw data. Every idiom with a zero for both variations bolsters the significance level without providing meaningful information about relative ranks. Consequently, while Kendall's Tau provides some information about the



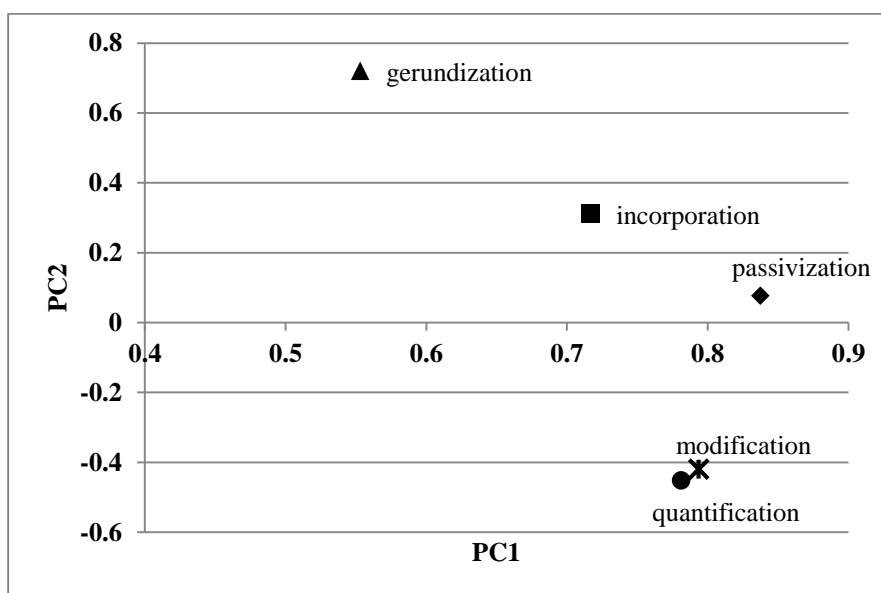
likelihood that two variations are correlated, the following section again explores dimensionality reduction as another means of quantifying these subset relationships.

#### 3.4.2.3 *Principal Component Analysis*

Recall that Multiple Correspondence Analysis was used to interpret the binary data from Google Books (Section 3.3.2.3). Principal Component Analysis (PCA) is another type of dimensionality reduction that operates on continuous rather than binary factors. Like MCA, PCA seeks to account for the largest amount of variables in a dataset with the smallest number of dimensions. Before such an analysis can be performed on the COCA results, however, the data must be transformed to account for the non-normal distribution. Consider again Table 3.7; the numbers for *pay attention* and, to a lesser extent, *break the news* are disproportionately large compared to the rest of the sample. Given this skewing in the data, any attempt to analyze the raw counts will naturally result in a bias. For instance, running PCA on the raw count data will find a natural division between passivization, modification, and quantification on the one hand and gerundization and incorporation on the other, since *pay attention* participates in the former but not the latter. This is because PCA identifies variables that account for variance in the data, and *pay attention* itself accounts for much of the variance here.

I normalized these data using log transformations, which reduces the difference in scale between the data points. However, log transformations cannot be performed on data with zeroes, as  $\log(0) = -\infty$ . Given these considerations, the following procedure was adopted, following Tukey (1977:397): Zeroes were eliminated from the dataset by

adding 0.1 to each data point. The resulting data were then log transformed. Finally, PCA was performed on the log transformed results.<sup>60</sup> The results of the PCA analysis are given in Figure 3.3 below.



**Figure 3.3: Principal Component Analysis (PCA) for COCA**

As with MCA, the axes in Figure 3.3 represent the correlation coefficients for the first and second principal components. Here as in the Google Books analysis, all five variations positively correlate with PC1, shown on the horizontal axis: passivization ( $R^2=0.84$ ,  $p<.001$ ), modification ( $R^2=0.79$ ,  $p<.001$ ), quantification ( $R^2=0.78$ ,  $p<.001$ ), incorporation ( $R^2=0.72$ ,  $p<.001$ ) and gerundization ( $R^2=0.55$ ,  $p<.001$ ). All five correlations are quite large, as indicated by the correlation coefficients. On the other hand, only gerundization ( $R^2=0.72$ ,  $p<.001$ ) is positively correlated with PC2, while

<sup>60</sup> To reiterate, running PCA on the non-transformed data would not be useful because PCA is a measure of variance, and almost all of the variance in this dataset is tied to a single item. Normalizing the data in this way is a standard procedure that allows PCA to function as intended.

quantification ( $R^2 = -0.45$ ,  $p < .001$ ) and modification ( $R^2 = -0.42$ ,  $p < .001$ ) are negatively correlated with PC2.

Some patterns emerge here. First, modification, quantification, and passivization are very close with respect to PC1, due to their high correlation. The other variations, while also highly correlated with PC1, are slightly less so and thus not grouped as tightly along this dimension. With respect to PC2, quantification and modification are both grouped together due to their negative correlation with this dimension. Gerundization is positively correlated and thus farther away from the modification-type variations on this dimension. Passivization and incorporation are uncorrelated with PC2. From the figure, it is clear that modification and quantification are patterning together with respect to both PC1 and PC2, while passivization, incorporation, and gerundization have less clear patterns.

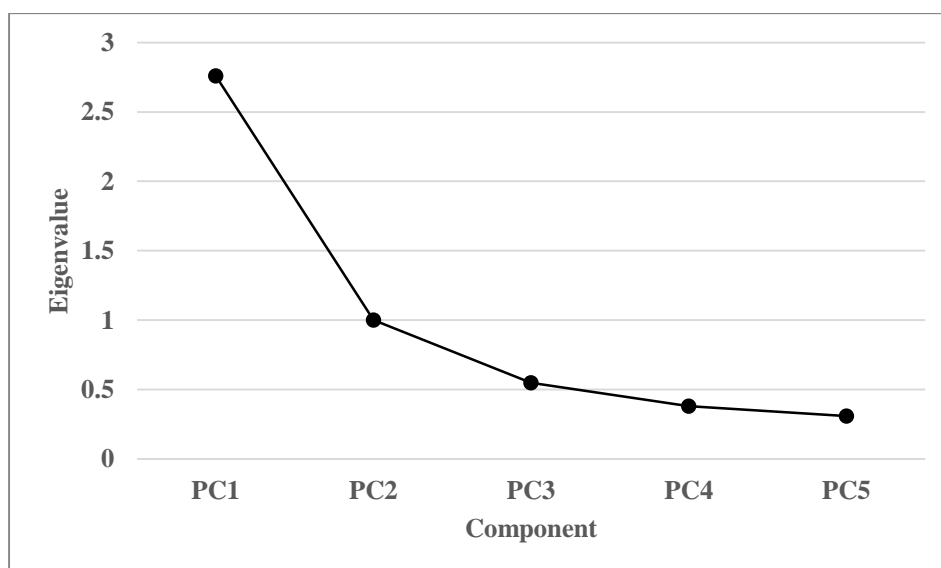
Table 3.10 contains the corresponding numerical data.

	<b>PC1</b>	<b>PC2</b>	<b>PC3</b>	<b>PC4</b>	<b>PC5</b>
<b>Eigenvalue</b>	2.76	1.00	0.55	0.38	0.31
<b>Percentage of Variance</b>	55.23	20.03	10.93	7.69	6.12
<b>Cumulative Percentage</b>	55.23	75.27	86.20	93.88	100.00

**Table 3.10: Principal Component Analysis (PCA) for COCA**

We can see that both the first and second components are making significant contributions in accounting for the variance in the data (56% and 21%, respectively), while the last two make minimal contributions. The third component makes a somewhat

meaningful contribution at around 12%. The first three components combine to account for over 88% of the total variance in the data.



**Figure 3.4: Scree Plot for COCA PCA**

The scree plot in Figure 3.4 shows that there is no obvious cut-off point for dimensions that should be retained for analysis. That is, there is no distinct “elbow” (see footnote 45); instead, there is more of a gradual curve. However, the line does flatten out after PC4, so at most the first three dimensions should be retained. Another common metric for determining this cut-off point is based on the eigenvalues: Only dimensions with an eigenvalue of one or greater should be retained. In that case, only PC1 and PC2 meet the threshold.

These results are not as clear-cut as those from the Google Books analysis. In particular, PC1 is highly correlated with all five variations, making it difficult to interpret. This could be due to the prevalence of modification as a sort of *ubervariation*, a superset

of all of the other variations. However, PC2 appears to have some relevance to the distinction between argument-structure-type and modification-type variations; the modification-type variations are both highly negatively correlated with this dimension, while the argument-structure-type variations all have positive correlations (though only gerundization has a statistically significant positive correlation).<sup>61</sup>

### 3.4.3 *Summary*

The results of the COCA study are overall less transparent than those of the Google Books study. While the subset relationships between gerundization, incorporation, and passivization appear to hold in a slightly weaker form in the COCA data, the discovery of modification as an *ubervariation* colors those results to some extent. Unlike in Google Books, where modification did not stand out in frequency, modification was much more frequent in COCA overall, and no idiom occurred in another variation if it did not also occur in the modification variation. These data are consistent with the predictions of the structural requirement model to the extent that the subset relationship holds between gerundization and passivization. No prediction was made concerning modification. The source of the differences between corpora is discussed more below, but it seems likely that corpus size is a contributing factor.

---

<sup>61</sup> Because the COCA data are continuous and the Google Books data are binary, the statistical tests used to evaluate the results are slightly different, although comparable. I did, for the sake of direct comparison, collapse the COCA data to categorical yes/no results and subject them to the exact same statistical tests that were used on the Google Books data. The results are reported in Appendix C.

### 3.5 Discussion

The results reported in Section 3.3.2 and Section 3.4.2 present an interesting and complicated picture of idiom flexibility. The Google Books study revealed a robust subset relationship among idioms with respect to three of the variations investigated here: The set of idioms that occurred in nominal gerundization also occurred in object incorporation, and those that occurred in object incorporation also occurred in passivization. This is a clear pattern with just two exceptions. The COCA data were suggestive of a similar pattern, but the results were less robust, with more exceptions. Furthermore, in COCA but not in Google Books, object modification was found to be a sort of *ubervariation* such that, if an idiom occurred in any other variation, it also occurred in object modification. The inconsistencies between the two corpora are puzzling and require further explanation, a topic I discuss briefly below in Section 3.5.1. I then explore the implications of these results, in light of the fact that—in spite of their differences—the results of both corpus studies confirm to varying degrees the predictions of the structural requirement model and are not easily captured by other extant models of idiomaticity.

#### 3.5.1 Comparing corpora

There are several differences between the corpora that could contribute to the observed differences in the results. As mentioned above, the most obvious reason that a pattern might emerge in the Google Books corpus but not in COCA is corpus size. Recall that Google Books contains roughly 350 billion words of English, while COCA is 450

million. At three orders of magnitude larger, Google Books is likely to contain a much larger number of idiomatic expressions. Furthermore, while idioms themselves are fairly common, idiom variations are very infrequent, making corpus size a serious issue, even at the scale of hundreds of millions versus hundreds of billions.

Another serious consideration mentioned above specifically in connection with object modification are the search mechanisms themselves. While variations are expected to be more common in Google Books given the relative sizes of the corpora, object modification was more common in COCA, a contradictory result. This could be attributed to the mechanisms used to search for modification in the two corpora. In COCA, the search interface's collocate finder was used in conjunction with part-of-speech tags to limit the search to any lemma of the idiomatic verb, followed within four words by an adjective or noun, immediately followed by the idiomatic object. This search mechanism allowed for standard object modification (e.g. *breaking the conversational ice*), as well as modification in which the determiner was omitted (e.g. *pulling political strings*) and modification with multiple modifiers (e.g. *clearing the bitter political air*). That is, the search was flexible, yet specific. By contrast, the Google Books search was conducted using the Boolean search string "V (DET) \* N" (e.g. "break the \* ice"). Although information is not readily available on exactly what the \* operator does during Google's search algorithm, the search results included expressions with one or more words intervening. However, I was unable to limit the number or part of speech of the intervening word(s); nor was I able to optionally omit the determiner in the search. This both limited the number of actual idiomatic results returned and also inflated the

number of irrelevant results; that is, the search was both too limited to get some instances of idiomatic object modification and so broad that it would get many instances of things other than idiomatic object modification. Because I was only able to look through a small number of results for each Google Books search, it is possible—likely, even—that modification is more common in the corpus than my results reflect.

The search methods also differed in another important way: the COCA interface allowed me to use lemmatized versions of the verbs, whereas Google Books forced me to choose a specific verb tense. Like the collocate searching discussed above, this actually should have increased the number of COCA results relative to the number of Google Books results.

Another obvious difference between the two corpora is the content. COCA is a conscientiously designed corpus, limited to American English, balanced across both genres and time, and spanning just two decades. Unfortunately, Google Books does not allow the researcher to control for these kinds of factors. Instead, it is a hodgepodge collection of published books—whatever the Google Books teams had to hand and chose to scan—heavily biased toward recent years but going back five centuries, and limited in no way by dialect. Given these rather drastic differences, it is not at all surprising that differences in idiomatic usage emerged. For instance, one could easily imagine that idiomatic variations—a form of linguistic creativity—would be used most frequently in certain genres of writing. Indeed, the COCA examples, which are tagged for source, come largely from popular magazines. It is likely that the stylistic writing of these magazines readily lends itself to idiomatic variations, whereas academic journals would



be largely devoid of such colloquialisms. It is also possible that certain variations are characteristic of certain genres. For instance, the fact that object modification was much more prevalent in COCA than in Google Books could also be an artefact of the genre differences between the two corpora. Similarly, the historical nature of Google Books might admit usages that have since gone out of style, and the lack of dialectal specification could allow for distinctly British or Australian usages to be represented in the data as well.

Finally, in addition to differences between the corpora themselves, methodological differences in data collection could well be influencing the results. Recall that the Google Books data were recorded as categorical and binary—either an idiom occurred in a given variation or it did not. The COCA data, however, were recorded as raw counts: How many times did an idiom occur in a given variation? Although the same categorical statistics were run on the COCA data, with very different results (see Appendix C), it is likely that the Google Books results were affected by the methodology, that is, by the fact that I was unable to check every result for idiomaticity and get raw counts rather than searching a small subset. It remains an open question whether the robust subset relationships identified in the current iteration of the Google Books corpus search would be replicated under different search procedures.

Given these considerations, definite conclusions regarding idiom flexibility are difficult to draw. Nonetheless, the following section presents some plausible interpretations of the results discussed here.

### 3.5.2 *Implications of results*

The results of both studies—and those of the Google Books study reported in Section 3.3.2 in particular—confirm the predictions of the structural requirement model of idiom flexibility (Folli & Harley 2007). This account predicts subset relationships among idioms due to the hierarchical nature of the verb phrase and the associations between certain variations and functional projections within the verbal functional structure. Importantly, the subset relationship predicted by the model is exactly the one discovered here: The set of passivizable idioms is a superset of the set of gerundizable idioms. This is a direct reflex of the fact that VoiceP, which is the locus of passivization, is higher in the verbal functional structure than *v*P, which is the locus of nominal gerundization.

Object incorporation plays an interesting role here. Unlike passivization and gerundization, it is not strongly associated with a particular verbal functional projection. However, it demonstrates the same subset behaviors as the other two argument-structure-type variations. It is reasonable to suggest that a functional projection should be added to accommodate these data. Section 4.3.2 explores this possibility in some detail.

No other prevailing model of idiomaticity predicts such subset relationships. The following chapter examines several such models in detail and argues that most are unable to account for these empirical observations. Those that can accommodate the data do so by stipulation only. The structural requirement hypothesis alone actually predicts these empirical results.

### 3.6 Conclusion

In this chapter, I have discussed the results of two corpus studies investigating the flexibility of English verb-object idioms. These results reveal systematic variation in idiom flexibility, particularly among the argument-structure-type variations. Specifically, the following subset relationships were found in Google Books: The set of idioms that occurred in nominal gerundization is a subset of the idioms that occurred in object incorporation; and the set of idioms that occurred in incorporation is subset of the idioms that occurred in passivization. The picture was less clear with the COCA data, where object modification was found to be a sort of *ubervariation*: If an idiom occurred in any other variation, it also occurred in the modification variation. Nonetheless, the presence of such systematicity in idiom flexibility suggests that idiomaticity is more than a matter of mere stipulation or accident; rather, it indicates that there are underlying differences among these idioms revealed in their patterns of flexibility.

One might wonder how the results of this corpus study can be reconciled with the experimental results reported in the previous chapter. Recall that neither experiment found a systematic difference among the idioms investigated. On the contrary, idioms could not be distinguished based on their ability to be passivized, either according to native English speakers' overt acceptability judgments or their underlying intuitions as measured by reading times.

Taken at face value, the experimental and corpus results are in direct contradiction. However, as suggested in Section 2.4, it is likely that the results of both experiments do not actually bear on the hypothesis under consideration. Recall that

subjects in Experiment 1 acknowledged that they were ignoring syntactic structure and focusing on the relevance of the idiom to the accompanying passage. In Experiment 2, I suggested that reading times were not a good measure of grammaticality due to the fact that even canonically nonpassivizable idioms like *kick the bucket* failed to show the expected difference in reading times. In fact, active *kick the bucket* was read more slowly than passive *kick the bucket*, which is the opposite of what was expected.

Taking these facts into consideration, the corpus results seem to provide a good basis for a model of idiomatic representations, in harmony with canonical judgments of idiom manipulability, namely the structural requirement model. This model is the focus of Chapter 4.

## CHAPTER 4: A THEORETICAL MODEL OF IDIOM FLEXIBILITY

### 4.1 Introduction

We have seen in the previous chapter that idiom flexibility is patterned in a very specific way: It is possible for an idiom's appearance in one variation to predict its appearance in another variation. These subset relationships hold among the three argument-structure-type variations. If an idiom occurred in gerundization, it also occurred in incorporation; if it occurred in incorporation, it also occurred in passivization. The other variations under investigation—modification and quantification—did not reliably exhibit such subset behavior.<sup>62</sup>

It should be clear that whatever theory of idiomatic flexibility we adopt must differentiate between the argument-structure-type variations in a way that predicts this subset behavior. This chapter reviews various candidate theories and evaluates their ability to account for these findings. Many proposed theories are unable to make the kinds of distinctions necessary to generate these subset relationships. Other theories, while powerful enough to capture the data, do not provide a motivation for the particular subset behavior that we have observed. Only the structural requirement model (Folli & Harley 2007; Stone 2009; Harley & Stone 2013; Punske & Stone 2014; Stone to appear) correctly predicts this behavior. In a theory where the passivizing projection is outside

---

<sup>62</sup> More precisely, modification was a superset of all four of the remaining variations in COCA but not in Google Books. While interesting, I will not address this for the remainder of the dissertation.

the nominalizing projection, passivizable idioms have to be a superset of gerundizable idioms, and it would be impossible for it to be the other way around.

In what follows, I lay out the case for the structural requirement approach, explaining in detail why the alternative theories fail to adequately account for the empirical observations of Chapter 3. I begin in Section 4.2 by providing a comprehensive overview of extant theories of idiom flexibility, both to situate my own theoretical contribution and, more importantly, to demonstrate that the subset relationships among idiom types are not adequately explained by the other theories. Then, in Section 4.3, I provide a detailed account of the structural requirement approach and show how it is not only compatible with the subset relationships from the corpus study but actually predicts them. Finally, Section 4.4 offers a summary and conclusion of these findings.

## **4.2 Inadequate approaches to idiom flexibility**

### ***4.2.1 Overview***

Extant models of idiom flexibility are many and varied; I review a selection of them below. However, I show that, despite their diversity, none of these theories is able to adequately account for the subset relationships among the idioms that were observed in passivization, incorporation, and gerundization.

The theories fall into six broad categories based on the types of predictions they make regarding idiom flexibility.<sup>63</sup> Some models predict no patterns of flexibility (Section 4.2.2). At the opposite end of the spectrum is one model that predicts that all variations should follow patterned subset behavior (Section 4.2.3). A third type of model predicts a cline of flexibility, from flexible to frozen (Section 4.2.4). The most common type of model, Type IV, predicts two distinct classes of idioms, flexible and inflexible (Section 4.2.5). Finally, some theories can descriptively account for the subset behavior under consideration here, but do so by stipulation rather than actually predicting that it should exist (Section 4.2.6). The remainder of this section discusses how each of these approaches fares when confronted with the subset relationships among the passive, incorporation, and gerundization variations, ultimately showing that they are all inadequate.

#### **4.2.2 *Model Type I: Predicts no regularities***

Many early approaches to idiomaticity (e.g. Weinreich 1969; Katz 1973) relied on the assumption that the locus of idiosyncrasy is the lexicon. Because idioms are, in name as well as behavior, idiosyncratic, these approaches held that an idiom's behaviors should be

---

<sup>63</sup> Note that not all theories of idioms make predictions about idiom flexibility, having as their primary goal some other explanandum. Svenonius (2005), for instance, employs multidominance structures ("banyan trees") as a way of accounting for idioms with discontinuous constituents (see discussion in Section 1.3). Another approach, the selection account of Bruening (2010) and O'Grady (1998), is intended to restrict the class of possible idioms. Both approaches are concerned with generalizations about what makes a possible idiom and how to capture those generalizations within a given theory; neither is concerned with the flexibility of idioms. As such, neither approach is strictly incompatible with the data from this dissertation. Instead, either would need to be supplemented with some means of handling idiom flexibility in order to capture the subset behavior under consideration.

lexically specified. If each idiom's individual behaviors are marked one by one, no regularities are expected; the prediction is that any given idiom could be specified to participate (or not) in any given variation. If such a theory were correct, the subset relationships among passivization, incorporation, and gerundization from the previous chapter would have to be taken as purely coincidental.

Lexical specification approaches generally assume that there are no patterns to idiomatic flexibility—an assumption that has now been shown to be false. Nonetheless, with that assumption as a starting point, it is not hard to see why simple stipulation was a logical solution. Weinreich (1969), for instance, proposes that each idiom is marked in the lexicon with a special syntactic feature for each transformational deficiency—i.e. for each variation that the idiom does not allow.<sup>64</sup> Thus, a verb-object idiom that was entirely inflexible would need to be marked as [-Passive], [-Gerundization], etc. Under such an account, patterns like the one discussed in the previous chapter—where an idiom's occurrence in gerundization implies its occurrence in passivization, for instance—would be unexpected and incredibly surprising, given that there is no underlying basis for the assignment of these syntactic features.

A similar approach from Katz (1973) assigns the syntactic feature [+Idiom] to one or more subconstituents of a given idiom, where each of those subconstituents represents a “transformational deficiency”. Consider the idiom *kick the bucket*. Katz suggests that [<sub>V</sub> kick], [<sub>DET</sub> the], [<sub>N</sub> bucket], and [<sub>NP</sub> the bucket] all bear the lexical feature

---

<sup>64</sup> Chomsky (1981:146, f.n. 94) has a very similar proposal.



[+Idiom]. Katz's account is simpler than Weinreich's in that it requires only a single equipollent feature rather than an ill-defined set of privative features (one for each possible transformation). The lack of syntactic flexibility is accounted for by stipulating that "a transformation does not apply to a phrase marker if the structural change of the transformation specifies a formal operation on a component of the substring that is marked [+Idiom]" (Katz 1973:365). That is, constituents marked with the [+Idiom] feature cannot be targeted by a transformation.<sup>65</sup> Returning to *kick the bucket*, because [NP the bucket] bears the lexical feature [+Idiom], it cannot be targeted by the passive transformation. Although slightly more sophisticated, Katz's account is still entirely stipulative and does not obviously explain systematic relationships between variations like those documented here.

Lexical specification models mostly represent early attempts to account for idiom flexibility. These models simply stipulate which variations an idiom does or does not allow. As such, it is difficult to see how any such account could explain the patterned subset behavior identified in the previous chapter.

#### **4.2.3 Model Type II: Predicts exhaustive subset relationships**

Unlike traditional lexical specification approaches, which assume that idiom flexibility is idiosyncratic, Fraser (1970) acknowledges that there are patterns to idiom flexibility. In fact, Fraser proposes that all idioms exhibit subset behaviors with respect to all

---

<sup>65</sup> In some ways, this account is surprisingly similar to the structural requirement approach detailed in Section 4.3 below. Both approaches rely on the fact that certain constituents are required to appear in a certain configuration in order for the idiomatic interpretation to be licensed.

variations. To account for this assumption, he proposes a “Frozenness Hierarchy” for idioms, ranging from L0 to L6. The variations are arranged on this hierarchy according to Fraser’s observations, and idioms assigned to a given level will allow all variations at that level, as well as all variations at all lower levels.<sup>66</sup> Idioms at the lowest level, L0, allow no variations whatsoever and are therefore completely frozen. Level L6 is unrestricted and is the level of nonidiomatic usage; Fraser claims that no idioms are found at this level.

(47) Fraser’s Frozenness Hierarchy

<b>L6 – Unrestricted</b>	all variations
<b>L5 – Reconstitution</b>	nominal gerundization [Fraser’s “action nominalization”]
<b>L4 – Extraction</b>	passivization, (some forms of) particle movement, preposition preposing
<b>L3 – Permutation</b>	(some forms of) indirect object movement, (some forms of) particle movement
<b>L2 – Insertion</b>	(some forms of) indirect object movement
<b>L1 – Adjunction</b>	gerundive nominalization, adverbial insertion
<b>L0 – Completely Frozen</b>	no variations

---

<sup>66</sup> More precisely, it will allow all variations that can apply to it (e.g. particle shift does not apply to idioms lacking a particle).

Fraser's hierarchy is implemented via lexical features: Each idiom is marked with a feature specifying at which level of the hierarchy it occurs. This allows Fraser to simplify greatly from the large set of features required by Weinreich (1969) (although he still has a larger feature set than Katz (1973)).

It should be easy to see how this hierarchy can account for the subset behavior observed among the argument-structure-type variations. As seen in (47), Fraser positions passivization at L4, below nominal gerundization, which is at L5 on the hierarchy. This suggests that idioms which occur in nominal gerundization should always occur in passivization, since gerundization is higher on the hierarchy, and is consistent with the findings of the corpus studies in Chapter 3. (Object incorporation is notably absent from Fraser's hierarchy.)

Fraser's Frozenness Hierarchy represents an attempt to understand the patterned behavior of idiomatic flexibility, but it has some drawbacks. First of all, it is inconsistent with empirical observation. As Katz (1973) and many others have noted—and as the data from Chapter 3 corroborate—not all variations follow a strict subset pattern. For instance, in the Google Books corpus, neither quantification nor modification appears to be in a subset relationship with any other variation. Thus, while Fraser's Frozenness Hierarchy correctly accounts for passivization and gerundization, it makes incorrect predictions about the behavior of other variations.

Furthermore, like other accounts that rely on lexical specification, Fraser's approach is weakened by its stipulative nature. Under this account, every idiom is marked with a feature specifying its level on the hierarchy, and there is no explanation for

why a given idiom falls at a particular level, or, more relevantly, why the levels are ordered in this particular way.<sup>67</sup>

Fraser's Frozenness Hierarchy is an improvement over the other lexical specification accounts in that it acknowledges the existence of patterns among the syntactic variations. However, this model is too restrictive: It predicts exhaustive subset relationships—that is, subset relationships among all of the variations—where no such relationships exist. Additionally, insofar as it can account for the data, it does so by stipulation; there is no principled reason why the variations occur in the order they do. The Frozenness Hierarchy is based solely on Fraser's observations, with no independent motivation.

#### **4.2.4 Model Type III: Predicts a cline of flexibility**

Pretense is a radically different account of idiomatization proposed by Egan (2008) to account for idiomatic creativity of the type seen in (48) below. Egan claims that idiomatic interpretation is essentially a matter of engaging in highly formalized make-believe. Although this proposal has a lot of empirical coverage, it falls short in surprisingly similar ways to the lexical specification approach. In particular, Egan's account predicts that, while some idioms are more flexible than others, the variations themselves should be randomly distributed over these idioms. That is, Egan predicts a

---

<sup>67</sup> The structural requirement model might seem to wither under a similar criticism; after all, there is no explanation for why *kill the audience* can occur in only passivization, while *steal the show* can occur in both passivization and object incorporation. However, the structural requirement approach ties these differences to independently motivated functional structure within the verb phrase, providing motivation for the various distinctions. See Section 4.3 for further discussion.

cline of flexibility with more and less flexible idioms, but no patterns among the variations.

(48) Figurative modification and extendibility in idioms; from (Egan 2008:391–393)

- (a) The strings we've been pulling to keep you out of prison are fraying badly.
- (b) I know the bathwater was really dirty, but you still shouldn't have thrown the baby out with it.
- (c) Brazil is a team with more rabbits in its hat than most.
- (d) If you let this cat out of the bag, a lot of people are going to get scratched.
- (e) I had my ducks in a row for about a week, but then they just went flapping and squawking all over the park.
- (f) **Tony:** Has he bought the farm yet?  
**Silvio:** Nope. The offer's been accepted and the loan's been approved, but he's taking his time closing on it.

“Pretense” is essentially a formal term for systematized pretending, a framework that accounts for the interactions that take place within an identified make-believe world (e.g. *pretend the cars are buffalo*, to use Egan's example). Within a given pretense, there are principles of generation that determine what is fictional, i.e. what is true in the pretend world (e.g. *if a car honks its horn, pretend that the buffalo bellowed*). Pretense also works in reverse: We can say things about the real world by saying something about the make-believe world (e.g. *There's a stampede!* could mean there are many cars coming down the road).

A pretense account of idioms holds that each idiom has its own associated pretense. Sentences containing idioms are interpreted compositionally, and that interpretation tells the listener what to pretend. The principles of generation aid the listener in determining what to infer about the real world, which ultimately provides the idiomatic truth conditions. For example, if someone says, “Briana kicked the bucket,” the listener interprets the phrase literally, and thus pretends that Briana’s foot has come in contact with an actual bucket. According to the principles of generation, “what it takes for it to be *fictional* that [Briana] kicked the bucket is for [Briana] *actually* to have died” (Egan 2008:387). This provides the following truth conditions: *Briana kicked the bucket* is true iff Briana died.

The idiomatic interpretation is activated via cuing, and use of the paradigmatic form of words serves as a cue to initiate the relevant pretense. Idiom inflexibility is a result of this cuing: “...if you gratuitously alter the form of words, you’re not providing your audience with the right cue” (Egan 2008:399). Thus, variations are only acceptable if they serve some communicative benefit<sup>68</sup>, since the cue will be weaker and the listener will be left to wonder whether the pretense has actually been activated. And word play effects are the result of using non-canonical forms for some reason other than communicative benefit.

Because cuing involves using canonical word patterns, Egan (2008:400) suggests that variation in idiom flexibility is a factor of “the relative importance of the particular

---

<sup>68</sup> Unfortunately, Egan is not at all clear about what he means by the term “communicative benefit”.

form of words as a cue to engage in the appropriate pretense.” Thus, weakly unpredictable idioms (i.e. those that can be more easily interpreted through metaphorical processes, etc.) tend to be flexible and extendible, perhaps because the pretense is easier for the listener to determine.

The pretense approach to idiomatic interpretation assumes that flexibility is strongly associated with predictability, in the sense that a less predictable idiom will require a stronger cue in order for the listener to determine the pretense. However, Egan is unclear about what this means in practice. Nunberg, Sag & Wasow (1994:496–497) are careful to point out that flexibility is independent of transparency, which they characterize as the relative ability of a speaker to understand the motivation behind an idiomatic use.<sup>69</sup> It seems clear, then, that the solution cannot simply be that flexible idioms are more easily or intuitively understood than inflexible ones. This leaves the pretense account with little to differentiate among various types of flexibility in idioms.<sup>70</sup>

---

<sup>69</sup> For example, they claim that *saw logs* is transparent but inflexible. See Section 4.2.5.2 below for a detailed discussion of Nunberg, Sag & Wasow’s idiom decomposition hypothesis.

<sup>70</sup> Wearing (2012) offers several other criticisms of Egan’s pretense account of idiomaticity. First, she claims that the pretense account is overkill, suggesting that simple analogical comparison is sufficient. For instance, if cuing is not for a pretense but rather a “rule schema” that stipulates equivalence, e.g. *pulling strings* = *exerting influence*, then the process of analogy will lead to the interpretation of *fraying strings* as *waning influence*, etc. The elaborate technology of pretense is simply unnecessary. Second, there are empirical concerns. The two-step interpretation process implied by pretense predicts that idioms should be interpreted first literally and then figuratively, but psycholinguistic evidence suggests the opposite (e.g. Ortony et al. 1978; Gibbs 1980; McGlone, Glucksberg & Cacciari 1994). Wearing is also concerned with overgeneration and is dissatisfied with Egan’s explanation that cuing (via canonical form) and communicative benefit conspire to give us the attested range of (un)acceptable variations. She presents the following contrasting pairs of idioms as counterexamples:

- i. (a) *The ice was broken by Paulie.*
- (b) *\*The bucket was kicked by Stevie.*

In the end, the predictions of this account are unclear due to the vagueness of the notion of “pretense”. However, what seems clear is that Egan predicts a scale of flexibility, according to the “relative importance” of the canonical form as a cue for the associated pretense, however that importance is measured. Taken at surface value, this predicts that some idioms will be completely flexible, some will be completely inflexible, and there should be a range of more or less flexible idioms between these two extremes. What it does not account for is structurally-conditioned subset behavior—that is, this approach cannot account for the fact that gerundizable idioms are a subset of passivizable idioms.

This type of behavior could be accommodated if Egan had not only a scale of idiom flexibility—the aforementioned scale based on the importance of cuing to an idiom’s interpretation—but also a second scale that ordered the variations, reminiscent of Fraser (1970)’s Frozenness Hierarchy. Provided that passivization, gerundization, and incorporation were positioned properly on the scale, the subset behavior could be accounted for. However, as with Fraser (1970), this version of the pretense approach would run into trouble with the other variations, such as quantification, which do not seem to exhibit the same kind of subset behavior.

- 
- ii. (a) *Jane had a bone to pick with Susan, and Anne had one to pick with Ian.*
  - (b) *\*Tony shot the breeze with Junior, and Paulie shot it with Silvio.*

I agree with Wearing’s judgments, but neither she nor Egan provides an explanation of “communicative benefit”, so it is difficult for me to judge whether the variations have equal communicative benefit in these examples. Nonetheless, she makes a convincing argument that pretense is both unnecessary and insufficient to account for idiom flexibility.



While Egan's pretense approach purports to explain radical flexibility in idioms, it actually fails on a much more basic level. Egan's explanation of what licenses flexibility is vague and lacking specificity. Because of this, the predictions of this approach are not entirely clear, but Egan seems to be fundamentally unable to account for the subset relationships of Chapter 3 without a major modification to his account, a modification that would ultimately fare no better at accounting for the full range of data.

#### **4.2.5 *Model Type IV: Predicts two classes of idioms***

So far, we have seen theories that make three types of predictions: no patterns to idiom flexibility, exhaustive (but stipulated) subset relationships, and a cline of flexibility. A fourth type of theory—and by far the most common in the literature—predicts essentially two classes of idioms, flexible and inflexible. Although there are myriad specific implementations, the basic idea is that some idioms are syntactically or semantically or conceptually structured in a way that allows them to undergo variations, while others lack the necessary kind of structuring. The intuitive appeal of this approach gives it traction, but it cannot be the full story. The subset behavior among idioms that occur in passivization, gerundization, and incorporation shows that there must be at least four different types of idiom displaying varying degrees of flexibility. The approaches in this section cannot account for such diversity among flexible idioms.

#### 4.2.5.1 Quasi-Inference

The quasi-inference account of Pulman (1993) takes a pragmatic approach to idiomaticity, proposing a model that relies on patterns of inference. Pulman's approach foreshadows the more familiar idiom decomposition hypothesis of Nunberg, Sag & Wasow (1994), suggesting that idioms are either flexible or inflexible. In spite of his attempts to be more explanatory and less stipulative than earlier accounts, Pulman's quasi-inference approach is unable to adequately predict the subset behavior of certain idioms with respect to the argument-structure-type variations.

This approach is intended to account for the widely variable syntactic behavior of idioms, as exemplified in (49-51) below.

(49) Syntactic versatility in idiomatic *turn the tables*

- (a) He turned the tables on me.
- (b) The tables have turned.

(50) Syntactic versatility in idiomatic *let the cat out of the bag*

- (a) Now he's let the cat out of the bag.
- (b) Now the cat's out of the bag.

(51) Syntactic versatility in idiomatic *tie up loose ends*

- (a) I've got some loose ends to be tied up.
- (b) I'm tying up a few loose ends.
- (c) A few loose ends need tying up.

In each of these examples, the (a), (b), and (c) sentences cannot be related derivationally, which was problematic in the theory of Pulman’s time. Furthermore, he claims that some idioms are not even constituents.<sup>71</sup> This poses an issue for other models of idiomaticity that try to state idiom meanings at a single level of representation, since there is no single level that unites these different forms.

Pulman’s solution is “quasi-inference”: The relationships among these sentences are not syntactic but rather pragmatic, computed by an “inference mechanism”. For example, (50a) entails (50b). Pulman proposes that it is inferential relationships between the forms of sentences that license idiomatic interpretations, as follows. If the literal interpretation of an idiom is among the inferences allowed by a given sentence, then the corresponding idiom rule, housed in the “inference mechanism”, is triggered. This allows the associated figurative meaning as an alternative interpretation. As an illustration, consider Pulman’s sentence *John let the cat out of the bag*. This sentence entails *The cat is out of the bag*, which in turn matches the left-hand side of an idiom rule. The match has the effect of triggering the rule and activating the right-hand side of the rule, which is the (optional) meaning ‘revealed a secret’.<sup>72</sup> On the other hand, if the

---

<sup>71</sup> Pulman cites several questionable examples of subject-verb idioms (see Harley & Stone (2013) for some counterarguments), but there are other reasons to believe that the elements of an idiom do not always form a constituent, e.g. the “proxy clause” data from Higgins (1981) (see Section 1.4.1, example (22)).

<sup>72</sup> More formally (Pulman 1993:262–263):

*From:* John let the cat out of the bag  
 $\text{let}(\text{John}, \exists c, b. \text{cat}(c) \wedge \text{bag}(b) \wedge \text{out-of}(c, b))$   
*via:*  $\exists c, b. \text{cat}(c) \wedge \text{bag}(b) \wedge \text{out-of}(c, b)$   
*and:*  $\forall x, y. \text{cat}(x) \wedge \text{bag}(y) \wedge \text{out-of}(x, y) \approx \exists a, z. \text{secret}(z) \wedge \text{revealed}(a, z)$   
*to:*  $\exists a, z. \text{secret}(z) \wedge \text{revealed}(a, z)$

direct object were *the dog* rather than *the cat*, the rule would not be triggered and no idiomatic meaning could be accessed.

The notion of inferences—rather than exact pattern-matching—accounts for the attested idiom flexibility. However, Pulman must also avoid “false positives”, such as having a sentence like *We watched the chandelier being switched on* trigger the idiom *We saw the light*. In order to do so, he introduces “lexical indexing”, whereby he requires certain words from the idiom to be present in the target sentence in order to trigger the rule. To continue with the previous example, the arbitrary idiom rule #45 might be associated with the set of lexical items {cat, out, of, bag}. Rule #45 would only apply if this set is a proper subset of the words in the target sentence. So *Let the cat out of the sack*, although it entails *Let the cat out of the bag*, would fail to trigger the idiom rule.

When it comes to variation within idiom flexibility, Pulman holds that the grammar should, in principle, allow any idiom to undergo any variation. Certain variations are ruled out, he claims, by information structuring. For Pulman, this amounts to whether or not discourse entities can be associated with elements of the idiom. He contrasts *kick the bucket*, where *the bucket* seems to have no corresponding associated meaning, with *spill the beans*, where it is possible to associate *spill* with ‘reveal’ and *the beans* with ‘a secret’. This approach, although not fully worked out, prefigures that of Nunberg, Sag, and Wasow (1994), discussed in detail in Section 4.2.5.2 below.

In terms of flexibility, Pulman’s approach predicts just one uniform class of flexible idioms, with exceptions based on the nuances of information structure. It seems from his example that he is suggesting that some idioms are wholly flexible while others

resist all variation due to lack of meaning association between the literal and figurative elements, in which case two classes of idioms would emerge. What is unexpected is subset behavior within the class of flexible idioms among three particular variations. Unless these variations can be tied in some principled way to differences in information structure, Pulman has no way of accounting for the data presented in the previous chapter.

It is possible in principle to divorce the “quasi-inference” component of Pulman’s approach from the information structuring approach to flexibility. In this case, quasi-inference actually makes no predictions regarding the syntactic flexibility of idioms; this sub-version of Pulman’s account would be grouped with the other theories that make no predictions (see footnote 63). It is the information-structuring component that makes incorrect predictions about idiom flexibility.

#### *4.2.5.2 Idiom Decomposition*

Probably the most well-known theory of idiomatic flexibility is laid out in Nunberg, Sag & Wasow (1994) (henceforth NSW) and will be referred to here as *idiom decomposition* (see also Nunberg 1978; Wasow, Sag & Nunberg 1984).<sup>73,74</sup> According to this idea, some idioms show a kind of post-hoc compositionality whereby elements of the figurative meaning can be mapped onto literal elements in the expression. Because of

---

<sup>73</sup> Many researchers have adopted similar approaches, assuming idiom decomposition as a given (e.g. Pulman 1993; Ruwet 1991).

<sup>74</sup> In previous work (Stone 2013; Punske & Stone 2014; Stone 2014; Stone to appear), I have referred to idiom decomposition as the semantic mapping hypothesis.

this mapping, the pieces of the idiom are meaningful and can undergo semantic variations. Due to the fact that idioms are either decomposable or not, this kind of theory predicts all-or-nothing syntactic flexibility, not the subset behavior revealed by the experiments in the previous chapter. The partial function approach, a precursor to idiom decomposition, is briefly discussed below; however, it is similarly ill-equipped to deal with subset behavior within the class of flexible idioms.

According to NSW, there are two basic classes of idioms: An idiomatic phrase (IDP) is an idiom whose meaning is distributed over the entire phrase (e.g. *kick the bucket*), while an idiomatically combining expression (ICE) is an idiom whose meaning is doled out to its component parts (e.g. *spill the beans*). IDPs are understood in the same way that we traditionally understand idiomatic expressions; they can be regarded as a single unit with a single associated meaning. In contrast, ICEs are built up out of smaller semantic units that are associated with the syntactic units of the idiom. For instance, in the idiom *spill the beans* ‘divulge a secret’, the figurative meaning ‘divulge’ is mapped onto the syntactic verb *spill*; similarly, the figurative meaning ‘a secret’ is mapped onto the syntactic object *the beans*. This allows NSW to ascribe a kind of post-hoc compositionality to ICEs. Although a listener unfamiliar with idiomatic *spill the beans* would not be able to ascertain the meaning of the whole from the meanings of the parts (because the meanings assigned to the parts are not the typical nonidiomatic ones), NSW maintain that the meaning of the whole is indeed composed of the meanings of its parts, albeit in this nonstandard way. As such, a kind of weak compositionality is preserved, for ICEs only.

Because the meaning of the whole idiom is distributed among its parts, the idiom is allowed to undergo semantic variations, i.e. those which—according to NSW—rely on the assignment of meaning to the syntactic units undergoing movement. ICEs can therefore participate in semantically-motivated variations such as passivization (52a), topicalization (52b), internal object modification (52c),<sup>75</sup> object quantification (52d), VP-ellipsis (52e), and pronominal reference (52f).

(52) Variations allowed by ICEs (from NSW:522, 500-502)

- (a) Too much advantage has been taken of the homeless.
- (b) Those windmills, not even he would tilt at.
- (c) kick the filthy habit
- (d) touch a couple of nerves
- (e) My goose is cooked, but yours isn't.
- (f) We thought tabs were being kept on us, but they weren't.

IDPs, on the other hand, do not license any of these variations because their parts do not carry meaning. Rather, it is the idiom as a unit which carries the meaning as a whole. As a consequence, IDPs are syntactically inflexible.

At its most basic, then, this theory predicts a dichotomy between ICEs and IDPs, which will be reflected in their behavior. ICEs should allow all semantic variations, and IDPs should allow none of them. It should be immediately clear that these predictions

---

<sup>75</sup> In the sense of Ernst (1981).

are incompatible with the results of the corpus studies from the previous chapter: The subset relationships involve three variations which are allowed by some but not all idioms. If occurrence in variations were all-or-nothing, then this subset behavior would be unexpected.<sup>76</sup>

However, to abandon idiom decomposition at this point would be premature, as Nunberg (1978) and NSW allow for some deviation from this model due to semantic variation. For instance, Nunberg (1978) discusses several reasons why a decomposable idiom (similar to NSW's ICE) might not be passivizable. First, the verb—or particular use of the verb—might not passivize at all, even in literal contexts. For instance, while some uses of *hit* are passivizable, others are not (53). Idioms that share corresponding meanings of the verb seem to inherit the passivization properties of the verb as well (54).

(53) (Non)passivizability of different uses of the verb *hit* (from Nunberg 1978:234)

- (a) The lucky number was hit.
- (b) \*The floor was hit by the paper I dropped.
- (c) \*The Cross-Bronx was hit by the travelers at rush hour.

---

<sup>76</sup> There are (at least) two reasons why a proponent of idiom decomposition might call the results from the previous chapter into question. The first is that the results were unable to distinguish internal from external modification, while NSW are insistent that only internal modification distinguishes ICEs from IDPs. The second is that I have not shown that either object incorporation or gerundization is a semantic variation, in the sense intended by NSW. Because they provide no means for assessing whether a variation is semantic or syntactic in nature, I have little more to say about that. However, I have already argued (see Chapter 3) that all three variations—passivization, incorporation, and gerundization—are similar in their effect on the argument structure of a sentence and thus should receive a similar treatment.



(54) (Non)passivizability of idiomatic uses of the verb *hit* (from Nunberg 1978:233)

- (a) The jackpot was hit by the little old lady.
- (b) \*Once the sack was hit, they all slept soundly.
- (c) \*The headlines were hit by the story the very next day.

Other more general properties of passives are similarly inherited by idioms, such as an apparent prohibition against coreference between (a subpart of) the idiomatic object (promoted to subject in the passive) and the subject (demoted to the *by*-phrase), accounting for the oddity of such sentences as *\*John's mind was made up by him* (Nunberg 1978:238). Another apparently general prohibition exists against passivizing an indefinite object over a definite subject, resulting in ungrammatical sentences such as *\*Time was marked by John* (Nunberg 1978:238). In each case, the inability to passivize the idiom can be seen as part of a prohibition against passives more generally.

Nunberg (1978) also has an entire class of idioms that can be modified but not passivized, what he calls “abnormally decomposable” idioms. Like other decomposable idioms (i.e. ICEs), we can identify the referents of the constituent terms; however, unlike other decomposable idioms, we only understand the relation that is being invoked by virtue of some conventional metaphor. In *make tracks* ‘to leave in a hurry’, for example, the literal-to-figurative mapping is relatively clear only if the idiom is understood metaphorically as a relation between the leaver and the path (cf. Nunberg 1978:224). In this case, then, the mapping is different: the elements of this idiom do not map directly onto elements of the figurative meaning but instead map onto elements of the relation by which—via conventional metaphor—the figurative meaning is identified. Thus,

passivization is prohibited for the same reason it is disallowed in nondecomposable idioms (i.e. IDPs), namely, there is no direct mapping between the literal and figurative components of meaning (and, thus, no semantic referent for the object).

It is fair to ask whether the kinds of “exceptions” discussed here can redeem the idiom decomposition story. In other words, is it possible that the subset relationship could be accounted for by idiom decomposition if we accepted Nunberg (1978)’s class of abnormally decomposable idioms and, additionally, allow for principled exceptions based on independent properties of the verb? I believe that the answer is no. First of all, Nunberg’s class of abnormally decomposable idioms consists of those which appear to be decomposable (i.e. they are ICEs) yet fail to passivize. He claims that this is due to the fact that “the object NP of abnormally decomposable idioms does not itself refer to some component of the idiomatic referent, but only to a component of the relation by means of which that referent is conventionally identified” (Nunberg 1978:228). That is, object NPs of abnormally decomposable idioms do not refer in the same way those of normally decomposable idioms do. Consider again *make tracks*, the object of which has a clear referent. However, according to Nunberg, this is not a referent of the literal meaning itself, but rather a referent of some relation from the figurative to literal meaning. According to NSW, then, *make tracks* and other abnormally decomposable idioms are not likely to undergo any of the argument-structure-type variations, which require semantic reference. Thus Nunberg’s abnormally decomposable idioms should be completely inflexible, not exhibit partial flexibility, at least with respect to the relevant variations.

In contrast, independent properties of the verb might give rise to some more unique patterns of flexibility. However, because these are idiosyncratic properties of the verb itself, they would not be expected to result in the patterned subset behavior seen here. Only if there were three independent classes of verbs that corresponded with these subset behaviors—i.e. one class that can occur in passive but not object incorporation or nominal gerundization, one class that can occur in passive and incorporation but not gerundization, and one class that can occur in all three—and only if the verbs of each class corresponded to the (interpretations of the) verbs in the idioms that shared each of these three patterns, respectively, would there be evidence that independent properties of the verbs, rather than some property of the idioms themselves, is affecting their flexibility. This does not appear to be the case, as the verbs do not always exhibit the same limited syntactic behavior as the idioms; that is, in most cases, these variations are fine on the non-idiomatic reading (55-57).

(55) Idiomatic vs. literal *kick the bucket*

- (a) #The bucket was kicked by Briana. (idiomatic)
- (b) The bucket was kicked by the angry cow. (literal)

(56) Idiomatic vs. literal *hit the sack*

- (a) #The sack was hit by the weary travelers. (idiomatic)
- (b) The large sack of concrete was hit by the careening truck. (literal)

(57) Idiomatic vs. literal *rock the boat*

- (a) #Lidia's rocking of the boat caused panic at work. (idiomatic)
- (b) Lidia's rocking of the boat almost caused it to capsize. (literal)

It should be clear from these examples that properties of the verb alone are not governing the variations in which the idiom can appear.

While the idiom decomposition hypothesis faces a major hurdle in being unable to account for the subset patterns seen in the corpus experiments presented above, it also has an even bigger problem: testability. As it is outlined in both NSW and Nunberg (1978), “idiom decomposition” is just a description of the intuition that certain syntactic variations have their root in something about the semantic properties of the idiom. However, no independent test or tool is suggested to diagnose those semantic properties; they do not correlate with anything other than the syntactic behaviors themselves. Consequently, idiom decomposition as presented in NSW is just a description, an idea about a difference between idioms that exhibit certain syntactic behaviors and idioms that do not.

NSW claim that an idiom's status as an ICE or an IDP determines its syntactic flexibility. In other words, if an idiom is an ICE, it will be syntactically flexible; if it is an IDP, it will be inflexible. However, they provide no criteria for determining whether an idiom is an ICE or an IDP aside from its flexibility. Without an independent means of establishing an idiom's status as an ICE or IDP, it is impossible to test the hypothesis that idiom decomposability is a key factor in determining an idiom's flexibility; in fact, the whole enterprise is quite circular.

For NSW, idiom decomposition relies on a homomorphism between certain properties of the literal and idiomatic interpretations (Nunberg, Sag & Wasow 1994:504). This homomorphism is elaborated for the case of *pull strings*, but the reader is left to guess how one might determine whether or not other idioms exhibit a similar homomorphism. Furthermore, NSW show only how an idiom that (they assert) exhibits such a homomorphism is syntactically flexible, not how this homomorphism plays a role in determining syntactic flexibility; the latter is crucial in having an explanatory rather than a merely descriptive theory.

Although NSW are not explicit about how to determine whether or not an idiom has the requisite homomorphism, aside from observing its flexibility, one might try to operationalize the intuition behind the proposal. One possibility is to propose that it is the idiom's adicity—which must allow for a one-to-one mapping between the literal and figurative elements of the idiom—that determines whether or not an idiom is decomposable. More explicitly, if the idiom's literal meaning and figurative interpretation share the same adicity, the mapping applies and the idiom is an ICE. For example, *kick the bucket* is syntactically transitive, with an adicity of two, but the gloss of its figurative interpretation, 'die', is intransitive, with an adicity of one; hence no mapping can apply. In contrast, *pull strings* is syntactically transitive, as is its interpretation, 'exert influence'. As both expressions have an adicity of two, we could hypothesize that a one-to-one mapping from parts of the literal expression to parts of the figurative expression could apply in this case. Under this proposal, idiom

decomposability could be motivated independently of flexibility by comparing the adicity of an idiom's figurative and literal interpretations.<sup>77</sup>

However, we now have a new problem: How does one determine the adicity of an idiom's figurative interpretation; or, perhaps more aptly, how does one determine an idiom's "true" figurative interpretation? Two different people are likely to provide two different paraphrases for a given idiom. Even dictionaries provide different definitions for the same idiom, sometimes with different adicities. (In the case of transitive *bite the bullet*, for example, the figurative meaning could be paraphrased as intransitive 'persevere', with no mapping predicted; or transitive 'accept consequences', with mapping predicted; or as the V+PP expression 'proceed despite negative consequences', with unclear predictions for mapping.) What kinds of definitions should be adopted if we are to try to implement an independent mapping criterion for determining whether an idiom is an ICE or IDP?

Assuming that lexicographers are able to give precise and apposite characterizations of the figurative meanings of idioms, I used the Oxford English Dictionary and the Cambridge Dictionary of Idioms to see whether dictionary definitions could provide independent criteria with which to operationalize the NSW idiom

---

<sup>77</sup> This is similar to a proposal by Newmeyer (1974). In his account, an idiom's flexibility is determined solely by its semantic properties: A flexible idiom is one with a transitive paraphrase (i.e. its figurative meaning is transitive), and an inflexible idiom is one with an intransitive paraphrase. For instance, the figurative meaning or paraphrase of *bury the hatchet* is 'make peace', which can be passivized. From this, Newmeyer says, we can predict that idiomatic *bury the hatchet* should be passivizable.

Nunberg (1978) points out that this cannot be the whole story, citing several exceptions (e.g. *throw in the sponge* is passivizable, although 'resign' is intransitive; and *have a shot at* cannot be passivized, even though 'try' can (Nunberg 1978:212–213)).

decomposition story. Each of the 38 idioms from the Google Books corpus study in the previous chapter was coded as an ICE or an IDP based on its dictionary definition: If the number of verbal arguments in the definition (a proxy for its adicity) matched the number of verbal arguments in the idiom, it was coded as an ICE; otherwise, it was coded as an IDP (58)-(59).

(58) Coding for *have the final say*

definition: ‘to have the authority to make decisions’

*have* maps to *have* and *the final say* maps to *the authority to make decisions* and the subject of the former maps onto the subject of the latter

∴ *have the final say* is an ICE

(59) Coding for *drop the ball*

definition: ‘to fail to keep working to reach a goal’

*drop* maps to *fail*, perhaps, but there is no nominal argument onto which to map *the ball*, and *the ball* cannot be conceived of as *to keep working to reach a goal*

∴ *drop the ball* is an IDP

I then tested to see whether an idiom’s decomposability was a good predictor of any of the variations under consideration, using logistic regression. As shown in Table 4.1, an idiom’s coding as an ICE or IDP did not predict its occurrence (or failure to occur) in any of the variations under consideration.

<b>Independent Variable</b>	<b>Dependent Variable</b>	<b>Beta-coefficient</b>	<b>SE</b>	<b>p-value</b>
<b>Quantification</b>	Decomposability	-0.1335	0.7480	$p=0.858$
<b>Modification</b>	Decomposability	0.08701	0.68692	$p=0.899$
<b>Passivization</b>	Decomposability	0.08701	0.68692	$p=0.899$
<b>Object Incorporation</b>	Decomposability	0.2342	0.6854	$p=0.733$
<b>Nominal Gerundization</b>	Decomposability	-0.08701	0.68692	$p=0.899$

**Table 4.1: Decomposability as a predictor of idiomatic flexibility**

This experiment shows that dictionary definitions do not provide an independent criterion for determining an idiom's decomposability status. Even if an idiom's decomposability should be established in this way (a questionable premise), this information does not help predict the idiom's occurrence in any of the variations under consideration here.

We have seen that trying to match the adicity of the literal and figurative interpretations of an idiom is not a reliable way of determining whether that idiom exhibits the homomorphism required to be considered an ICE. However, Nunberg (1978) suggests another possibility. In his conception, the key is general speaker agreement. According to Nunberg, a given idiom is decomposable only if the following two conditions are met: (1) it is possible to decompose the idiom's figurative meaning into an entity and a relation such that there is a function from each to a distinct component of the literal expression; and (2) speakers will generally agree on the most appropriate functions.



Nunberg contrasts three familiar idioms that all mean ‘die’, showing how they fare differently in this approach (Nunberg 1978:222–224). Consider the idiom *give up the ghost*. If death is seen as a two-place relation between “persons and their spirits”, it is not hard to determine a function from ghosts to spirits, and another from giving something up to releasing one’s hold on one’s spirit. With *shuffle off this (one’s) mortal coil*, the two-place relation is between “souls and bodies”. Thus, we get a function from mortal coils to bodies and another from the act of shuffling off to the soul “discarding” the body. These two cases differ crucially from that of *kick the bucket*, according to Nunberg, in that “we cannot suppose that speakers will be able to agree to analyze dying into a relation and entity such that there is a function to each of them from the components of *kick the bucket*” (Nunberg 1978:223–224). What Nunberg seems to be saying here is that there are no readily apparent literal-to-figurative functions that the majority of speakers would agree on in the case of *kick the bucket*; whereas, with *give up the ghost* and *shuffle off this mortal coil*, the literal-to-figurative functions are obvious enough that general speaker agreement can be assumed.

This account of idiom decomposition, while much more clearly articulated than that found in NSW, does little to fix the original problem, instead simply shifting the ambiguity up a level. Under NSW’s approach, it was impossible to identify an independent means of determining whether an idiom exhibited the requisite homomorphism. The solution proposed by Nunberg (1978) is that speakers will generally agree on whether or not such a homomorphism exists; but now we have no way of determining whether or not there is general speaker agreement. I am aware of no

independent metric of speaker agreement, nor does Nunberg offer one. He simply asserts that in certain cases speakers are likely to agree, while in other cases they are not. This is effectively similar to the claim in NSW and suffers from the same problems.

It should by now be clear that idiom decomposition is not a viable approach to idiomatic flexibility, given the corpus results from the previous chapter. On the one hand, idiom decomposition itself is just a descriptive observation, lacking the independent motivation required to achieve explanatory adequacy. Further, this approach lacks the formal mechanisms needed to account for the syntactic subset behavior that idioms exhibit.

A precursor to idiom decomposition is the partial function approach of Gazdar et al. (1985), situated within the framework of Generalized Phrase Structure Grammar (GPSG). This approach shares the central claim of idiom decomposition, namely that (for a certain class of idioms) the parts of an idiom—not just the idiom as a whole—have meaning. The semantic relationship between the flexible parts of an idiom is represented using partial functions. Consider *spill the beans*. If **spill'** is the literal denotation of *spill*, then **spill''** is the denotation of idiomatic *spill*, meaning roughly 'divulge'. The predicate **spill''** is a partial function because it takes as an argument only **beans''** (the denotation of idiomatic *beans*, meaning roughly 'information'). NSW point out a subtle difference between this approach and their own: Gazdar et al. "claimed that there is a semantic incompatibility between the idiomatic *the beans* and any verb other than the idiomatic *spill*. While the present analysis requires the idiomatic NP and the idiomatic verb to co-occur for semantic reasons, it does not exclude the possibility that an idiom chunk could

also be semantically compatible with other expressions” (Nunberg, Sag & Wasow 1994:505). Thus, using partial functions, each idiom chunk must be semantically specified as to which other chunk(s) it can occur with; with idiom decomposition, the restriction is looser, requiring co-occurrence for interpretability but not strictly prohibiting idiom chunks from occurring elsewhere.<sup>78</sup>

Like idiom decomposition, the partial function approach fails to account for the subset behavior observed here. It predicts the same two classes of idioms—flexible and inflexible—with no (patterned) diversity in the class of flexible idioms.

#### **4.2.6 Model Type V: *Anything is possible***

In the previous section, we saw that many well-known approaches to idiom flexibility predict two classes of idioms—flexible and inflexible—and are therefore unable to account for subset behavior of the kind exhibited by the passivization, incorporation, and gerundization variations. We have also discussed one theory that overgenerates the subset behavior, one that predicts an unpatterned cline of flexibility, as well as those that predict no patterns of flexibility at all, all of which are also inadequate.

Finally, we turn to a fifth type of theory, one that is descriptively much more powerful than those previously discussed. The models outlined in this section have extensive theoretical machinery, which enables them to account for the observed subset

---

<sup>78</sup> Gazdar et al. claim that the partial function approach predicts the existence of families of idioms (e.g. *pack a punch/wallop* and *stop/turn on a dime*) because idiomatic arguments can be specified by more than one predicate, and idiomatic predicates can specify more than one argument. However, it is not clear how this is actually predicted to occur by their theory. The phenomenon can be accounted for descriptively within their framework, but nothing about the framework actually predicts its existence.

relationships. However, I show that these theories fail to reach the level of explanatory adequacy, as their descriptive power is accomplished by means of stipulation. As such, while they are better suited to the task than the theories discussed above, they are still insufficient.

#### *4.2.6.1 Synchronous Tree Adjoining Grammar*

Abeillé (1995) provides an account of idiom flexibility within the framework of Synchronous Tree Adjoining Grammar. For her, idioms are stored as whole syntactic units with their associated semantic representations, and flexibility is licensed via inheritance from non-idiomatic supertypes.<sup>79</sup> This arrangement allows her to account for the fact that idioms tend to be more flexible than not. Although Abeillé does not specifically address the possibility, Tree Adjoining Grammars have type hierarchies that provide a mechanism for capturing subset relationships such as the ones observed among the argument-structure-type variations. However, Abeillé's account stops short of providing an explanation for this behavior—why, for instance, passivation should be a superset of gerundization and not the other way around.

Abeillé maintains that idioms are, by and large, noncompositional (contra Nunberg, Sag & Wasow 1994). They are stored as extended elementary trees in the lexicon, along with their associated semantic representations. Because idioms inherit

---

<sup>79</sup> In TAG, syntactic structures (“treelets”) are stored in a type of lexicon. Different treelets can be more or less general, so, for example, intransitive VP treelets are a subtype of a more general VP treelet type. Subtypes of treelets have all the properties of their supertypes, plus some more specific ones that distinguish them from their supertype. See Abeillé (1995) for an overview of TAG. This is similar to the inheritance structure of more familiar HPSG-type theories (Sag, Wasow & Bender 2003).

from their compositional counterparts (e.g. transitive and intransitive verbs), they are systematically flexible, in spite of their noncompositionality.

Abeillé's system reflects the results of her investigation of more than 20,000 French idioms. She found that "frozenness is the exceptional case...and syntactic flexibility, the general one" (Abeillé 1995:18). Furthermore, she found systematic relationships between variations; for example, "[c]left extraction of an idiomatic NP (or PP) is directly determined by the insertion of a modifier (cf. example 60) or a free determiner, as in (61), on this idiomatic element" (Abeillé 1995:19).<sup>80</sup>

(60) Cleft extraction with a frozen determiner and a free modifier

C'est le taureau **des privatisations** qu'il a pris par les cornes de l'actionariat populaire.

'It is the privatization bull that he took by the horns of popular shareholding.'

(From Abeillé 1995:20)

---

<sup>80</sup> Note that this is a stricter relationship than the subset relationships identified in the previous chapter. Here, Abeillé is claiming that an idiom must not only allow modification or a free determiner in order to occur in clefting, but that it must actually display a free modifier or free determiner to license the clefting variation. This seems natural in that clefting involves a relative operator, which means that it should relate to determiner variation in some way; see discussion of a related point regarding the passivization and quantification variations in Section 3.3.2.1 above.

## (61) Cleft extraction with a free determiner

C'est sur **ton** dos que Jean a cassé du sucre.

(lit. it is on your back that Jean broke sugar.)

'It is on you that Jean put the blame.'

(From Abeillé 1995:19)

This and similar systematic relationships led Abeillé to posit that inflexibility in idioms is marked grammatically, and she accomplishes this marking via stipulations on idiomatic subtypes. Continuing with the clefting example, the idiomatic clefted subtype is stipulated to have  $\langle \text{det} \rangle = -$  and  $\langle \text{modif} \rangle = -$  for its bottom features and  $\langle \text{det} \rangle = +$  and  $\langle \text{modif} \rangle = +$  for its top features. This feature clash has the effect of forcing either a free determiner or a free modifier in the clefted structure.<sup>81</sup>

Based on our discussion, it should be clear that this is the kind of theory that can capture subset relationships like the ones among the argument-structure-type variations. If idiom subtypes can be marked with specific requirements that affect their occurrence in a given variation, then nested subtypes should give the desired subset behavior. However, in an approach like Abeillé's, these subset relationships are simply stipulated; there is no principled reason why, for instance, passive should apply more freely than gerundization and not the other way around. (In this regard, it is similar to Fraser

---

<sup>81</sup> Top and bottom features unify at each node at the end of a derivation. If there are incompatible top and bottom values at a given node, the clash can only be resolved by adjunction of an auxiliary tree at that node (Abeillé 1995:26–27).

(1970)’s approach, outlined above.) A better theory would have an explanation for the ordering of the hierarchy, not simply be able to capture its existence.

The Synchronized Tree Adjoining Grammar model of Abeillé (1995) has the theoretical machinery to capture almost any observable idiomatic behavior, including subset behavior. However, such behavior is unmotivated and stipulative under this approach; there is no principled reason why gerundization should be a subset of passivization and not the other way around.

#### 4.2.6.2 *Sign-Based Construction Grammar*

Sag (2012)’s Sign-Based Construction Grammar (SBCG) is a blend of Head-Driven Phrase Structure Grammar (HPSG; Pollard & Sag 1994; Sag, Wasow & Bender 2003) and Berkeley Construction Grammar (Fillmore 1988; Kay & Fillmore 1999; Michaelis & Ruppenhofer 2001). Sag emphasizes that there is no “unified analysis” of idioms, insisting that the grammar treats completely frozen idioms differently than “semi-fixed expressions”, and those differently from completely flexible ones. Ultimately, he describes three classes of idioms, with varying levels of flexibility. Strictly interpreted, this predicts that there should be three types of idioms: completely inflexible, nearly or partially inflexible,<sup>82</sup> and completely flexible. This is inconsistent with the idiomatic subset behavior of the argument-structure-type variations. Nonetheless, like

---

<sup>82</sup> Note that Sag’s “semi-fixed expressions”, which comprise this nearly inflexible category, are what most others consider completely frozen, allowing only inflectional variants and external modification. This type of modification, although located internally to the NP object, scopes over the entire VP or proposition rather than the NP object alone, e.g. *Bob kicked the proverbial bucket*. See Ernst (1981) for extensive discussion.

Synchronous Tree Adjoining Grammar, SBCG is a flexible theory with the mechanisms in place to account for subset behavior. Yet, this same flexibility means that SBCG, again like Synchronous Tree Adjoining Grammar, is unable to predict specific patterns of subset behavior and, thus, cannot provide an explanation for the observed patterns.

For Sag, “fixed expressions” (e.g. *by and large*, *in short*, *jack of all trades*, *every which way*) are fully lexicalized “words with spaces”.<sup>83</sup> These idioms are treated as single listemes, each with a non-singleton list as the FORM value.<sup>84</sup> There is little else to say here; fixed expressions are treated simply—as simply as their behavior and the grammar will allow.

In contrast, due to the complexities of their behavior, flexible idioms must be treated differently by the grammar. Like Nunberg, Sag & Wasow (1994), Sag maintains that these idioms are semantically decomposable. In order to accommodate this intuition, he proposes that the different components (e.g. the verbal head and its dependents) are encoded via separate listemes. For example, the idiom *pull strings* is composed of two listemes: idiomatic *pull*, meaning ‘manipulate’, and idiomatic *strings*, meaning ‘connections’. In order to prevent these listemes from occurring in illicit contexts,

---

<sup>83</sup> Heidi Harley (p.c.) remarks that this account is implausible if it means these idioms should behave like words throughout the grammar. These expressions have the prosody of phrases, not words, and prosody is determined by syntactic structure. Therefore, even these “fixed expressions” must have structure beyond the word level.

<sup>84</sup> The FORM value specifies the morphological content of the listeme. (“The value of the feature FORM is of type *morphological-object* (*morph-obj*); these are the elements that will be phonologically realized within the sign’s PHON value....FORM values [are] lists whose members include stems and more complex entities built up from stems” (Sag 2012:79).) In the case of fixed expressions, the idiom is comprised of a single listeme whose morphological content consists of more than one element.



idiomatic *pull* specifies the LID, or semantic value, of its direct object. For idiomatic *pull*, the LID of its direct object is “*i-strings<sup>connections</sup>-fr*”, which is an identifier unique to idiomatic *strings*.<sup>85</sup> Finally, Sag uses the familiar HPSG mechanism of type inheritance to ensure that flexible idioms allow variations like passivization. Because idiomatic *pull* is of the type *strict-transitive-verb-lexeme*, it will inherit from its mother type *transitive-verb-lexeme* the ability to occur in the passive variation. This neatly captures meaning decompositionality as well as the modifiability and flexibility facts.

Finally, Sag addresses “semi-fixed expressions”, which allow only inflectional variants and external modification (e.g. *kick the bucket*). Like flexible idioms, these are comprised of two listemes, but unlike them, the direct object actually makes no semantic contribution. This is formalized by marking its IND feature as *none* and its FRAMES list as empty. The IND feature identifies the referent of the expression; the FRAMES feature specifies “the list of predications that together determine the meaning of a sign” (Sag 2012:89). This entails that the idiom cannot be internally modified, since there is no content there to modify. Lack of syntactic flexibility in these idioms is achieved via the type hierarchy: Unlike the verbal heads of flexible idioms (e.g. *pull*), which are of the type *strict-transitive-verb-lexeme* and inherit from type *transitive-verb-lexeme*, the verbal heads of semi-fixed expressions (e.g. *kick*) are stipulated to be of type *pseudo-transitive-verb-lexeme*, which is not a subtype of *transitive-verb-lexeme* and therefore does not

---

<sup>85</sup> In point of fact, this should prevent idiomatic *pull* from occurring with other direct objects, but it cannot prevent idiomatic *strings* from occurring as the object of another verb; some other mechanism will be required for that task.

inherit passivization. As a result, idioms with verbs of type *pseudo-transitive-verb-lexeme* (i.e. semi-fixed expressions) are not passivizable.

As it stands, Sag's theory predicts two types of idioms: completely flexible and completely inflexible.<sup>86</sup> This is inconsistent with the data from the previous chapter. However, Sag's SBCG has the kind of theoretical machinery that can capture subset behavior, with its sophisticated system of nested types. In principle, then, it would be possible to modify Sag's proposal with a more complex set of types to accommodate the subset relationships among the passive, incorporation, and gerundization variations. However, again, the resulting model would be purely descriptive; it would not follow from any independent properties of the constructional types involved. Hence it would be largely stipulative, lacking any external correlates that would suggest a deeper underlying motivation for the type of structures proposed. Within SBCG, there is no reason why all idioms that occur in gerundization should also occur in passivization, and it would be just as easy to construct the type system in such a way that the reverse entailment held. It would be better, then, to find a theory that can provide not just a description of this phenomenon but an explanation.

We have seen that Sag's SBCG approach to idioms predicts a simple flexible/inflexible dichotomy among idioms, which is not borne out in the data. Even if this approach were modified using the existing theoretical machinery to account for the observed subset behavior, SBCG would still fall short of explanatory power.

---

<sup>86</sup> There are technically three types of idioms, of course, but none of his "fixed expressions" are verbal idioms, and all verbal idioms appear to at least allow inflection, suggesting that they are either "semi-fixed expressions" or flexible idioms.

#### 4.2.7 *Summary*

In this section, we discussed many theories of idiomaticity that make predictions which are not in line with the empirical observations of this dissertation. In particular, the corpus experiments in Chapter 3 have shown that any viable theory of idiom flexibility must be able to capture the fact that gerundizable idioms are a subset of the idioms that allow incorporation, which are in turn a subset of passivizable idioms. We have seen that it is the systematicity of this strict subset relationship which many theories have difficulty capturing. For instance, most lexical specification theories predict no patterns of idiom flexibility. Fraser (1970)'s Frozenness Hierarchy (and a modified version of Egan (2008)'s pretense account) has no difficulty with the subset behavior, but in fact overgenerates that behavior: it predicts that all idioms should be ordered on a scale of least to most frozen, and that the variations are also ordered so that some only apply to the least frozen idioms, while others apply to more frozen idioms as well. While this could capture the subset relationships among the argument-structure-type variations, it fails to account for the fact that the other variations under investigation were not similarly involved in subset relationships. As a sort of middle ground between those two extremes, idiom decomposition, partial functions, and quasi-inference predict two classes of idioms—flexible and inflexible—but have no way to account for patterns within the set of flexible idioms. Finally, Synchronous Tree Adjoining Grammar and Sign-Based Construction Grammar have the theoretical machinery to account for the attested subset relationships, but to do so would be a matter of sheer stipulation within either theory. In the following section, we look at the model sketched in previous chapters, which is not

only equipped with the machinery to account for such subset behavior, but in fact can provide an explanation for why that behavior exists for the three variations under investigation.

### **4.3 The Structural Requirement Model**

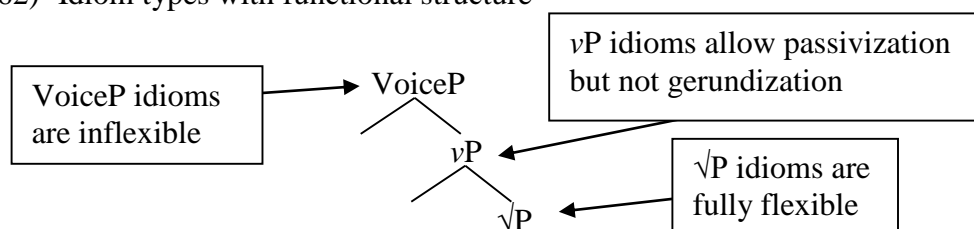
We have now discussed several theories that make predictions that are incompatible with the subset behavior among the argument-structure-type variations. I have shown that the gold standard for a theory of idiomatic flexibility is one that is able to both capture the descriptive generalizations from Chapter 3 and provide an explanation for those generalizations. More specifically, this theory should be able to adequately represent the fact that idioms that occur in gerundization also occur in incorporation, and those that occur in incorporation also occur in passivization. Furthermore, rather than simply doing so by stipulation, this theory should provide some independent motivation for the fact that it is passivization that is a superset of gerundization and not the other way around. In this section, I describe in detail the structural requirement model, which meets this standard.

#### ***4.3.1 Details of the model***

According to the structural requirement model of idiom flexibility (Folli & Harley 2007; Stone 2009; Harley & Stone 2013; Punske & Stone 2014; Stone to appear), there is a link between patterns of flexibility and independently motivated functional structure within the verbal hierarchy. More specifically, the amount of functional structure required to

license an idiom's special interpretation determines which variations that idiom will allow, since different amounts of functional structure can be independently shown to be involved in different variations.<sup>87</sup> Punske & Stone (2014) propose a typology of idioms based on the availability of VoiceP, *v*P, and  $\sqrt{P}$  as idiomatic domains: VoiceP idioms require both VoiceP and *v*P to license the idiomatic interpretation and therefore allow neither passivization (which is associated with VoiceP) nor gerundization (which is associated with *v*P); *v*P idioms require only the *v*P layer of functional structure and are therefore passivizable but not gerundizable; and  $\sqrt{P}$  idioms are true root + object idioms, placing no requirements on their functional structure, and are therefore both passivizable and gerundizable.

(62) Idiom types with functional structure



It should be easy to see how this approach can be extended to capture the empirical observations of this dissertation. Furthermore, because it is rooted in independently

<sup>87</sup> The independently syntactically motivated structure could itself have its roots in something else, such as the semantic requirements of the conceptual formatives that the structure represents (elements like CAUSE, BECOME, STATE, etc.). However, the theory of these conceptual formatives is not sufficiently well-developed to motivate the present analysis. While we continue to look for deeper explanations, the syntactic account adopted here represents the current state of the theory. (Thanks to Mike Hammond, p.c. for inquiring after deeper explanations.)

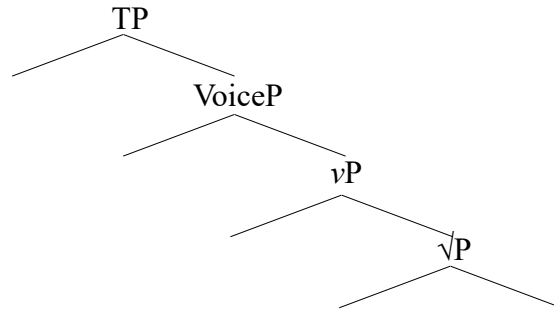
established functional structure, the structural requirement approach has more explanatory power than any of the accounts discussed above.

This approach is grounded in a long tradition of work on the extended functional projections within the verbal domain (Hale & Keyser 1993; Chomsky 1995; Harley 1995; Kratzer 1996; Marantz 1997; Borer 1998; Marantz 2001; Pylkkänen 2002; Cuervo 2003; Alexiadou, Anagnostopoulou & Everaert 2004; Harley 2007; Pylkkänen 2008; Ramchand 2008; Harley 2013; Merchant 2013; Harley 2014a; Alexiadou, Anagnostopoulou & Schäfer 2015, among many others). The result of this work is a well-established hierarchy of core verbal functional structure. These projections are relatively stable across many nuanced models of syntax (though the particular names may be different), and their relative ordering has been fixed by years of research and argumentation. This core skeleton is provided in (63) below.<sup>88</sup>

---

<sup>88</sup> While other functional structure likely intervenes between these projections in a more fully articulated verb phrase (see, e.g., Cinque (1999) for a more fine-grained approach), only VoiceP and vP are required to capture the subset behavior under discussion here.

## (63) Verbal functional structure



The structural requirement model also assumes that particular variations are associated with specific projections within the verbal functional domain. Again, these associations are well-established in the literature. VoiceP is the locus of passivization (Pylkkänen 2002; Folli & Harley 2005; Pylkkänen 2008). vP is the locus of nominal gerundization (Harley 2008; Punske 2012).<sup>89</sup> This functional structure and the associations that hold between variations and verbal projections provide the basis for this model's predictive power.

Having established the relevance of key verbal projections, we can now explore how they align to give the range of attested patterns in idiomatic variations and prevent the unattested patterns.<sup>90</sup> In particular, I show that this functional hierarchy predicts the

---

<sup>89</sup> This claim is not without its detractors. Alexiadou & Iordăchioaia (2015), for example, claim that nominal gerundization targets a much larger structure (AspP above VoiceP) and explicitly argue that verbal idioms should not allow nominal gerundization. In fact, the data discussed in this dissertation can help distinguish between these opposing theoretical positions. We have seen that (some) verbal idioms do indeed allow gerundization, suggesting that the low analysis of nominal gerundization may be more plausible.

<sup>90</sup> Because it is not considered in the relevant literature, discussion of incorporation is reserved until later in this section.

patterns of variation in Table 3.5, where “yes” indicates that a variation is allowed by an idiom and “no” indicates that it is not allowed. According to the predictions of the functional hierarchy, we expect to find idioms that occur in both passivization and gerundization, idioms that occur in passivization but not gerundization, and idioms that occur in neither passivization nor gerundization. We do not expect, however, to find idioms that occur in gerundization but not passivization.

<b>Passivization</b>	<b>Nominal Gerundization</b>	<b>Possible?</b>
yes	yes	✓
yes	no	✓
no	no	✓
no	yes	✗

**Table 4.2: Patterns of idiomaticity predicted by the structural requirement model**

These predicted patterns are derived from the tree in (63). Because VoiceP is higher than  $\nu$ P, it is possible for an idiom to allow variation at the VoiceP level but not the  $\nu$ P level. However, the opposite is not true—since  $\nu$ P is lower in the tree than VoiceP, it is impossible for an idiom to allow variation at the  $\nu$ P level but not the VoiceP level. In terms of flexibility, this means that an idiom can be freely passivizable without being gerundizable, since passivization is associated with the higher projection. However, an idiom cannot be gerundizable without being passivizable, since gerundization is associated with the lower projection.

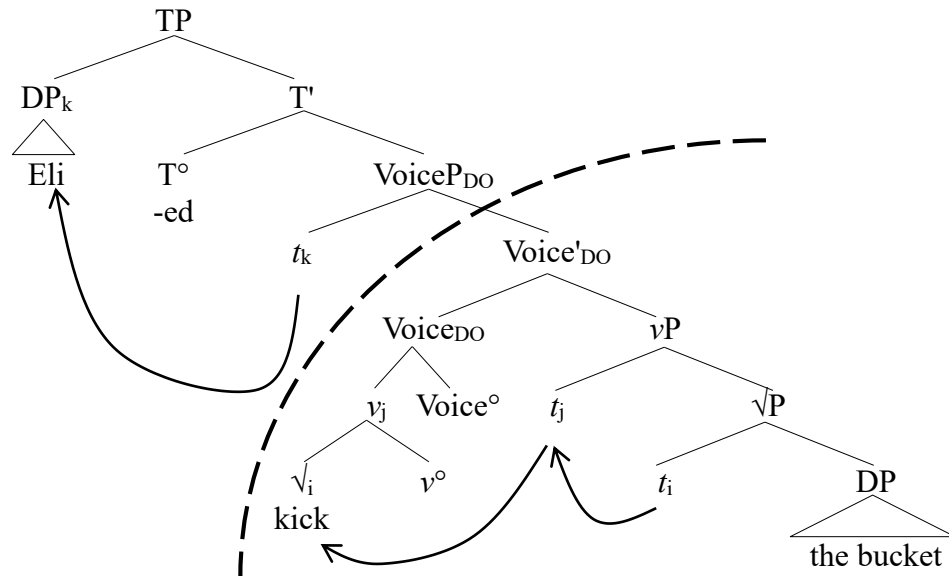


The hierarchical nature of the verbal functional structure, combined with the fact that these variations are associated with nested functional projections, gives rise to the precise subset pattern of behavior that was attested in the previous chapter. If an idiom requires only  $\sqrt{P}$  to receive its figurative interpretation, then it will be completely flexible. If it is fixed at the  $\nu P$  level, it will require this verbalizing head and so will not allow any category-changing variations. Finally, if an idiom requires VoiceP, it will specify the particular Voice head with which it must occur—e.g. Voice<sub>DO</sub> for agentive idioms or Voice<sub>PASS</sub> for passive ones. As such, the idiom's Voice—active or passive—will be fixed, as will all of the structure below that level.

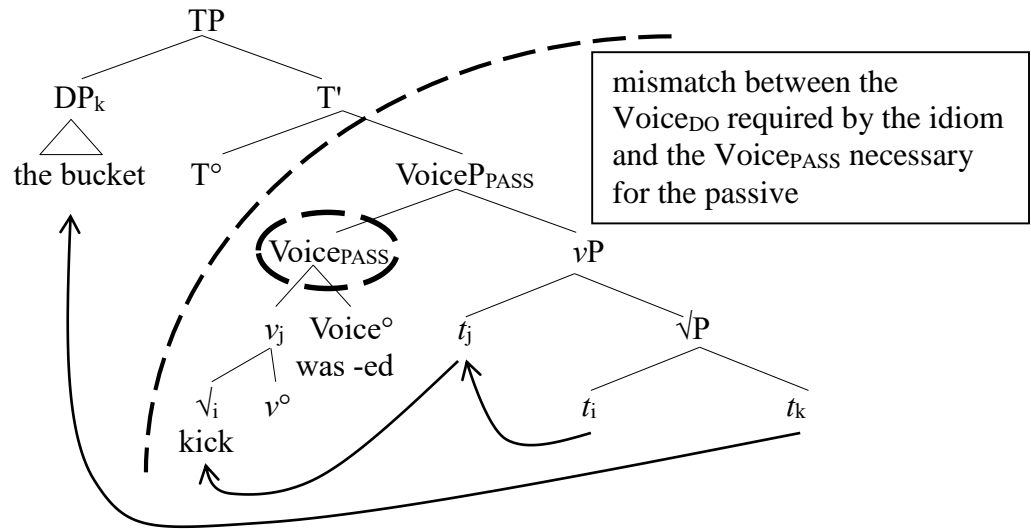
This last part is the key to the nested subset behavior and bears repeating: Idioms that are fixed at one level of the verbal hierarchy are fixed at all lower levels. This means that idioms which disallow passivization—a VoiceP variation—will also not allow gerundization—a  $\nu P$  variation. In contrast, idioms fixed at one level will be free at all higher levels. Thus idioms which disallow gerundization will allow passivization if they are fixed at the  $\nu P$  level:  $\nu P$  level variations will be prohibited, but variations higher in the structure will be acceptable.

*Kick the bucket* is an example of a VoiceP idiom. It requires the active, agentive Voice head, Voice<sub>DO</sub>. Thus, the idiomatic meaning is not accessible in the passive voice, where Voice<sub>DO</sub> is replaced by the passive Voice head, Voice<sub>PASS</sub>.

(64) VoiceP idiom – *Eli kicked the bucket*



In (31), the dashed line shows the idiomatic domain; all of the functional structure (and lexical content) below this line must be present in order to license the idiomatic interpretation. When the structure is passivized, VoiceP<sub>DO</sub> is replaced with VoiceP<sub>PASS</sub>. Consequently, the passive structure does not contain the required functional material, and the structure cannot be interpreted idiomatically, as in (32).

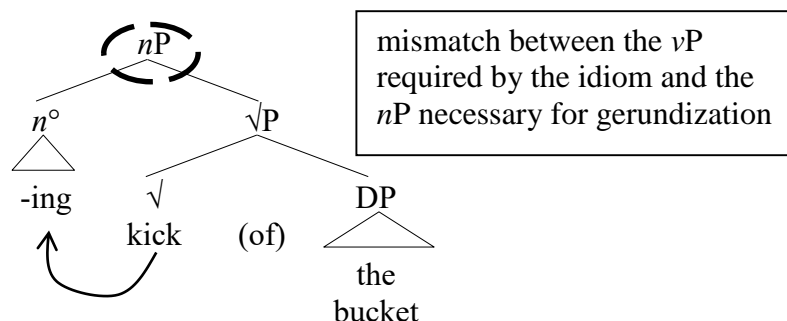
(65) VoiceP idiom – #*The bucket was kicked*

There is a further consequence of idiomatic domains, mentioned above: If an idiom is fixed at a given level, all functional material below that level is also fixed. In the case of *kick the bucket*, because this idiom is fixed at the VoiceP level, νP must also be fixed. This has the effect of preventing idiomatic *kick the bucket* from appearing in category-changing variations that target νP below VoiceP, such as nominal gerundization<sup>91</sup>, shown in (33) below.

<sup>91</sup> Note that verbal gerunds seem to be fine with all idioms, including *kick the bucket*:

- i. John's kicking the bucket left his mother completely distraught.

Verbal gerunds target a larger piece of structure (Alexiadou 2013) and therefore represent an *nP* merging above VoiceP. This is consistent with the fact that all idioms seem to allow verbal gerunds; the locus of verbal gerundization is higher than the domain for idiomatization of even the most "fixed" of idioms.

(66) VoiceP idiom – #*kicking of the bucket*

When vP is replaced by a non-verbal categorizer—e.g. nP, as in (33)—the idiomatic interpretation of *kick the bucket* is no longer available. This can be explained if that functional structure, though unpronounced, is required to license the idiomatic meaning.

While *kick the bucket* is considered a VoiceP idiom because it is fixed very high in the structure, *steal the show* is a vP idiom, fixing the functional structure within the verbal hierarchy just up to the vP head.

## (67) vP idiom – steal the show

- (a) The show was stolen (by the flying dolphin).
- (b) #The stealing of the show (by the flying dolphin)...

Idioms of this type allow passivization because they are free at the VoiceP level; this means that their idiomatic interpretations are not dependent on the presence of a specific Voice head, and either Voice<sub>DO</sub> or Voice<sub>PASS</sub> is fine—both are outside the domain of idiomatic interpretation.

Finally, there are √P idioms such as *turn the tables*. These are true root-plus-object idioms, requiring no additional functional structure to license their idiomatic

interpretations. As such, they occur in both the passivization and gerundization variations. Being free at both the VoiceP and *v*P levels, these idioms can appear with a variety of different Voice heads, as well as different category-changing functional heads that occur low in the tree (e.g. *n*P, which forms nominal gerunds).

The typology developed so far is consistent with the predicted patterns from Table 3.5: A given idiom might allow both passivization and gerundization; it might allow neither; or it might allow passivization but not gerundization. However, an idiom should never allow gerundization but not passivization.<sup>92</sup> This is a direct reflex of the hierarchical nature of the verb phrase. VoiceP is the locus of passivization, and it is higher in the tree than *v*P, the locus of gerundization; it follows that passivization should occur more freely than gerundization and other variations involving structures lower in the tree, closer to the root. Thus far, the typology of idioms is as follows (Punske & Stone 2014):

(68) Preliminary typology of idioms

- (a) **VoiceP idioms** (e.g. *kick the bucket*): Fixed at the VoiceP level, these idioms require both VoiceP and *v*P to license their idiomatic interpretations. Therefore, they allow neither passivization nor gerundization.

---

<sup>92</sup> Of course, for a given idiom, gerundization might be much more frequent than passivization. In a corpus, this could reveal itself as the reverse pattern. This is one of many reasons why it is important to pair corpus data with grammaticality judgments. See Section 1.1 for a more detailed discussion.

- (b) **vP idioms** (e.g. *steal the show*): Fixed at the vP level, these idioms require only vP and are therefore passivizable but not gerundizable.
- (c) **√P idioms** (e.g. *spill the beans*): True root-plus-object idioms, these idioms have no requirements on functional structure. They therefore allow both passivization and gerundization.

Having established these facts, we now turn to object incorporation. What can this variation, which has received almost no attention in the literature, tell us about the nature of the model?

#### 4.3.2 *Re-examining the verbal functional structure*

This typology works nicely for passivization and gerundization, which have already been considered in the literature. I now turn to object incorporation, the third argument-structure-type variation that was found in the previous chapter to participate in this subset behavior. Recall that an idiom that occurred in gerundization also occurred in incorporation, and one that occurred in incorporation also occurred in passivization. This positions incorporation between passivization and gerundization, suggesting that there is a fourth type in the typology of idioms and that another functional projection, call it FP, is located between VoiceP and vP and is the locus of incorporation.

#### (69) Revised typology of idioms

- (a) **VoiceP idioms** (e.g. *kick the bucket*): Fixed at the VoiceP level, these idioms require VoiceP, FP, and vP to license their idiomatic

interpretations. Therefore, they disallow passivization, incorporation, and gerundization.

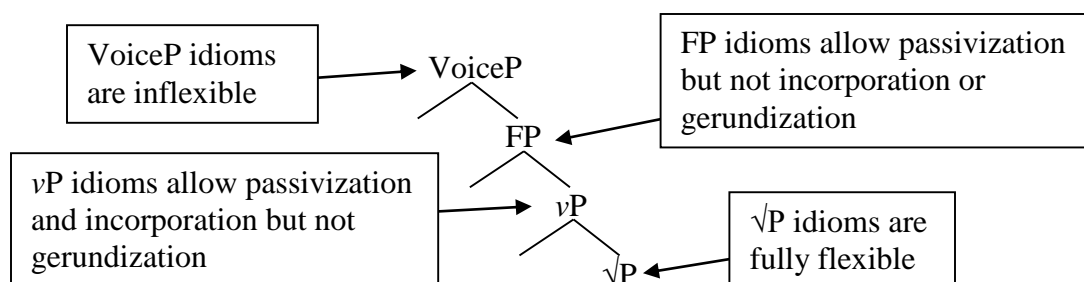
- (b) **FP idioms** (e.g. *kill the audience*): Fixed at the FP level, these idioms require both FP and vP. Thus, they allow passivization but not incorporation or gerundization.
- (c) **vP idioms** (e.g. *steal the show*): Fixed at the vP level, these idioms require only vP and therefore allow passivization and incorporation but not gerundization.
- (d) **√P idioms** (e.g. *spill the beans*): True root-plus-object idioms, these idioms have no requirements on functional structure. They therefore allow passivization, incorporation, and gerundization.

Although the identity of FP is unknown, its existence is predicted by the corpus results. This highlights an interesting—and, I think, important—interplay between theoretical and experimental research. The structural requirement model gains traction from the fact that the empirically attested idiom types correspond to functional structure within the verbal hierarchy that has been established independently of work on idioms. That is, VoiceP and vP are independently motivated, and are known to affect passivization and nominal gerundization, respectively. This is a key component of the current approach: Established theoretical machinery provides the scaffolding to support empirical observation.

However, the role of object incorporation shows that information can—and indeed should—flow both ways. The systematic behavior of the object incorporation

variation suggests that it is tied to syntactic structure, along with passivization and gerundization. If those two variations are linked to functional heads, it stands to reason that this third variation should be linked to a functional head of its own. Furthermore, the fact that incorporation is a subset of passivization and a superset of gerundization suggests that this functional head should be located above  $\nu$ P but below VoiceP. This is shown schematically in (70) below.

(70) Idiom types with functional structure



The findings of this study, then, present an empirical challenge for theoreticians working in this field: What is the identity of FP?

Although it is too early to identify FP with certainty, AgrOP is a good candidate (Koizumi 1993; Chomsky 1995; Bobaljik & Carnie 1996).<sup>93</sup> As the locus of object incorporation, this projection serves as the landing site for the incorporated object. Because AgrOP was proposed to check the case feature of the object, it is reasonable to assume that it could serve this function. Furthermore, although the position of AgrOP in the extended verbal domain is not as certain as the other projections under discussion

<sup>93</sup> Thanks to Andrew Carnie (p.c.) for suggesting this.



here, it is again reasonable to think that it might be located between *vP* and *VoiceP*.

Further research is necessary to corroborate this preliminary assessment.

This empirical prediction is confounded slightly by two factors. First of all, what I have labeled “object incorporation” throughout this dissertation is actually (at least) two distinct operations working in tandem: the object has combined with the verb, and the resulting structure has been recategorized as an adjective. English is notoriously resistant to verbal compounds (Harley 2011), and that is apparent in this domain. The corpus search for object incorporation focused on forms ending in *-ing*, which limited the results to participles and gerunds (i.e. incorporated structures that are, ultimately, nonverbal). Because these forms are at least twice removed from simple active verb phrases, it is therefore possible that more than a single functional head is involved.

Furthermore, and perhaps more importantly, the data from the previous chapter are somewhat inconclusive with respect to object incorporation. While passivization and gerundization are in a clear superset-subset relationship, as are incorporation and gerundization, all but one of the exceptions to the subset relationships across both corpora involve hierarchical reversals of passivization and incorporation. The complete list of exceptions is provided in (71) below:

## (71) Exceptions to the subset relationships

- (a) *bust a move* occurred in incorporation but not passive (Google Books)
- (b) *hit the spot* occurred in incorporation but not passive (Google Books)
- (c) *jump the gun* occurred in incorporation but not passive (COCA)
- (d) *rock the boat* occurred in incorporation but not passive (COCA)
- (e) *take a chance* occurred in incorporation but not passive (COCA)
- (f) *spill the beans* occurred in gerundization but not passive or incorporation (COCA)

These exceptions indicate that the subset relationship between object incorporation and passivization is less robust than the others. There are a number of possible explanations.

As suggested in the previous chapter (Section 3.3.2 and Section 3.4.2), it is possible that these are not exceptions at all, but rather simple anomalies in the data, which could be attributed to various factors, such as the relative frequencies of the individual forms in question. For instance, *gun-jumping* has an established meaning in the stock market world: it is the (illegal) act of trading securities on the basis of information that has not yet been disclosed to the public. Given this, it seems unsurprising that the incorporated structure would be more frequent than its passive counterpart, which could be possible but accidentally unattested in the corpus. With *hit the spot*, there was a type of interference in the search itself (in Google Books) from the distinct idiom *sweet spot*. Because I only looked at 100 total hits for each search string, it is possible that a passive version of this idiom was present in the corpus but was missed due to the large number of instances of *the sweet spot was hit*. Similar explanations were

proffered for the remaining exceptions in the previous chapter, and it seems plausible that more data are needed to see the full effects of this subset relationship.

Another possibility is that FP is not positioned below VoiceP but rather above.<sup>94</sup> The number of examples differentiating VoiceP from FP idioms is actually quite small; in Google Books only four idioms exhibited the FP pattern of behavior (occurrence in passivization only) and in COCA only five did. In contrast, there were two idioms in Google Books and four in COCA that occurred only in incorporation (both listed above as exceptions). The preference for passivization in these cases is thus only slight, and it is not immediately clear if this is accidental or the result of an underlying systematicity. A larger body of data would help to truly distinguish the two idiom types and solidify the location of FP in the verbal hierarchy.<sup>95</sup>

A third alternative is that object incorporation is simply not part of the hierarchy, in spite of initial appearances. If this were the case, the exceptions listed in (71) would not be exceptions, *per se*, but would be considered evidence of the fact that object incorporation does not behave in a principled way with respect to the other argument-structure-type variations; more precisely, that incorporation does not have a subset (or superset) relationship with passivization and/or gerundization. This kind of behavior would in fact be expected if incorporation were not actually participating in subset

---

<sup>94</sup> This, however, could be problematic for phase-based accounts of idiomatic domains; cf. Marantz (1984); Svenonius (2005); Stone (2009); Harley & Stone (2013); Harwood (2013).

<sup>95</sup> The higher position of FP is consistent with an analysis of object incorporation structures as deverbal adjectives, but only if they are structurally quite large—including at least VoiceP inside them. This is not a typical analysis of these forms. For instance, Punske (in prep) suggests that they include only the *vP* level.

relationships with passivization and gerundization. As with the previous suggestion, additional data could help tease apart whether the examples in (71) are exceptions to some kind of systematic behavior or rather evidence that no such systematic behavior exists.<sup>96</sup>

Of these three alternatives, not all are equally plausible. The possibility that the behavior of object incorporation only appears to be systematic at first glance but is in fact random seems unlikely. In that scenario, we would expect more balanced results, i.e. more exceptions; instead, the overwhelming majority of idioms that occurred in object incorporation followed the expected pattern. In Google Books, there was an 89% probability that an idiom that occurred in incorporation also occurred in passivization (two exceptions out of 19 idioms that occurred in incorporation); in COCA, the probability was 71% (three exceptions out of 14 idioms). This is at least suggestive of a subset pattern between these variations.

The second alternative is more plausible. Because the number of idioms appearing only in the incorporation variation was roughly equal to the number of idioms appearing only in the passive variation, it is difficult to establish the relative positions of FP and VoiceP in the verbal hierarchy. Theoretical research suggests that VoiceP is a phase boundary and thus a likely cut-off point for idiomatic interpretations, but only additional data can truly help distinguish these two options.

---

<sup>96</sup> Object incorporation is quite systematic in English in general. See Harley (2011) for discussion.

Given the nature of the exceptions, the first alternative seems most likely. As discussed above, the relative frequency of individual forms (e.g. *gun-jumping*) can have an effect on corpus results. In any case, more data will help distinguish among the three alternatives. Whether these exceptions are actual anomalies that will resolve themselves with a larger corpus, a sign that FP is actually higher—not lower—than VoiceP, or evidence that object incorporation is in fact not part of the hierarchy at all, only a larger amount of data can tell with any measure of certainty.

Given the current limitations and scope of the project, it seems reasonable to tentatively conclude that FP is a functional head located above *v*P and below VoiceP, and that this functional head is the “control center” for object incorporation. This conclusion is based on the relative strength of the subset relationships involving incorporation and the other argument-structure-type variations, as well as the nature of the exceptions discussed above. Chapter 5 provides some speculation about other possible data that could be brought to bear on this discussion.

Regardless of the status of FP, the relationship between passivization and gerundization is clear and unambiguous: Within the domain of idioms, these two variations stand in a superset-subset relationship, with passivization being a superset of gerundization. This has been observed independently (Punske 2010; Punske 2012; Bruening 2013) and correlates with existing functional structure in the verbal hierarchy.

### 4.3.3 *Summary*

The structural requirement model of idiom flexibility is best suited to address the empirical observations of this dissertation. Not only does it have the theoretical machinery to account for the observed subset behavior—namely that gerundizable idioms are a subset of the idioms that allow incorporation, which are in turn a subset of passivizable idioms—but that theoretical machinery is motivated on independent grounds that are unrelated to idioms. The four idiom types are associated with functional projections in the verbal hierarchy, and these functional projections are independently associated with the relevant variations in the appropriate hierarchical arrangement. Systematic patterns of syntactic flexibility are easily accommodated within such an account, which leads to a theory of idioms built on independent components of the grammar.

## 4.4 **Conclusion**

Idioms have been a mystery in the generative tradition of linguistics nearly since its inception. Their syntactic and semantic behaviors are different from more obviously compositional language, and generativists have long sought to explain these facts (Katz & Postal 1963; Chafe 1968; Weinreich 1969; Fraser 1970; Fraser & Ross 1970; Newmeyer 1974; Ruhl 1975, etc.). The previous chapter brings additional data to bear on the most suitable theory of idiomaticity: Because three syntactic variations were found to be in a subset relationship, any theory of idioms must be able to capture this type of

behavior. Of the many existing theories, only the structural requirement approach to idiom flexibility can provide not just a description but an explanation for these new data.

Of the theories that make explicit claims about idiom flexibility, most are unable to adequately account for this subset relationship. Lexical specification approaches like those of Weinreich (1969) and Katz (1973) predict that there should be no pattern to flexibility in idioms, so the subset relationships in these data are unexpected. Approaches like Fraser (1970)'s Frozenness Hierarchy and Egan (2008)'s pretense (with some modifications) can provide the scaffolding for subset relationships like the one revealed here, but they take it too far, predicting that all variations are in a subset relationship, not just those involving changes in argument structure. Many other approaches, such as Nunberg, Sag & Wasow (1994)'s idiom decomposition, Gazdar et al. (1985)'s partial functions, and Pulman (1993)'s quasi-inference, predict two classes of idioms—flexible and inflexible—and are therefore unable to account for differences within the class of flexible idioms. Finally, Abeillé (1995)'s Synchronous Tree Adjoining Grammar and Sag (2012)'s Sign-Based Construction Grammar are both theories with enough power to account for the relevant distinctions, but they fail to reach the level of explanatory adequacy. In either case, the specific subset relationships would be a matter of stipulation.

Only the structural requirement model actually predicts the very patterns in idiom flexibility that were found in Chapter 3. Because idiomatic flexibility is tied to differences in the amount of functional structure required by a given idiom, and because the functional structure of the verb phrase is inherently hierarchical, subset relationships

can not only be explained but are actually predicted. In this theory, it is not by coincidence or stipulation that gerundization is a subset of passivization; it is a direct result of the fact that  $\nu$ P is lower in the verbal hierarchy than VoiceP.

Given all of these considerations, the structural requirement model emerges as the most suitable theory of idiomatic flexibility. This approach alone can fully explain the subset relationships from the previous chapter, namely that idioms that occurred in gerundization also occurred in incorporation, and those that occurred in incorporation also occurred in passivization.



## CHAPTER 5: CONCLUSION

This dissertation has explored the issue of idiom flexibility from three perspectives: experimental, corpus-based, and theoretical. The primary empirical finding is that idiom behavior is patterned and systematic—certain variations display subset behaviors that are not easily explained by the most prominent models of idiomaticity. More specifically, if an idiom occurred in nominal gerundization, it also occurred in object incorporation; and if it occurred in incorporation, it also occurred in passivization. I have argued that these subset relationships are not just compatible with but are actually predicted by the structural requirement model of Folli & Harley (2007), Harley & Stone (2013), and Punske & Stone (2014). According to this model, different idiom types are associated with differences in the functional structure required to license an idiomatic expression's figurative interpretation. The subset relationships among idiom types mirror the hierarchical structure of the verbal functional domain. In what follows, I summarize the key contribution of each chapter. I then turn to some promising avenues for future work in the area of idiom flexibility.

The studies in Chapter 2 tested one prediction of the structural requirement model of idiom flexibility, namely that idioms should fall into two clear groups with respect to passivizability: passivizable and nonpassivizable. It ultimately became clear that certain methodologies were more effective for testing idiom behavior than others. The experiments in Chapter 2 highlighted some of the many problematic aspects of the experimental methodologies. Experiment 1 was a straightforward acceptability task

where subjects were given context paragraphs and then asked to rate target sentences containing idioms in one of three variations (active, passive, or nominal gerund) on a 1-7 scale. Subjects reported that they ignored the syntactic variation and rated idioms instead on how well the figurative meaning aligned with the context paragraph. For Experiment 2, subjects' reading times were recorded for congruent and incongruent sentences in both the active and passive variations. We expected incongruent sentences to be read more slowly than congruent ones, because there is a processing delay when we encounter a semantic mismatch. Given this semantic incongruity effect, we expected nonpassivizable idioms to be incongruent in the passive condition, while passivizable idioms would be congruent. The results showed no such effect, and I speculated that the reading times were measuring a stage in the comprehension process prior to idiom interpretation (Townsend & Bever 2001).

Chapter 3 approached idioms from a different angle: corpora. Whereas experiments measure individuals' responses to carefully crafted stimuli, corpus studies provide access to actual "real-world" usage. The question here is not about what a speaker *might* do, but about what speakers *actually* do. The goal of these studies was to test another prediction of the structural requirement model, namely that certain variations should participate in subset relationships. Using Google Books for the first study and COCA for the second, I recorded the occurrences of idioms in five variations. The results of the Google Books study and, to a lesser extent, the COCA study confirmed the predictions of the structural requirement approach. I found that the set of gerundizable

idioms is a subset of the idioms that allow incorporation; and the set of idioms that allow incorporation is a subset of passivizable idioms.

Chapter 4 elaborated the structural requirement model, arguing that the subset relationship between passivization and gerundization from Chapter 3 is precisely the one predicted by the model. Furthermore, I showed that several other theories of idiom flexibility are unable to account for these patterns. Many theories are simply incapable of capturing subset relationships. For others that are capable of modelling the data, they must do so by stipulation. The structural requirement approach alone can provide a motivation for these subset relationships. The chapter also commented on the interplay between theory and data and speculated about the inclusion of AgrOP in the verbal functional domain to account for the subset relationships involving object incorporation.

The findings reported here open the door for substantial future work. One logical next step is to seek independent confirmation for the idiom types proposed above. In English, Voice and *v* morphology is typically null. Thus, there is no overt marking on idioms that require one of these functional heads.<sup>97</sup> An obvious place to begin looking for such confirmation is in languages with overt marking of these categories, such as Japanese, Persian, or Greek. The structural requirement model predicts that we should find idioms that can appear with both different *v* and different Voice heads; idioms with a fixed *v* head that can be embedded under different Voice heads; and idioms with fixed *v* and Voice heads. However, we do *not* expect to find idioms with a fixed Voice head that

---

<sup>97</sup> In fact, part of the difficulty in determining what differentiates canonically passivizable *spill the beans* and canonically nonpassivizable *kick the bucket* lies in the fact that their surface syntax is identical.

can have different  $v$  heads. This kind of evidence would be a significant next step toward building this model.

Determining the identity of FP and its position in the verbal hierarchy is another priority for future work. In Section 4.3.2, I speculated that FP might be AgrOP, noting that the case-checking properties of AgrOP align with the fact that FP must be the locus for object incorporation. However, AgrOP has been little-used in recent syntactic analyses, so its position with respect to the more highly articulated verb phrase is a bit unclear; it is definitely below TP and above  $\sqrt{P}$ , but different studies place it in different locations unordered with respect to  $vP$  and/or VoiceP (e.g. Koizumi 1993; Bobaljik & Carnie 1996). Another possibility for FP is AspectP (Travis 2000), which has also been proposed to check accusative case.

It will be impossible to determine the identity of FP with any degree of confidence, however, until the corpus results have been confirmed. The subset relationships were most robust in Google Books, but there were some methodological concerns, such as the ambiguity of the “hits” returned by the search algorithm. There may be new corpus resources available to confirm these findings and, more importantly, establish with certainty whether object incorporation is a subset or a superset of passivization. This will be crucial in determining whether FP is situated above or below VoiceP in the verbal hierarchy.

Finally, idiom flexibility extends far beyond the six variations that I examined in these studies. In Section 3.2, I discussed the methodology used by Fellbaum (2007) for her large-scale study of idiom flexibility in German. Rather than choosing a select

number of variations and searching for idioms in those variations, she and her colleagues were able to look for a wide range of idiom flexibility. Conducting a similar study for English idioms would be ideal. One reason to conduct such a study would be to learn more about variations which were not included in the present study but which the structural requirement model predicts to be related. For instance, relativization, question formation, and quantification all involve *wh*-operators. As such, this model predicts that these three variations should be correlated—if a given idiom occurs in any one of these variations, it should also occur in the other three. Conducting another large-scale corpus investigation would be one way to test this prediction.

To summarize the key findings of this dissertation, the research reported here found that idioms behave in predictable, patterned ways with respect to three variations: gerundizable idioms are a subset of the idioms that occur in object incorporation, which are in turn a subset of passivizable idioms. This result is surprising given the most well-known theories of idiom flexibility, particularly that of Nunberg, Sag & Wasow (1994), which predicts that idioms should fall into two classes—flexible and inflexible. Only the structural requirement model is able to account for, and even predict, these subset relationships.

## APPENDIX A: STIMULI FOR EXPERIMENT ONE

Context Paragraph	Stimulus Sentences
Cindy hated tomatoes and complained every time her husband put them on his salad. Even though he knew how she felt about them, she refused to stop bringing it up.	<p>I heard that Cindy beat a dead horse.</p> <p>I heard that a dead horse was beaten by Cindy.</p> <p>I heard about Cindy's beating of a dead horse.</p>
Lisa was working on a project that required a \$500 piece of software. She bought the software in spite of the fact that it was so expensive.	<p>I heard that Lisa bit the bullet.</p> <p>I heard that the bullet was bitten by Lisa.</p> <p>I heard about Lisa's biting of the bullet.</p>
Stanley was a grumpy old man. After many years of tormenting his neighbors, he finally died.	<p>I heard that Stanley bit the dust.</p> <p>I heard that the dust was bitten by Stanley.</p> <p>I heard about Stanley's biting of the dust.</p>
Amy was driving home from work one evening when she saw some gang members vandalizing a building. She immediately reported their illegal activity by calling 911.	<p>I heard that Amy blew the whistle.</p> <p>I heard that the whistle was blown by Amy.</p> <p>I heard about Amy's blowing of the whistle.</p>
Roger was waiting for class to begin on the first day of the semester. He struck up a conversation with the girl sitting next to him.	<p>I heard that Roger broke the ice.</p> <p>I heard that the ice was broken by Roger.</p> <p>I heard about Roger's breaking of the ice.</p>
Grace got distracted while driving to school one day and sideswiped a telephone pole. When she got home, she told her parents that she wrecked their car.	<p>I heard that Grace broke the news.</p> <p>I heard that the news was broken by Grace.</p> <p>I heard about Grace's breaking of the news.</p>
Victor was double-majoring in engineering and English literature. Though the two fields are very different, he often wrote papers that incorporated aspects of both.	<p>I heard that Victor bridged the gap.</p> <p>I heard that the gap was bridged by Victor.</p> <p>I heard about Victor's bridging of the gap.</p>
Donna went to a party with several of her friends. She was one of the first people on the dance floor, and everyone was watching her.	<p>I heard that Donna busted a move.</p> <p>I heard that a move was busted by Donna.</p> <p>I heard about Donna's busting of a move.</p>
Jerry didn't like Alexander because he had acne. He started making fun of Alexander, and all his friends joined in.	<p>I heard that Jerry cast the first stone.</p> <p>I heard that the first stone was cast by Jerry.</p> <p>I heard about Jerry's casting of the first stone.</p>
Teresa forgot to put her hair up before her chemistry lab. It was hanging over a Bunsen burner and burst into flames.	<p>I heard that Teresa caught fire.</p> <p>I heard that fire was caught by Teresa.</p> <p>I heard about Teresa's catching of fire.</p>
Joan was having a conversation with her co-worker about their boss. When he came in the front door, she quickly switched the topic of conversation to avoid being overheard.	<p>I heard that Joan changed the channel.</p> <p>I heard that the channel was changed by Joan.</p> <p>I heard about Joan's changing of the channel.</p>

Context Paragraph	Stimulus Sentences
Larry and his girlfriend had a fight one night and he called her a bad name. The next day he apologized and they made up.	I heard that Larry cleared the air. I heard that the air was cleared by Larry. I heard about Larry's clearing of the air.
Deborah was walking through a busy parking lot. Someone started backing out of a nearby spot, and she had to hurry out of the way to avoid getting hit.	I heard that Deborah cleared the decks. I heard that the decks were cleared by Deborah. I heard about Deborah's clearing of the decks.
Richard and his wife had 7 children. As he was preparing to leave for the grocery store, he counted the children in the van to make sure they were all present.	I heard that Richard counted heads. I heard that heads were counted by Richard. I heard about Richard's counting of heads.
Elizabeth was taking a test in class one day. She knew the material very well, but she couldn't remember one of the key vocabulary words.	I heard that Elizabeth drew a blank. I heard that a blank was drawn by Elizabeth. I heard about Elizabeth's drawing of a blank.
Aaron was a manager who tried to be lenient with his employees. The one thing he would not tolerate, however, was people telling dirty jokes on the job.	I heard that Aaron drew the line. I heard that the line was drawn by Aaron. I heard about Aaron's drawing of the line.
Betty was at the flea market looking for some new furniture. She found a futon that she really wanted, and she negotiated the seller down to half the posted price.	I heard that Betty drove a hard bargain. I heard that a hard bargain was driven by Betty. I heard about Betty's driving of a hard bargain.
Jeremy met a girl at a party one night, and they decided the next day that they would get married. When he told his friends, they were very shocked.	I heard that Jeremy dropped a bomb. I heard that a bomb was dropped by Jeremy. I heard about Jeremy's dropping of a bomb.
Stephanie's boss asked her to prepare a report for an important client. She completely forgot about the report because she was having a fight with her husband.	I heard that Stephanie dropped the ball. I heard that the ball was dropped by Stephanie. I heard about Stephanie's dropping of the ball.
Norman worked at a local law firm. His boss always made him do the crappy jobs and talked bad about him, but he stayed around because he really needed the money.	I heard that Norman ate shit. I heard that shit was eaten by Norman. I heard about Norman's eating of shit.
Joshua was a competitive NCAA athlete. When he was exposed for taking steroids, he had to deal with all of the resulting negative media coverage.	I heard that Joshua faced the music. I heard that the music was faced by Joshua. I heard about Joshua's facing of the music.
Ellen was tall, blonde, and thin. She auditioned for a modeling job, and the company decided that she was exactly what they were looking for.	I heard that Ellen fit the bill. I heard that the bill was fit by Ellen. I heard about Ellen's fitting of the bill.
Edward was running for governor. His popularity in the polls increased when his opponent announced that she was anti-education.	I heard that Edward gained ground. I heard that ground was gained by Edward. I heard about Edward's gaining of ground.

Context Paragraph	Stimulus Sentences
Lillian had a big presentation in class on Monday. Even though she didn't have to, she worked until 5:00 a.m. making sure it was the very best it could possibly be.	I heard that Lillian went the distance. I heard that the distance was gone by Lillian. I heard about Lillian's going of the distance.
Ronald was going to the convenience store. His route was diverted when he made a left-hand turn to avoid an accident in the road ahead.	I heard that Ronald hung a left. I heard that a left was hung by Ronald. I heard about Ronald's hanging of a left.
James was driving to Kris's house for the first time. When he got to Elm Street, he made a right-hand turn, following the directions that Kris had provided.	I heard that James hung a right. I heard that a right was hung by James. I heard about James's hanging of a right.
Anna was driving to work one day when she realized that she had left an important document at home. She made a U-turn and headed back to her house to get the document.	I heard that Anna hung a U-ey. I heard that a U-ey was hung by Anna. I heard about Anna's hanging of a U-ey.
Christine was having a bad week after her boyfriend broke up with her and her boss told her she could not have her requested vacation days. As if things weren't bad enough, she found out that her sister was in the hospital after falling off her horse.	I heard that Christine hit rock bottom. I heard that rock bottom was hit by Christine. I heard about Christine's hitting of rock bottom.
Joe got very angry after his car was towed. He yelled at the tow truck driver until his face turned red.	I heard that Joe hit the ceiling. I heard that the ceiling was hit by Joe. I heard about Joe's hitting of the ceiling.
Dorothy was sitting at home one evening when she heard gunshots outside. She quickly dropped to the floor to avoid getting hit by a stray bullet.	I heard that Dorothy hit the deck. I heard that the deck was hit by Dorothy. I heard about Dorothy's hitting of the deck.
Peter went out drinking with his friends one evening. When he finally got home at 3:00 a.m., he went straight to bed.	I heard that Peter hit the hay. I heard that the hay was hit by Peter. I heard about Peter's hitting of the hay.
Rachel was very excited about going on vacation. She loaded all her bags into the car and kissed her parents goodbye, then drove away.	I heard that Rachel hit the road. I heard that the road was hit by Rachel. I heard about Rachel's hitting of the road.
Edith was waiting for her husband to get home from work. He had been late every day that week without telling her why, and when he got home she yelled at him for 10 minutes.	I heard that Edith hit the roof. I heard that the roof was hit by Edith. I heard about Edith's hitting of the roof.
Paula was exhausted after a long day of hiking with her friends. She went to bed almost as soon as she got home to avoid the headache she could feel coming.	I heard that Paula hit the sack. I heard that the sack was hit by Paula. I heard about Paula's hitting of the sack.
Marcus was conducting an investigation of an explosion at a power plant. He announced to the press that it was accidental, but found out later that there was foul play involved.	I heard that Marcus jumped the gun. I heard that the gun was jumped by Marcus. I heard about Marcus's jumping of the gun.



Context Paragraph	Stimulus Sentences
Sarah was terminally ill. After many years alone, she passed away peacefully in her sleep one night.	<p>I heard that Sarah kicked the bucket.</p> <p>I heard that the bucket was kicked by Sarah.</p> <p>I heard about Sarah's kicking of the bucket.</p>
Rodney had been a smoker since he was in high school. When he turned 50, he gave up smoking, much to the delight of his children.	<p>I heard that Rodney kicked the habit.</p> <p>I heard that the habit was kicked by Rodney.</p> <p>I heard about Rodney's kicking of the habit.</p>
Janice worked for the same company for 13 years. Whenever someone new was hired, she was responsible for their orientation because she was so familiar with the company's policies and procedures.	<p>I heard that Janice knew the ropes.</p> <p>I heard that the ropes were known by Janice.</p> <p>I heard about Janice's knowing of the ropes.</p>
Antonio went outside one day and saw that his neighbor had a moving van in front of his house. He spent most of the day helping his neighbor load the truck.	<p>I heard that Antonio lent a hand.</p> <p>I heard that a hand was lent by Antonio.</p> <p>I heard about Antonio's lending of a hand.</p>
Jacqueline was ahead in the contest to collect canned food for the homeless. Then one day Francis brought in several boxes full of cans, which put him in the lead.	<p>I heard that Jacqueline lost ground.</p> <p>I heard that ground was lost by Jacqueline.</p> <p>I heard about Jacqueline's losing of ground.</p>
David was attending a new student orientation, and he was very bored. When they brought out the snacks, he went directly to the table and grabbed a cookie before anyone else even noticed.	<p>I heard that David made a beeline.</p> <p>I heard that a beeline was made by David.</p> <p>I heard about David's making of a beeline.</p>
Lucille was an astute trader. She made a lot of money on the stock market last weekend.	<p>I heard that Lucille made a killing.</p> <p>I heard that a killing was made by Lucille.</p> <p>I heard about Lucille's making of a killing.</p>
Jimmy wrote a novel about his years as an undergraduate. Even though it was his first novel, it made the best-seller list and was critically-acclaimed.	<p>I heard that Jimmy made a splash.</p> <p>I heard that a splash was made by Jimmy.</p> <p>I heard about Jimmy's making of a splash.</p>
Eric bought his son a toy at the store, but he forgot to buy one for his daughter. The next day, in order to make it up to her, he let her pick out her own toy.	<p>I heard that Eric made amends.</p> <p>I heard that amends were made by Eric.</p> <p>I heard about Eric's making of amends.</p>
Amber was a scientist who was working on a cure for Alzheimer's disease. After much testing, she determined that a drug she had developed could delay the onset of the disease by up to 5 years.	<p>I heard that Amber made headway.</p> <p>I heard that headway was made by Amber.</p> <p>I heard about Amber's making of headway.</p>
Ryan and his friends decided to go to the symphony. His friends all bought their tickets, but by the time he went to buy them there were no more tickets available.	<p>I heard that Ryan missed the boat.</p> <p>I heard that the boat was missed by Ryan.</p> <p>I heard about Ryan's missing of the boat.</p>
Thelma was an attorney representing a defendant who claimed that excessive burping caused him to commit the crime. As a result, several defendants with similar stories sought her services.	<p>I heard that Thelma opened the floodgates.</p> <p>I heard that the floodgates were opened by Thelma.</p> <p>I heard about Thelma's opening of the floodgates.</p>

Context Paragraph	Stimulus Sentences
Gary was a very famous politician. One evening, he made a speech that was particularly elegant and really caught people's attention.	I heard that Gary packed a punch. I heard that a punch was packed by Gary. I heard about Gary's packing of a punch.
Alice was very upset after having a fight with her boyfriend. She spent the next several hours hopping from bar to bar and dancing with every guy she could find.	I heard that Alice painted the town. I heard that the town was painted by Alice. I heard about Alice's painting of the town.
Bradley was always very attentive in class, taking notes during every lecture. Not surprisingly, he always did well on quizzes and tests.	I heard that Bradley paid attention. I heard that attention was paid by Bradley. I heard about Bradley's paying of attention.
Anne was at a restaurant with her friends. The waitress warned her that the veggie burger was not very good, so she ordered a salad instead.	I heard that Anne paid heed. I heard that heed was paid by Anne. I heard about Anne's paying of heed.
Don was accused by his wife of putting an empty milk carton back in the fridge. When she confronted him about it, he kept quiet so he wouldn't get himself into trouble.	I heard that Don pled the fifth. I heard that the fifth was pled by Don. I heard about Don's pleading of the fifth.
Bill and Lauren had been dating for three years, and Lauren wondered if they would ever get engaged. One evening, Bill finally asked her to marry him.	I heard that Bill popped the question. I heard that the question was popped by Bill. I heard about Bill's popping of the question.
Shawn's daughter had just graduated from high school and was looking for a job. He knew someone who worked at a local law firm and got her a job as a receptionist in the office.	I heard that Shawn pulled strings. I heard that strings were pulled by Shawn. I heard about Shawn's pulling of strings.
Leslie surprised all of her friends by getting a tattoo. They were even more surprised when they found out that it was on her chin.	I heard that Leslie raised eyebrows. I heard that eyebrows were raised by Leslie. I heard about Leslie's raising of eyebrows.
Wayne took a job at a company that had been in business for over fifty years without ever updating its manufacturing methods. Even though he had been told that the managers were uncomfortable with change, he decided to bring up a more efficient method at the shareholders' meeting.	I heard that Wayne rocked the boat. I heard that the boat was rocked by Wayne. I heard about Wayne's rocking of the boat.
Crystal was raised in a racist neighborhood. After moving away and going to college, she realized that racism was harmful and unjustified.	I heard that Crystal saw the light. I heard that the light was seen by Crystal. I heard about Crystal's seeing of the light.
Helen was going to a party, but she was running late because she couldn't decide what to wear. She really had to hurry to make it to the party on time.	I heard that Helen shook a leg. I heard that a leg was shaken by Helen. I heard about Helen's shaking of a leg.

Context Paragraph	Stimulus Sentences
Jesse always stopped by the front desk when he got to work in the morning. He chatted briefly with the receptionist before heading to his office.	I heard that Jesse shot the breeze. I heard that the breeze was shot by Jesse. I heard about Jesse's shooting of the breeze.
Tony's best friend told him about his secret crush on Kathleen. Rather than keeping the secret, he blurted the news to Kathleen.	I heard that Tony spilled the beans. I heard that the beans were spilled by Tony. I heard about Tony's spilling of the beans.
Travis had always wanted to have children, but his wife had had several miscarriages. Finally, after many years, their first healthy child was born.	I heard that Travis started a family. I heard that a family was started by a family. I heard about Travis's starting of a family.
Elaine had a bit part in an off-Broadway show. In spite of the fact that she had a minor role, she was widely lauded as the star.	I heard that Elaine stole the show. I heard that the show was stolen by Elaine. I heard about Elaine's stealing of the show.
Fred knew that Diane loved dogs. When he asked her for a donation for his charity, he made sure to point out that the money would go to helping abandoned dogs find homes.	I heard that Fred struck a chord. I heard that a chord was struck by Fred. I heard about Fred's striking of a chord.
Alfred had been running the family business for 50 years. He finally decided to let his son take over, allowing Alfred to relax.	I heard that Alfred took a back seat. I heard that a back seat was taken by Alfred. I heard about Alfred's taking of a back seat.
Gladys really liked one of the boys in her homeroom class. After waiting several weeks to see if he would ask her out, she took matters into her own hands and asked him herself.	I heard that Gladys took a chance. I heard that a chance was taken by Gladys. I heard about Gladys's taking of a chance.
Barry was very upset about a comment he overheard. He decided to go for a long run, and he finally managed to relax.	I heard that Barry took a chill pill. I heard that a chill pill was taken by Barry. I heard about Barry's taking of a chill pill.
Jeff was out fishing along a river when he had to go to the bathroom. He went into the nearby bushes and relieved himself.	I heard that Jeff took a leak. I heard that a leak was taken by Jeff. I heard about Jeff's taking of a leak.
Cynthia made plans to have lunch with her advisor to talk about the paper she was writing. Unfortunately, something came up at the last minute, and she had to reschedule.	I heard that Cynthia took a raincheck. I heard that a raincheck was taken by Cynthia. I heard about Cynthia's taking of a raincheck.
Anita was walking down the street when she heard gunfire nearby. She ducked inside the nearest doorway where she waited until the shooting stopped.	I heard that Anita took cover. I heard that cover was taken by Anita. I heard about Anita's taking of cover.
Samuel was feeling very down about his job search. Then one week he had three job interviews, which greatly lifted his spirits.	I heard that Samuel took heart. I heard that heart was taken by Samuel. I heard about Samuel's taking of heart.

Context Paragraph	Stimulus Sentences
Pamela was working very hard on a project. Her boss told her to take a break, and she happily obliged.	I heard that Pamela took ten. I heard that ten was taken by Pamela. I heard about Pamela's taking of ten.
Todd had always wanted to travel the world. After he graduated from college, he decided to fulfill his dream.	I heard that Todd took the plunge. I heard that the plunge was taken by Todd. I heard about Todd's taking of the plunge.
Patricia and her coworkers went out to dinner after work on Friday. Even though they should have been starting the weekend, conversation kept drifting back to work projects.	I heard that Patricia talked shop. I heard that shop was talked by Patricia. I heard about Patricia's talking of shop.
Eugene would do anything he could to win on the basketball court. He often taunted and insulted his opponents to try to throw them off their game.	I heard that Eugene talked trash. I heard that trash was talked by Eugene. I heard about Eugene's talking of trash.
Barbara wanted to throw a surprise party for her parents' 50th wedding anniversary. She checked with her brother and sister to see if they thought it was a good idea.	I heard that Barbara tested the waters. I heard that the waters were tested by Barbara. I heard about Barbara's testing of the waters.
Joel was president of a private college. Many of the students were angry because he proposed major tuition increases, something the students strongly resisted because their tuition was already higher than the national average.	I heard that Joel touched a nerve. I heard that a nerve was touched by Joel. I heard about Joel's touching of a nerve.
Brenda had dreams of a better life, but her current job did not pay well enough to help her move up. She lived from paycheck to paycheck without making much progress.	I heard that Brenda treaded water. I heard that water was treaded by Brenda. I heard about Brenda's treading of water.
Donald knew that his daughter was drinking alcohol, even though she was only 18. He chose to ignore her behavior.	I heard that Donald turned a blind eye. I heard that a blind eye was turned by Donald. I heard about Donald's turning of a blind eye.
Judith was a very cheap boss. She repeatedly ignored her employees' requests for a paid lunch break.	I heard that Judith turned a deaf ear. I heard that a deaf ear was turned by Judith. I heard about Judith's turning of a deaf ear.
Ralph was a model with the kind of face that made everyone stop and stare. When he walked down the street, everyone looked.	I heard that Ralph turned heads. I heard that heads were turned by Ralph. I heard about Ralph's turning of heads.
Steve was being filmed by a camera crew. He surprised them all by grabbing a camera of his own and filming them instead.	I heard that Steve turned the tables. I heard that the tables were turned by Steve. I heard about Steve's turning of the tables.

## APPENDIX B: STIMULI AND FILLERS FOR EXPERIMENT TWO

### B.1 Stimuli

In the following table, A-C = Active Congruent, P-C = Passive Congruent, A-I = Active Incongruent, and P-I = Passive Incongruent.

	First Clause	Second Clause
A-C	Deborah beat a dead horse,	and she talked nonstop about winning the game.
P-C	A dead horse was beaten by Deborah,	and she talked nonstop about winning the game.
A-I	Deborah planted a fruit tree,	and she talked nonstop about winning the game.
P-I	A fruit tree was planted by Deborah,	and she talked nonstop about winning the game.
A-C	Elaine bit the bullet,	and the project succeeded as a result.
P-C	The bullet was bitten by Elaine,	and the project succeeded as a result.
A-I	Elaine saw the sneakers,	and the project succeeded as a result.
P-I	The sneakers were seen by Elaine, Did Elaine's project succeed?	and the project succeeded as a result. Yes No
A-C	Lucille bit the dust,	and her family and friends were shocked and upset.
P-C	The dust was bitten by Lucille,	and her family and friends were shocked and upset.
A-I	Lucille shook the rug,	and her family and friends were shocked and upset.
P-I	The rug was shaken by Lucille, Who was upset?	and her family and friends were shocked and upset. Lucille Lucille's family
A-C	Jerry broke the ice,	and everyone in the room felt more comfortable.
P-C	The ice was broken by Jerry,	and everyone in the room felt more comfortable.
A-I	Jerry pocketed the money,	and everyone in the room felt more comfortable.
P-I	The money was pocketed by Jerry, What did Jerry do?	and everyone in the room felt more comfortable. Eat ice cream Make people comfortable
A-C	Lisa broke the news,	and this made her parents very angry.
P-C	The news was broken by Lisa,	and this made her parents very angry.
A-I	Lisa loved the statue,	and this made her parents very angry.
P-I	The statue was loved by Lisa,	and this made her parents very angry.
A-C	Bradley bridged the gap,	and his boss was pleased with the connections that he had made.
P-C	The gap was bridged by Bradley,	and his boss was pleased with the connections that he had made.
A-I	Bradley groomed the dog,	and his boss was pleased with the connections that he had made.
P-I	The dog was groomed by Bradley, Who was pleased?	and his boss was pleased with the connections that he had made. Bradley Bradley's boss

	First Clause	Second Clause
A-C	Edward busted a move,	and everyone at the party was impressed with his dancing.
P-C	A move was busted by Edward,	and everyone at the party was impressed with his dancing.
A-I	Edward poked a lizard,	and everyone at the party was impressed with his dancing.
P-I	A lizard was poked by Edward,	and everyone at the party was impressed with his dancing.
A-C	Jimmy caught fire,	and his hair and clothes were singed.
P-C	Fire was caught by Jimmy,	and his hair and clothes were singed.
A-I	Jimmy froze water,	and his hair and clothes were singed.
P-I	Water was frozen by Jimmy,	and his hair and clothes were singed.
A-C	Eric cleared the air,	and the tension between him and his girlfriend was broken.
P-C	The air was cleared by Eric,	and the tension between him and his girlfriend was broken.
A-I	Eric chose the recipient,	and the tension between him and his girlfriend was broken.
P-I	The recipient was chosen by Eric,	and the tension between him and his girlfriend was broken.
A-C	Nicholas counted heads,	and that ensured that no one was left behind.
P-C	Heads were counted by Nicholas,	and that ensured that no one was left behind.
A-I	Nicholas built houses,	and that ensured that no one was left behind.
P-I	Houses were built by Nicholas,	and that ensured that no one was left behind.
A-C	Victor drew a blank,	and he missed a question on the test as a result.
P-C	A blank was drawn by Victor,	and he missed a question on the test as a result.
A-I	Victor stole a hammer,	and he missed a question on the test as a result.
P-I	A hammer was stolen by Victor,	and he missed a question on the test as a result.
A-C	Joan drew the line,	and her employees respected her for it.
P-C	The line was drawn by Joan,	and her employees respected her for it.
A-I	Joan swung the bat,	and her employees respected her for it.
P-I	The bat was swung by Joan,	and her employees respected her for it.
A-C	Jesse drove a hard bargain,	and that allowed him to get a good deal on the futon.
P-C	A hard bargain was driven by Jesse,	and that allowed him to get a good deal on the futon.
A-I	Jesse sliced a ripe pear,	and that allowed him to get a good deal on the futon.
P-I	A ripe pear was sliced by Jesse,	and that allowed him to get a good deal on the futon.
A-C	Sarah dropped a bomb,	and all of her friends were surprised by the news.
P-C	A bomb was dropped by Sarah,	and all of her friends were surprised by the news.
A-I	Sarah carried a backpack,	and all of her friends were surprised by the news.
P-I	A backpack was carried by Sarah,	and all of her friends were surprised by the news.
A-C	Jacqueline dropped the ball,	and her boss was left in a bad position.
P-C	The ball was dropped by Jacqueline,	and her boss was left in a bad position.
A-I	Jacqueline clarified the point,	and her boss was left in a bad position.
P-I	The point was clarified by Jacqueline,	and her boss was left in a bad position.

	First Clause	Second Clause
A-C	Tony faced the music,	and he endured public humiliation as a result.
P-C	The music was faced by Tony,	and he endured public humiliation as a result.
A-I	Tony broke the vase,	and he endured public humiliation as a result.
P-I	The vase was broken by Tony, Was Tony publicly humiliated?	and he endured public humiliation as a result. No    Yes
A-C	Amy fit the bill,	and the agent said she was exactly what he was looking for.
P-C	The bill was fit by Amy,	and the agent said she was exactly what he was looking for.
A-I	Amy released the butterfly,	and the agent said she was exactly what he was looking for.
P-I	The butterfly was released by Amy,	and the agent said she was exactly what he was looking for.
A-C	Leslie gained ground,	and her position in the mayoral race steadily improved.
P-C	Ground was gained by Leslie,	and her position in the mayoral race steadily improved.
A-I	Leslie practiced karate,	and her position in the mayoral race steadily improved.
P-I	Karate was practiced by Leslie, What was Leslie doing?	and her position in the mayoral race steadily improved. Running for mayor    Digging in the ground
A-C	Norman went the distance,	and his project succeeded because of his resolve.
P-C	The distance was gone by Norman,	and his project succeeded because of his resolve.
A-I	Norman rode the horse,	and his project succeeded because of his resolve.
P-I	The horse was ridden by Norman,	and his project succeeded because of his resolve.
A-C	Elizabeth hung a left,	and this allowed her to avoid a traffic jam.
P-C	A left was hung by Elizabeth,	and this allowed her to avoid a traffic jam.
A-I	Elizabeth published a manuscript,	and this allowed her to avoid a traffic jam.
P-I	A manuscript was published by Elizabeth,	and this allowed her to avoid a traffic jam.
A-C	Joshua hung a right,	and turning right allowed him to reach his destination.
P-C	A right was hung by Joshua,	and turning right allowed him to reach his destination.
A-I	Joshua forwarded an email,	and turning right allowed him to reach his destination.
P-I	An email was forwarded by Joshua,	and turning right allowed him to reach his destination.
A-C	Joe hit rock bottom,	and he cried for several hours.
P-C	Rock bottom was hit by Joe,	and he cried for several hours.
A-I	Joe ate raw sushi,	and he cried for several hours.
P-I	Raw sushi was eaten by Joe,	and he cried for several hours.
A-C	Dorothy hit the hay,	and she fell asleep almost immediately.
P-C	The hay was hit by Dorothy,	and she fell asleep almost immediately.
A-I	Dorothy cancelled the performance,	and she fell asleep almost immediately.
P-I	The performance was cancelled by Dorothy,	and she fell asleep almost immediately.

	First Clause	Second Clause
A-C	Aaron hit the road,	and he waved goodbye to his family as he drove away.
P-C	The road was hit by Aaron,	and he waved goodbye to his family as he drove away.
A-I	Aaron addressed the audience,	and he waved goodbye to his family as he drove away.
P-I	The audience was addressed by Aaron,	and he waved goodbye to his family as he drove away.
A-C	Travis hit the roof,	and his yelling could be heard by the neighbors next door.
P-C	The roof was hit by Travis,	and his yelling could be heard by the neighbors next door.
A-I	Travis traced the border,	and his yelling could be heard by the neighbors next door.
P-I	The border was traced by Travis,	and his yelling could be heard by the neighbors next door.
A-C	James hit the sack,	and his snores could be heard in a matter of moments.
P-C	The sack was hit by James,	and his snores could be heard in a matter of moments.
A-I	James confronted the bully,	and his snores could be heard in a matter of moments.
P-I	The bully was confronted by James,	and his snores could be heard in a matter of moments.
A-C	Janice jumped the gun,	and she later apologized for being impatient.
P-C	The gun was jumped by Janice,	and she later apologized for being impatient.
A-I	Janice arranged the flowers,	and she later apologized for being impatient.
P-I	The flowers were arranged by Janice,	and she later apologized for being impatient.
A-C	David kicked the bucket,	and his funeral will be next week.
P-C	The bucket was kicked by David,	and his funeral will be next week.
A-I	David labeled the items,	and his funeral will be next week.
P-I	The items were labeled by David,	and his funeral will be next week.
A-C	Helen kicked the habit,	and her breath no longer smelled like cigarette smoke.
P-C	The habit was kicked by Helen,	and her breath no longer smelled like cigarette smoke.
A-I	Helen held the handlebars,	and her breath no longer smelled like cigarette smoke.
P-I	The handlebars were held by Helen, Did Helen's breath smell like smoke?	and her breath no longer smelled like cigarette smoke. Yes    No
A-C	Amber knew the ropes,	and she was able to easily train her replacement.
P-C	The ropes were known by Amber,	and she was able to easily train her replacement.
A-I	Amber totaled the columns,	and she was able to easily train her replacement.
P-I	The columns were totaled by Amber, Who got trained?	and she was able to easily train her replacement. Amber    Amber's replacement
A-C	Grace lent a hand,	and her neighbor was very thankful for the help.
P-C	A hand was lent by Grace,	and her neighbor was very thankful for the help.
A-I	Grace submitted an application,	and her neighbor was very thankful for the help.
P-I	An application was submitted by Grace,	and her neighbor was very thankful for the help.
A-C	Rachel made a killing,	and she saved most of the money for retirement.
P-C	A killing was made by Rachel,	and she saved most of the money for retirement.
A-I	Rachel flipped a coin,	and she saved most of the money for retirement.
P-I	A coin was flipped by Rachel,	and she saved most of the money for retirement.



	First Clause	Second Clause
A-C	Alfred made a splash,	and his performance immediately received critical acclaim.
P-C	A splash was made by Alfred,	and his performance immediately received critical acclaim.
A-I	Alfred bought a ring,	and his performance immediately received critical acclaim.
P-I	A ring was bought by Alfred,	and his performance immediately received critical acclaim.
	What did Alfred do?	Splash someone    Perform well
A-C	Betty made amends,	and her sister felt much better once they had made up.
P-C	Amends were made by Betty,	and her sister felt much better once they had made up.
A-I	Betty wrapped presents,	and her sister felt much better once they had made up.
P-I	Presents were wrapped by Betty,	and her sister felt much better once they had made up.
A-C	Rodney made headway,	and he knew he would finally finish the report.
P-C	Headway was made by Rodney,	and he knew he would finally finish the report.
A-I	Rodney directed traffic,	and he knew he would finally finish the report.
P-I	Traffic was directed by Rodney,	and he knew he would finally finish the report.
A-C	Antonio packed a punch,	and his speech really got people's attention.
P-C	A punch was packed by Antonio,	and his speech really got people's attention.
A-I	Antonio drank a soda,	and his speech really got people's attention.
P-I	A soda was drunk by Antonio,	and his speech really got people's attention.
A-C	Fred painted the town,	and he had a hangover the next day.
P-C	The town was painted by Fred,	and he had a hangover the next day.
A-I	Fred standardized the test,	and he had a hangover the next day.
P-I	The test was standardized by Fred,	and he had a hangover the next day.
A-C	Cindy paid attention,	and she got good grades in school as a result.
P-C	Attention was paid by Cindy,	and she got good grades in school as a result.
A-I	Cindy dried basil,	and she got good grades in school as a result.
P-I	Basil was dried by Cindy,	and she got good grades in school as a result.
A-C	Paula pled the fifth,	and her silence just made everyone more suspicious.
P-C	The fifth was pled by Paula,	and her silence just made everyone more suspicious.
A-I	Paula scheduled the meeting,	and her silence just made everyone more suspicious.
P-I	The meeting was scheduled by Paula,	and her silence just made everyone more suspicious.
	Who was silent?	Paula    Paula's friends
A-C	Larry popped the question,	and he was pleased when his girlfriend said 'Yes'.
P-C	The question was popped by Larry,	and he was pleased when his girlfriend said 'Yes'.
A-I	Larry sabotaged the production,	and he was pleased when his girlfriend said 'Yes'.
P-I	The production was sabotaged by Larry,	and he was pleased when his girlfriend said 'Yes'.
	Who said 'Yes'?	Larry    Larry's girlfriend

	First Clause	Second Clause
A-C	Ellen pulled strings,	and she was able to get her friend a good job.
P-C	Strings were pulled by Ellen,	and she was able to get her friend a good job.
A-I	Ellen herded sheep,	and she was able to get her friend a good job.
P-I	Sheep were herded by Ellen,	and she was able to get her friend a good job.
A-C	Crystal raised eyebrows,	and her husband was thrown out of the country club.
P-C	Eyebrows were raised by Crystal,	and her husband was thrown out of the country club.
A-I	Crystal wrote poems,	and her husband was thrown out of the country club.
P-I	Poems were written by Crystal,	and her husband was thrown out of the country club.
A-C	Marcus rocked the boat,	and his boss was upset.
P-C	The boat was rocked by Marcus,	and his boss was upset.
A-I	Marcus carved the turkey,	and his boss was upset.
P-I	The turkey was carved by Marcus,	and his boss was upset.
A-C	Anna saw the light,	and her racist views were soon completely forgotten.
P-C	The light was seen by Anna,	and her racist views were soon completely forgotten.
A-I	Anna cut the ribbon,	and her racist views were soon completely forgotten.
P-I	The ribbon was cut by Anna,	and her racist views were soon completely forgotten.
A-C	Richard spilled the beans,	and his friend was mad at him for telling his secret.
P-C	The beans were spilled by Richard,	and his friend was mad at him for telling his secret.
A-I	Richard refused the offer,	and his friend was mad at him for telling his secret.
P-I	The offer was refused by Richard,	and his friend was mad at him for telling his secret.
A-C	Donna started a family,	and she loved her new baby more than life itself.
P-C	A family was started by Donna,	and she loved her new baby more than life itself.
A-I	Donna devised a scheme,	and she loved her new baby more than life itself.
P-I	A scheme was devised by Donna,	and she loved her new baby more than life itself.
A-C	Peter stole the show,	and he received more attention from critics than the lead actor.
P-C	The show was stolen by Peter,	and he received more attention from critics than the lead actor.
A-I	Peter ironed the clothes,	and he received more attention from critics than the lead actor.
P-I	The clothes were ironed by Peter,	and he received more attention from critics than the lead actor.
A-C	Anne struck a chord,	and the audience at her performance laughed and applauded.
P-C	A chord was struck by Anne,	and the audience at her performance laughed and applauded.
A-I	Anne initiated an investigation,	and the audience at her performance laughed and applauded.
P-I	An investigation was initiated by Anne,	and the audience at her performance laughed and applauded.

	First Clause	Second Clause
A-C	Edith took a back seat,	and her daughter took over the family business.
P-C	A back seat was taken by Edith,	and her daughter took over the family business.
A-I	Edith freed a trapped bird,	and her daughter took over the family business.
P-I	A trapped bird was freed by Edith,	and her daughter took over the family business.
	Who took over the business?	Edith's daughter   Edith
A-C	Stanley took a chance,	and his friends commended him for his bravery.
P-C	A chance was taken by Stanley,	and his friends commended him for his bravery.
A-I	Stanley joined a club,	and his friends commended him for his bravery.
P-I	A club was joined by Stanley,	and his friends commended him for his bravery.
A-C	Stephanie took a chill pill,	and it helped her calm down and finish her presentation.
P-C	A chill pill was taken by Stephanie,	and it helped her calm down and finish her presentation.
A-I	Stephanie clocked a fast time,	and it helped her calm down and finish her presentation.
P-I	A fast time was clocked by Stephanie,	and it helped her calm down and finish her presentation.
	Did Stephanie complete her presentation?	No   Yes
A-C	Ronald took a leak,	and he felt much better after he emptied his bladder.
P-C	A leak was taken by Ronald,	and he felt much better after he emptied his bladder.
A-I	Ronald danced a jig,	and he felt much better after he emptied his bladder.
P-I	A jig was danced by Ronald,	and he felt much better after he emptied his bladder.
A-C	Jeremy took a raincheck,	and he postponed his lunch plans with his coworker.
P-C	A raincheck was taken by Jeremy,	and he postponed his lunch plans with his coworker.
A-I	Jeremy housed a refugee,	and he postponed his lunch plans with his coworker.
P-I	A refugee was housed by Jeremy,	and he postponed his lunch plans with his coworker.
A-C	Alice took cover,	and she ducked into a doorway when she heard gunfire.
P-C	Cover was taken by Alice,	and she ducked into a doorway when she heard gunfire.
A-I	Alice sought recognition,	and she ducked into a doorway when she heard gunfire.
P-I	Recognition was sought by Alice,	and she ducked into a doorway when she heard gunfire.
A-C	Wayne took heart,	and his job search improved as his attitude improved.
P-C	Heart was taken by Wayne,	and his job search improved as his attitude improved.
A-I	Wayne suffered humiliation,	and his job search improved as his attitude improved.
P-I	Humiliation was suffered by Wayne,	and his job search improved as his attitude improved.
A-C	Lillian took ten,	and then she returned to practice after her break.
P-C	Ten was taken by Lillian,	and then she returned to practice after her break.
A-I	Lillian promoted peace,	and then she returned to practice after her break.
P-I	Peace was promoted by Lillian,	and then she returned to practice after her break.
	What did Lillian do?	Steal ten items   Take a break
A-C	Bill took the plunge,	and his family was excited about his new job in Japan.
P-C	The plunge was taken by Bill,	and his family was excited about his new job in Japan.
A-I	Bill browsed the internet,	and his family was excited about his new job in Japan.
P-I	The internet was browsed by Bill,	and his family was excited about his new job in Japan.
	Where was Bill's new job?	Japan   Near his family

	First Clause	Second Clause
A-C	Roger talked trash,	and his opponents were thrown off by his constant banter.
P-C	Trash was talked by Roger,	and his opponents were thrown off by his constant banter.
A-I	Roger hated math,	and his opponents were thrown off by his constant banter.
P-I	Math was hated by Roger,	and his opponents were thrown off by his constant banter.
A-C	Shawn tested the waters,	and the investors agreed that his idea was really good.
P-C	The waters were tested by Shawn,	and the investors agreed that his idea was really good.
A-I	Shawn chased the bus,	and the investors agreed that his idea was really good.
P-I	The bus was chased by Shawn,	and the investors agreed that his idea was really good.
A-C	Gary touched a nerve,	and some of his students were offended.
P-C	A nerve was touched by Gary,	and some of his students were offended.
A-I	Gary operated a machine,	and some of his students were offended.
P-I	A machine was operated by Gary,	and some of his students were offended.
A-C	Thelma treaded water,	and she lived from paycheck to paycheck in constant debt.
P-C	Water was treaded by Thelma,	and she lived from paycheck to paycheck in constant debt.
A-I	Thelma attained excellence,	and she lived from paycheck to paycheck in constant debt.
P-I	Excellence was attained by Thelma,	and she lived from paycheck to paycheck in constant debt.
A-C	Ryan turned a blind eye,	and he pretended not to know his friend cheated on the test.
P-C	A blind eye was turned by Ryan,	and he pretended not to know his friend cheated on the test.
A-I	Ryan served a delicious meal,	and he pretended not to know his friend cheated on the test.
P-I	A delicious meal was served by Ryan,	and he pretended not to know his friend cheated on the test.
A-C	Teresa turned heads,	and everybody stared when she walked by.
P-C	Heads were turned by Teresa,	and everybody stared when she walked by.
A-I	Teresa enjoyed pasta,	and everybody stared when she walked by.
P-I	Pasta was enjoyed by Teresa, Who was staring?	and everybody stared when she walked by. Teresa    Everyone else
A-C	Christine turned the tables,	and she came back to win the game after being down 40-love.
P-C	The tables were turned by Christine,	and she came back to win the game after being down 40-love.
A-I	Christine cured the disease,	and she came back to win the game after being down 40-love.
P-I	The disease was cured by Christine,	and she came back to win the game after being down 40-love.

## B.2 Fillers

In the following table, A = Active and P = Passive. All fillers are semantically congruent.

	First Clause	Second Clause
A	William argued the point, Who won the argument?	and his friends eventually gave in. William. William's friends.
A	Thomas read a how-to book,	and it helped him with his project.
A	Kevin updated his wardrobe,	and it made him more popular with the ladies.
A	Timothy baffled his opponent,	and he quickly won the game of chess.
A	Sandra pinched her son's cheeks, Who was embarrassed?	and he was very embarrassed that his friends saw. Sandra. Sandra's son.
A	Laura rested her eyes,	and she was soon fast asleep.
A	Jessica boarded the jet,	and she was happy to be sitting in first class.
A	Sharon ranked her priorities,	and this helped her get things accomplished.
P	Important information was omitted by Robert,	and he was accused of perjury.
P	The picture was cropped by Michael, What did Michael do?	and it looked much better when he was done. Paint a picture. Crop a picture.
P	A comeback was engineered by the team,	and they won the game in the final seconds.
P	The action was filmed by Joseph,	and then he posted the video on YouTube.
P	The plane was landed by Daniel,	and then it taxied along the runway to the gate.
P	A garden was grown by Mark,	and he frequently made meals from his fresh vegetables.
P	Bubbles were blown by the child, Who blew bubbles?	and they popped when he tried to catch them. The child. The child's mother.
P	The crime scene was investigated by the police,	and the cause of death was ruled to be a homicide.
P	The exam was administered by Kenneth,	and many of his students passed with flying colors.
P	The numbers were added by Steven,	and they totaled exactly three hundred.
P	The puppies were fed by Anthony,	and then they were ready for a nap.
P	The wall was punched by Jason, What did Jason do?	and his hand was in a cast for weeks. Punch the wall. Punch his friend.
P	Matthew was scared by the snake,	and he screamed like a little girl.
P	A hole was dug by Frank,	and he planted the young sapling.
P	Vehicles were swapped by Mary,	and she drove the truck to work instead of the car.
P	The applicants were screened by Patricia,	and she selected the most qualified person for the job.
P	The pasture was enclosed by a fence, What was kept in the pasture?	and the horses were confined. Horses. Cows.

	First Clause	Second Clause
P	The information was absorbed by Barbara,	and she was soon ready to start her new job.
P	The ingredients were measured by Jennifer,	and the cake was soon baking in the oven.
P	The important passages were underlined by Susan,	and that helped her when it came time to study later.
P	The night sky was searched by Margaret,	and she saw many shooting stars.
P	Music was composed by Nancy, Who was playing music?	and soon it was being rehearsed by a string quartet. Nancy. A string quartet.
P	The soup was tasted by Karen,	and she decided that it needed a bit more salt.
P	The plans were finalized by Carol,	and she was able to relax the evening before the event.
P	The employees were motivated by Kimberly,	and their performance improved as a result.
P	The needle was threaded by Shirley,	and she began sewing a button on her shirt.
P	The new furnace was installed by Brian, What did Brian do?	and the house was soon warming up. Install the furnace. Paint the house.
P	The toaster was fixed by Angela,	and her children were happy to have their breakfast.
P	Money was owed by Melissa,	and she had to work two jobs to pay off her loans.
P	The results were manipulated by Ruth,	and her supervisor fired her for being dishonest.
P	The occasion was celebrated by Christopher,	and everyone wore party hats.
P	Peaches were canned by Michelle, What did Michelle do?	and she was able to enjoy them for the rest of the year. Eat oranges. Can peaches.

## APPENDIX C: ANALYZING THE COCA RESULTS AS BINARY

One significant difference between the two studies reported here is the nature of the data: The Google Books results are binary—either an idiom occurred in a given variation or it did not—while the COCA results report the actual number of times an idiom occurred in a given variation. This difference affects the statistical measures chosen to analyze the results. One might wonder what the results would look like if both datasets were treated as binary so that identical statistical measures could be used across the two studies. As reported below, these statistical measures provide no additional insight into the COCA data, which remain less straightforward than the Google Books data under these analyses.

### C.1 Fisher’s Exact Test for COCA

In order to directly compare across the two corpora, the data from Table 3.7 above were converted to binary data, where any number greater than one was collapsed to “1” and all zeroes remain “0”. The procedure used to create 2x2 contingency tables for each pair of variations, as outlined for the Google Books data in Section 3.3.2.2, was then applied to the resulting binary data.

Fisher’s Exact Test is used to provide simple comparisons between two datasets, where a significant  $p$ -value indicates dependence between two variations. Again, we use Bonferroni correction to set the significance level at  $p < 0.005$ . Table C.1 shows the Fisher’s Exact Test results for COCA.

	<b>Modification</b>	<b>Passivization</b>	<b>Object Incorporation</b>	<b>Nominal Gerundization</b>
<b>Quantification</b>	$p=0.014$	$p=0.0011^*$	$p=0.052$	$p=0.37$
<b>Modification</b>		$p=0.029$	$p=0.075$	$p=0.57$
<b>Passivization</b>			$p=0.0059$	$p=0.029$
<b>Object Incorporation</b>				$p=0.014$

**Table C.1: Fisher's Exact Test results for COCA**

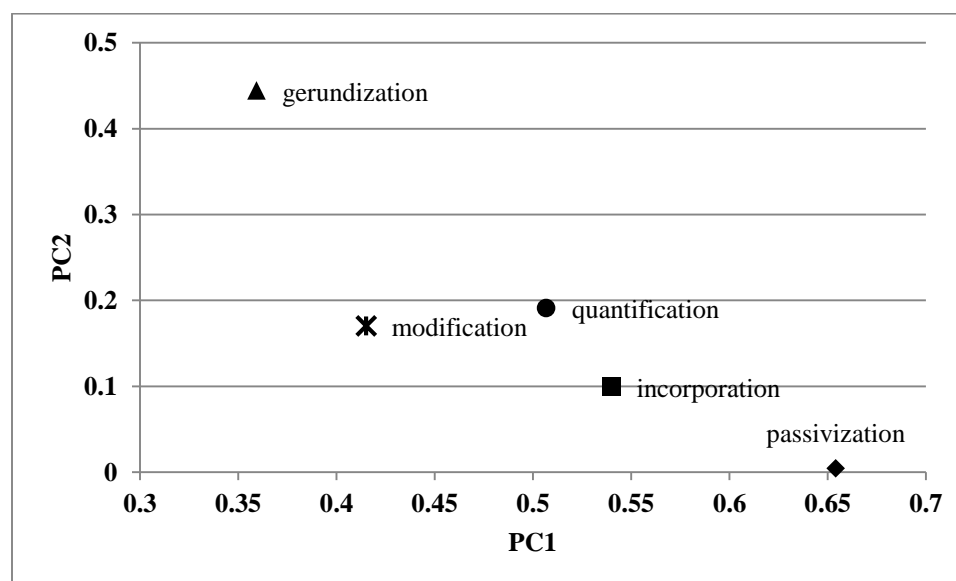
The results reported in Table C.1 are somewhat unexpected. First, the dependence between quantification and passivization does not represent any of the subset relationships under discussion here, though all of the statistical measures found a significant correlation. As speculated above, it is not particularly surprising that these two variations would be correlated given the well-known relationship between passivization and definiteness. Furthermore, none of the variation pairs involved in subset relationships meet the minimum significance threshold of  $p < 0.005$ . With passivization, incorporation, and gerundization, the subset relationships are weaker here than in Google Books, with more exceptions; however, only the passivization~incorporation relationship even approaches significance. Additionally, in the case of modification, our *ubervariation*, this variation is a perfect superset of each of the other variations, as shown by the conditional probabilities in Table 3.8, yet Fisher's Exact Test did not yield significant results here either.



Although the significance levels of Fisher's Exact Test seemed to mirror the high conditional probability values for Google Books, that is not the case here, and this was most likely a happy coincidence; conditional probability is a unidirectional relationship, whereas correlation is bidirectional.

## C.2 Multiple Correspondence Analysis for COCA

Recall from the discussion in Section 3.3.2.3 above that MCA is a type of dimensionality reduction. As such, it reduces the number of dimensions needed to explain a set of variables. In this case, the variables are the five variations under consideration. With the Google Books data, there was a clear division between the argument-structure-type variations and the modification-type variations. Figure C.1 shows the results of a Multiple Correspondence Analysis of the binary version of the COCA data.



**Figure C.1: Multiple Correspondence Analysis (MCA) for COCA**

Not surprisingly, all five variations have significant positive correlations with PC1: passivization ( $R^2=0.65$ ,  $p<.001$ ), incorporation ( $R^2=0.54$ ,  $p<.001$ ), quantification ( $R^2=0.51$ ,  $p<.001$ ), modification ( $R^2=0.42$ ,  $p<.001$ ), and gerundization ( $R^2=0.36$ ,  $p<.001$ ). Four of the five variations also correlate with the second dimension, PC2, although three of the correlations are much weaker: gerundization ( $R^2=0.44$ ,  $p<.001$ ), quantification ( $R^2=0.19$ ,  $p<.01$ ), modification ( $R^2=0.17$ ,  $p<.01$ ), and incorporation ( $R^2=0.10$ ,  $p<.05$ ).

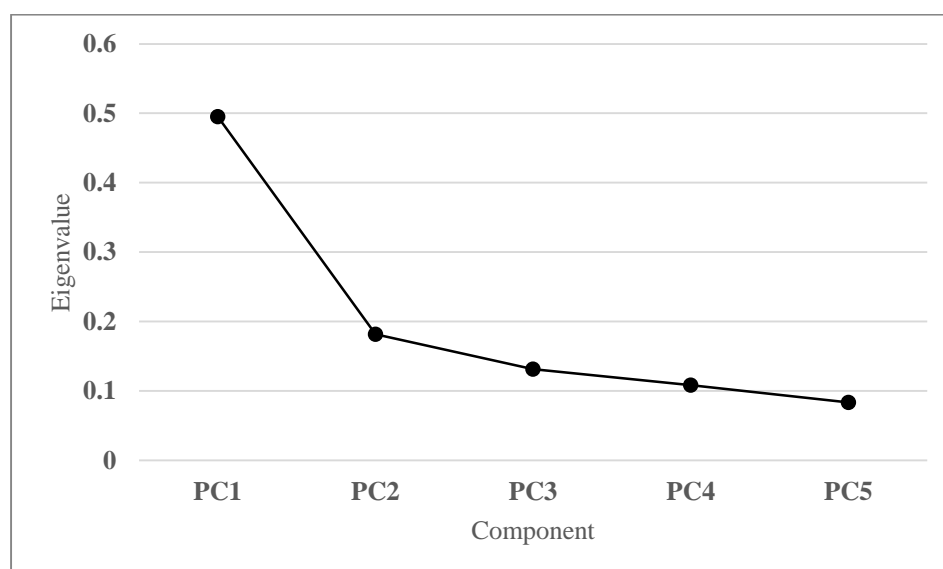
The most visually striking thing in this figure is the lack of pattern, a sharp contrast to the obvious clusters in the Google Books data in Figure 3.1 above. What we see instead is relatively equal spacing with respect to PC1, with all variations significantly correlated to this dimension. With respect to PC2, gerundization alone is strongly correlated, while modification, quantification, and incorporation are much more weakly correlated. Passivization is near zero, having no correlation with PC2. Neither of these dimensions seems to be distinguishing between the argument-structure-type and modification-type variations, as the dimensions did to various degrees in the MCA analysis of Google Books and the PCA analysis of COCA above.

The data table showing the eigenvalues and percentage of variance for each dimension from Figure C.1 is provided in Table C.2 below.

	<b>PC1</b>	<b>PC2</b>	<b>PC3</b>	<b>PC4</b>	<b>PC5</b>
<b>Eigenvalue</b>	0.50	0.18	0.13	0.12	0.08
<b>Percentage of Variance</b>	49.50	18.19	13.12	10.83	8.36
<b>Cumulative Percentage</b>	49.50	67.69	80.81	91.64	100.00

**Table C.2: Multiple Correspondence Analysis (MCA) for COCA**

The first component accounts for about 50% of the variance, and we see that the variations appear to be roughly equally spaced according to this dimension (the horizontal axis in Figure C.1). The remaining components account for 18%, 13%, 11%, and 8% of the variance, respectively. Unlike the Google Books analysis, where only three factors were needed to account for all but 7% of the variance, these data are much more diverse.



**Figure C.2: Scree Plot for COCA MCA**

The scree plot in Figure C.2 shows the “elbow” at PC2, indicating the PC1 alone should be retained for further analysis. However, the trend line does not level out at that point but retains a noticeable downward slope, suggesting that more of these dimensions could be playing a role. Although the data are remarkably muddled, one thing is clear: MCA on the binary data from COCA does not group the argument-structure-type and modification-type variations in the same way that the Google Books MCA did.

## REFERENCES

- Abeillé, Anne. 1995. The flexibility of French Idioms: A representation with lexicalized tree adjoining grammar. In Everaert, Martin and van der Linden, Erik-Jan and Schenk, André and Schreuder, Rob (ed.), *Idioms: Structural and psychological perspectives*, 15–42. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Alexiadou, Artemis. 2013. Nominal vs. verbal -ing constructions in the development of the English progressive. *English Linguistics Research* 2(2). 126–140.
- Alexiadou, Artemis, Elena Anagnostopoulou & Martin Everaert. 2004. *The unaccusativity puzzle: Explorations of the syntax-lexicon interface*. Oxford: Oxford University Press.
- Alexiadou, Artemis, Elena Anagnostopoulou & Florian Schäfer. 2015. *External arguments in transitivity alternations: A layering approach*. Oxford: Oxford University Press.
- Alexiadou, Artemis & Gianina Iordăchioaia. 2015. Idiomaticity and compositionality in deverbal compounds. Paper presented at the BCGL8: The Grammar of Idioms, Brussels, Belgium, 5 June 2015.
- Ayto, John. 2010. *Oxford dictionary of English idioms*. Oxford: Oxford University Press.
- Berman, Arlene. 1973. A constraint on tough-movement. In Corum, Claudia and Smith-Stark, T. Cedric and Weiser, Ann (ed.), *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society*, 34–43. Chicago: University of Chicago.
- Bever, Thomas G. & David J. Townsend. 1979. Perceptual mechanisms and formal properties of main and subordinate clauses. In Cooper, William E. and Walker, Edward C. T. (ed.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*, 159–226. Hillsdale, NJ: Erlbaum.
- Blais, Mary-Jane & Debra Titone. 2008. Idiom compositionality effects on word by word reading. Paper presented at the McGill's Canadian Conference for Linguistic Undergraduates, McGill University, Montreal, Quebec, 30 March 2008.
- Bobaljik, Jonathan David & Andrew Carnie. 1996. A minimalist approach to some problems of Irish word order. In Borsley, Robert D. and Roberts, Ian (ed.), *The Syntax of the Celtic Languages*, 223–240. Cambridge: Cambridge University Press.

- Bobrow, Samuel A. & Susan M. Bell. 1973. On catching on to idiomatic expressions. *Memory & Cognition* 1(3). 343–346.
- Boland, Julie E. 1997. The relationship between syntactic and semantic processes in sentence comprehension. *Language and Cognitive Processes* 12(4). 423–484.
- Bolinger, Dwight. 1977. Idioms have relations. *Forum Linguisticum* 2(2). 157–169.
- Borer, Hagit. 1998. Passive without theta-grids. In Lapointe, Steven G. and Brentari, Diane K. and Farrell, Patrick M. (ed.), *Morphology and Its Relation to Phonology and Syntax*, 60–99. Stanford: CSLI Publications.
- Brame, Michael. 1968. A new analysis of the relative clause: evidence for an interpretive theory. Unpublished manuscript, MIT.
- Braze, David, Donald Shankweiler, Weijia Ni & Laura Conway Palumbo. 2002. Readers' eye movements distinguish anomalies of form and content. *Journal of Psycholinguistic Research* 31(1). 25–44.
- Bruening, Benjamin. 2010. Ditransitive asymmetries and a theory of idiom formation. *Linguistic Inquiry* 41(4). 519–562.
- Bruening, Benjamin. 2013. By phrases in passives and nominals. *Syntax* 16(1). 1–41.
- Bruening, Benjamin. 2014. Word formation is syntactic: Adjectival passives in English. *Natural Language & Linguistic Theory* 32(2). 363–422.
- Burzio, Luigi. 1986. Italian syntax: A government-binding approach. Dordrecht: Reidel.
- Cacciari, Cristina & Patrizia Tabossi. 1988. The comprehension of idioms. *Journal of Memory and Language* 27(6). 668–683.
- Carlson, Greg. 1977. Amount relatives. *Language* 53. 520–542.
- Carrithers, Caroline. 1986. *The special status of non canonical sequences*. New York, NY: Columbia University dissertation.
- Carrithers, Caroline & Thomas G. Bever. 1984. Eye-fixation patterns during reading confirm theories of language comprehension. *Cognitive Science* 8(2). 157–172.
- Chafe, Wallace L. 1968. Idiomaticity as an anomaly in the Chomskyan paradigm. *Foundations of Language* 4(2). 109–127.

- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge: MIT press.
- Chomsky, Noam. 1970. Remarks on nominalization. In Jacobs, Roderick A. and Rosenbaum, Peter S. (ed.), *Readings in English transformational grammar*, 184–221. Waltham, MA: Ginn.
- Chomsky, Noam. 1981. Lectures on government and binding: The Pisa lectures. Dordrecht: Foris.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
- Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. (Oxford Studies in Comparative Syntax). New York/Oxford: Oxford University Press.
- Cuervo, María Cristina. 2003. *Datives at large*. Cambridge, MA: MIT dissertation.
- Cutting, J. Cooper & Kathryn Bock. 1997. That’s the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blends. *Memory & Cognition* 25(1). 57–71.
- Darnell, Regna. 1972. On “seeing the light.” *Papers in Linguistics* 5(2). 333–335.
- Davies, Mark. 2008-. The Corpus of Contemporary American English: 450 million words, 1990-present. Available online at <http://corpus.byu.edu/coca/>.
- Egan, Andy. 2008. Pretense for the complete idiom. *Noûs* 42(3). 381–409.
- Ernst, Thomas. 1981. Grist for the linguistic mill: Idioms and “extra” adjectives. *Journal of Linguistic Research* 1(3). 51–68.
- Fazly, Afsaneh, Paul Cook & Suzanne Stevenson. 2009. Unsupervised type and token identification of idiomatic expressions. *Computational Linguistics* 35(1). 61–103.
- Fellbaum, Christiane. 2002. VP idioms in the lexicon: Topics for research using a very large corpus. In Busemann, Stephan (ed.), *Proceedings of the KONVENS 2002 Conference*, 7–11. Saarbrücken, Germany: Deutsche Forschungszentrum für Künstliche Intelligenz.
- Fellbaum, Christiane (Ed.). 2007. *Idioms and collocations: Corpus-based linguistic and lexicographic studies*. Birmingham, UK: Continuum International Publishing Group.

- Fellbaum, Christiane, Alexander Geyken, Axel Herold, Fabian Koerner & Gerald Neumann. 2006. Corpus-based studies of German idioms and light verbs. *International Journal of Lexicography* 19(4). 349–360.
- Fillmore, Charles J. 1988. The mechanisms of “construction grammar.” *Annual Meeting of the Berkeley Linguistics Society* 14. 35–55.
- Folli, Raffaella & Heidi Harley. 2005. Flavors of v: Consuming results in Italian and English. In Kempchinsky, Paula and Slabakova, Roumyana (ed.), *Aspectual Enquiries*, 95–120. Springer: Dordrecht.
- Folli, Raffaella & Heidi Harley. 2007. Causation, obligation, and argument structure: On the nature of little v. *Linguistic Inquiry* 38(2). 197–238.
- Folli, Raffaella, Heidi Harley & Simin Karimi. 2005. Determinants of event type in Persian complex predicates. *Lingua* 115(10). 1365–1401.
- Forster, Kenneth I. & Jonathan C. Forster. 2003. DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers* 35(1). 116–124.
- Forster, Kenneth I. & Ilmar Olbrei. 1972. Semantic heuristics and syntactic analysis. *Cognition* 2(3). 319–347.
- Fraser, Bruce. 1970. Idioms within a transformational grammar. *Foundations of Language* 6(1). 22–42.
- Fraser, Bruce & John Robert Ross. 1970. Idioms and unspecified NP deletion. *Linguistic Inquiry* 1(2). 264–265.
- Frazier, Lyn, Lori Taft, Tom Roeper, Charles Clifton & Kate Ehrlich. 1984. Parallel structure: A source of facilitation in sentence comprehension. *Memory & Cognition* 12(5). 421–430.
- Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan A. Sag. 1985. *Generalized phrase structure grammar*. Oxford: Basil Blackwell.
- Ghameshi, Jila & Diane Massam. 1994. Lexical/syntactic relations without projection. *Linguistic Analysis* 24(3-4). 175–217.
- Gibbs, Jr. Raymond W. & Nandini P. Nayak. 1989. Psycholinguistic studies on the syntactic behavior of idioms. *Cognitive Psychology* 21(1). 100–138.

- Gibbs Jr., Raymond W. 1986. Skating on thin ice: Literal meaning and understanding idioms in conversation. *Discourse Processes* 9(1). 17–30.
- Gibbs Jr., Raymond W. & Gayle P. Gonzales. 1985. Syntactic frozenness in processing and remembering idioms. *Cognition* 20(3). 243–259.
- Gibbs Jr., Raymond W., Nandini P. Nayak & Cooper Cutting. 1989. How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of Memory and Language* 28(5). 576–593.
- Gibbs, Raymond W. 1980. Spilling the beans on understanding and memory for idioms in conversation. *Memory & Cognition* 8(2). 149–156.
- Gibbs, Raymond W. Jr., Nandini P. Nayak, John L. Bolton & Melissa E. Keppel. 1989. Speakers' assumptions about the lexical flexibility of idioms. *Memory & Cognition* 17(1). 58–68.
- Glasbey, Sheila. 2007. Aspectual composition in idioms. In Saussure, Louis de and Moeschler, Jacques and Puskás, Genoveva (ed.), *Recent advances in the syntax and semantics of tense, aspect and modality*, 71–88. (Trends in Linguistics. Studies and Monographs [TiLSM] 185). Berlin: De Gruyter Mouton.
- Glasbey, Sheila R. 2003. Let's paint the town red for a few hours: Composition of aspect in idioms. *LexFig '03: Proceedings of the ACL 2003 Workshop on Lexicon and Figurative Language, Volume 14*, 43–49. Stroudsburg, PA: Association for Computational Linguistics.
- Gough, Philip B. 1965. Grammatical transformations and speed of understanding. *Journal of Verbal Learning and Verbal Behavior* 4(2). 107–111.
- Gough, Philip B. 1966. The verification of sentences: The effects of delay of evidence and sentence length. *Journal of Verbal Learning and Verbal Behavior* 5(5). 492–496.
- Greenbaum, Sidney. 1976. Syntactic frequency and acceptability. *Lingua* 40. 99–113.
- Greenbaum, Sidney. 1977. Judgments of syntactic acceptability and frequency. *Studia Linguistica* 31(2). 83–105.
- Hale, Kenneth & Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Hale, Kenneth and Keyser, Samuel Jay (ed.), *The View from Building 20*, 53–109. Cambridge, MA: MIT Press.



- Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Hale, Kenneth and Keyser, Samuel Jay (ed.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111–176. Cambridge, MA: MIT Press.
- Halle, Morris & Alec Marantz. 1994. Some key features of Distributed Morphology. *MIT Working Papers in Linguistics: Papers on Phonology and Morphology* 21. 275–288.
- Hamblin, Jennifer L. & Raymond W. Gibbs Jr. 1999. Why you can't kick the bucket as you slowly die: Verbs in idiom comprehension. *Journal of Psycholinguistic Research* 28(1). 25–39.
- Harley, Heidi. 2003. Possession and the double object construction. *Linguistic Variation Yearbook* 2(1). 31–70.
- Harley, Heidi. 2007. The bipartite structure of verbs cross-linguistically, or, Why Mary can't exhibit John her paintings. In Silva, Thaïs Cristófaró and Mello, Heliana (ed.), *Conferências do V Congresso Internacional da Associação Brasileira de Linguística*, 45–84. Belo Horizonte, Brazil: ABRALIN and FALE/UFMG.
- Harley, Heidi. 2008. The morphology of nominalizations and the syntax of vP. In Giannakidou, Anastasia and Rathert, Monika (ed.), *Quantification, definiteness, and nominalization*, 320–342. Oxford: Oxford University Press.
- Harley, Heidi. 2011. Compounding in Distributed Morphology. In Lieber, Rochelle and Štekauer, Pavol (ed.), *The Oxford handbook of compounding*, 129–144. Oxford: Oxford University Press.
- Harley, Heidi. 2013. External arguments and the Mirror Principle: On the distinctness of Voice and v. *Lingua* 125. 34–57.
- Harley, Heidi. 2014a. On the identity of roots. *Theoretical Linguistics* 40(3-4). 225–276.
- Harley, Heidi. 2014b. Reply to commentaries, “On the identity of roots.” *Theoretical Linguistics* 40(3-4). 447–474.
- Harley, Heidi Britton. 1995. *Subjects, events and licensing*. Cambridge, MA: MIT dissertation.
- Harley, Heidi & Megan Schildmier Stone. 2013. The “No Agent Idioms” hypothesis. In Folli, Raffaella and Sevdali, Christina and Truswell, Robert (ed.), *Syntax and its Limits*, 251–275. (Oxford Studies in Theoretical Linguistics 48). Oxford: Oxford University Press.

- Hartigan, J. A. & P. M. Hartigan. 1985. The dip test of unimodality. *The Annals of Statistics* 13(1). 70–84.
- Harwood, William. 2013. *Being progressive is just a phase: Dividing the functional hierarchy*. Belgium: Ghent University dissertation.
- Healey, Alan. 1968. English idioms. *Kivung* 1(2). 71–108.
- Higgins, F. R. 1974. On the use of idioms as evidence for movement: A cautionary note. Paper presented at the Linguistics Society of America Annual Meeting, New York, December 1974.
- Higgins, F. R. 1981. On what are here called proxy clauses. Unpublished manuscript.
- Hockett, Charles F. 1958. *A course in modern linguistics*. New York: Macmillan.
- Horvath, Julia & Tal Siloni. 2002. Against the little-v hypothesis. *Rivista di Grammatica Generativa* 27. 107–122.
- Huddleston, Rodney. 1984. *Introduction to the Grammar of English*. Cambridge: Cambridge University Press.
- Idioms and Phrases - TheFreeDictionary. <http://idioms.thefreedictionary.com/> Accessed on 28 November 2016.
- Jackendoff, Ray. 1997. *The architecture of the language faculty*. Cambridge, MA: MIT Press.
- Just, Marcel Adam, Patricia A. Carpenter & Jacqueline D. Woolley. 1982. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General* 111(2). 228–238.
- Katz, J. J. & P. M. Postal. 1963. Semantic interpretation of idioms and sentences containing them. *Quarterly Progress Report No. 70*, 275–282. Cambridge, MA: MIT Research Laboratory of Electronics.
- Katz, Jerrold J. 1973. Compositionality, idiomaticity, and lexical substitution. In Halle, Morris and Anderson, Stephen R. and Kiparsky, Paul (ed.), *A Festschrift for Morris Halle*, 357–376. New York: Holt, Rinehart and Winston.
- Kay, Paul & Charles J. Fillmore. 1999. Grammatical constructions and linguistic generalizations: The What's X doing Y? construction. *Language* 75(1). 1–33.

- Koizumi, Masatoshi. 1993. Object agreement phrases and the split VP hypothesis. In Bobaljik, Jonathan and Philips, Colin (ed.), *Papers on Case and Agreement I*, 99–148. Cambridge, MA: MIT Press.
- Koopman, Hilda & Dominique Sportiche. 1991. The position of subjects. *Lingua* 85(2). 211–258.
- Kounios, John, Sonja A. Kotz & Phillip J. Holcomb. 2000. On the locus of the semantic satiation effect: Evidence from event-related brain potentials. *Memory & cognition* 28(8). 1366–1377.
- Kratzer, Angelika. 1996. Severing the external argument from its verb. In Rooryck, Johan and Zaring, Laurie (ed.), *Phrase structure and the lexicon*, 109–137. (Studies in Natural Language and Linguistic Theory). Dordrecht: Kluwer.
- Krifka, Manfred. 1998. The origins of telicity. In Rothstein, Susan (ed.), *Events and Grammar*, 197–235. Dordrecht: Kluwer.
- Lasnik, Howard & Robert Fiengo. 1974. Complement object deletion. *Linguistic Inquiry* 5(4). 535–571.
- Lin, Yuri, Jean-Baptiste Michel, Erez Lieberman Aiden, Jon Orwant, Will Brockman & Slav Petrov. 2012. Syntactic annotations for the Google Books Ngram Corpus. *Proceedings of the ACL 2012 System Demonstrations*, 169–174. Jeju, Republic of Korea: Association for Computational Linguistics.
- Makkai, Adam. 1966. *Idiom structure in English*. New Haven, CT: Yale dissertation.
- Makkai, Adam. 1972. *Idiom structure in English*. The Hague: Mouton.
- Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge, MA: MIT Press.
- Marantz, Alec. 1995. “Cat” as a phrasal idiom: Consequences of late insertion in Distributed Morphology. Unpublished manuscript, MIT.
- Marantz, Alec. 1997. No escape from syntax: Don’t try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics* 4(2). 201–225.
- Marantz, Alec. 2001. Words. Paper presented at the 20th West Coast Conference on Formal Linguistics, University of Southern California, Los Angeles, 23–25 February 2001.

- Marantz, Alec. 2013. Locality domains for contextual allomorphy across the interfaces. In Matushansky, Ora and Marantz, Alec (ed.), *Distributed Morphology Today: Morphemes for Morris Halle*, 95–115. Cambridge, MA: MIT Press.
- Massam, Diane. 2001. Pseudo noun incorporation in Niuean. *Natural Language & Linguistic Theory* 19(1). 153–197.
- McCawley, James D. 1981. The syntax and semantics of English relative clauses. *Lingua* 53(2). 99–149.
- McGinnis, Martha. 2002. On the systematic aspect of idioms. *Linguistic Inquiry* 33(4). 665–672.
- McGlone, Matthew S., Sam Glucksberg & Cristina Cacciari. 1994. Semantic productivity and idiom comprehension. *Discourse Processes* 17(2). 167–190.
- McMahon, Lee Edward. 1963. *Grammatical analysis as part of understanding a sentence*. Cambridge, MA: Harvard University dissertation.
- Mehler, Jacques. 1963. Some effects of grammatical transformations on the recall of English sentences. *Journal of Verbal Learning and Verbal Behavior* 2(4). 346–351.
- Mel'čuk, Igor. 1960. O terminax “ustojčivost” i “idiomatičnost” [About the terms “fixedness” and “idiomaticity”]. *Voprosy jazykoznanija* 4. 73–80.
- Merchant, Jason. 2013. Voice and ellipsis. *Linguistic Inquiry* 44(1). 77–108.
- Michaelis, Laura A. & Josef Ruppenhofer. 2001. *Beyond alternations: A constructional model of the German applicative pattern*. Stanford: CSLI Publications.
- Mitchell, Terry F. 1971. Linguistic “goings on”: Collocations and other lexical matters arising on the syntagmatic record. *Archivum Linguisticum* 2(1). 35–69.
- Moon, Rosamund. 1998. *Fixed expressions and idioms in English: A corpus-based approach*. Oxford: Oxford University Press.
- Newmeyer, Frederick J. 1974. The regularity of idiom behavior. *Lingua* 34(4). 327–342.
- Nunberg, Geoffrey D. 1978. *The pragmatics of reference*. New York, NY: City University of New York dissertation.
- Nunberg, Geoffrey, Ivan A. Sag & Thomas Wasow. 1994. Idioms. *Language* 70(3). 491–538.

- O'Grady, William. 1998. The syntax of idioms. *Natural Language & Linguistic Theory* 16(2). 279–312.
- Ortony, Andrew, Diane L. Schallert, Ralph E. Reynolds & Stephen J. Antos. 1978. Interpreting metaphors and idioms: Some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior* 17(4). 465–477.
- Palmer, Frank R. 1981. *Semantics*. Cambridge: Cambridge University Press.
- Perlmutter, David. 1970. The two verbs *begin*. In Jacobs, Roderick A. and Rosenbaum, Peter S. (ed.), *Readings in English transformational grammar*, 107–119. Waltham, MA: Ginn and Company.
- Peterson, Robert R., Curt Burgess, Gary S. Dell & Kathleen M. Eberhard. 2001. Dissociation between syntactic and semantic processing during idiom comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27(5). 1223–1237.
- Pollard, Carl & Ivan A. Sag. 1994. *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.
- Postal, Paul. 1974. *On Raising: One rule of English grammar and its implications*. Cambridge, MA: MIT Press.
- Pulman, Stephen G. 1993. The recognition and interpretation of idioms. In Cacciari, Cristina and Tabossi, Patrizia (ed.), *Idioms: Processing, structure, and interpretation*, 249–270. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Punske, Jeffrey. 2010. On nominal arguments. In Hogue, Alan and Schertz, Jessamyn (ed.), *Volume 17: Proceedings of the Arizona Linguistics Circle* 3.
- Punske, Jeffrey. 2012. Aspects of the internal structure of nominalization: Roots, morphology and derivation. Tucson, AZ: The University of Arizona dissertation.
- Punske, Jeffrey & Megan Schildmier Stone. 2014. Idiomatic expressions, passivization, and gerundization. Paper presented at the Linguistics Society of America 2014 Annual Meeting, Minneapolis, MN, 3 January 2014.
- Pylkkänen, Liina. 2002. *Introducing arguments*. Cambridge, MA: MIT dissertation.
- Pylkkänen, Liina. 2008. *Introducing arguments*. Cambridge, MA: MIT press.

- Ramchand, Gillian Catriona. 2008. *Verb meaning and the lexicon: A first phase syntax*. Cambridge: Cambridge University Press.
- Reagan, Robert Timothy. 1987. The syntax of English idioms: Can the dog be put on? *Journal of Psycholinguistic Research* 16(5). 417–441.
- Ruhl, Charles. 1975. Kick the bucket is not an idiom. *Interfaces: Linguistics and Psychoanalysis* 4. 2–4.
- Ruwet, Nicolas. 1991. *Syntax and human experience*. Chicago: University of Chicago Press.
- Sag, Ivan A. 2012. Sign-based construction grammar: An informal synopsis. In Boas, Hans C. and Sag, Ivan A. (ed.), *Sign-Based Construction Grammar*, 69–202. Chicago: CSLI.
- Sag, Ivan A., Thomas Wasow & Emily M. Bender. 2003. *Syntactic theory: A formal introduction*. Stanford, CA: Center for the Study of Language and Information.
- Salton, Giancarlo D., Robert J. Ross & John D. Kelleher. 2016. Idiom token classification using sentential distributed semantics. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, 194–204. Berlin: Association for Computational Linguistics.
- Sauerland, Uli. 2003. Unpronounced heads in relative clauses. In Schwabe, Kerstin and Winkler, Susanne (ed.), *The interfaces: Deriving and interpreting omitted structures*, 205–226. Amsterdam: John Benjamins.
- Saussure, Ferdinand de. 1986. *Course in General Linguistics*. Bally, Charles and Sechehaye, Albert and Riedlinger, Albert (ed.). Harris, Roy (trans.). Chicago: Open Court.
- Schachter, Paul. 1973. Focus and Relativization. *Language* 49(1). 19–46.
- Schütze, Carson T. 1996. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.
- Slobin, Dan I. 1966. Grammatical transformations and sentence comprehension in childhood and adulthood. *Journal of Verbal Learning and Verbal Behavior* 5(3). 219–227.
- Snyder, William. 2000. An experimental investigation of syntactic satiation effects. *Linguistic Inquiry* 31(3). 575–582.

- Speyer, Augustin. 2005. Topicalization in English and the Trochaic Requirement. *University of Pennsylvania Working Papers in Linguistics* 10(2). 243–256.
- Sprouse, Jon. 2007. *A program for experimental syntax: Finding the relationship between acceptability and grammatical knowledge*. College Park, MD: University of Maryland dissertation.
- Sprouse, Jon. 2008. Magnitude estimation and the non-linearity of acceptability judgments. In Abner, Natasha and Bishop, Jason (ed.), *Proceedings of the 27th West Coast Conference on Formal Linguistics*, 397–403. Somerville, MA: Cascadilla Proceedings Project.
- Stone, Megan Schildmier. To appear. Systematic flexibility in verb-object idioms. In Corver, Norbert and van Craenenbroeck, Jeroen and Harwood, William and Hladnik, Marko and Leufkens, Sterre and Temmerman, Tanja (ed.), *The Grammar of Idioms: Proceedings of the 8th Brussels Conference on Generative Linguistics*. Berlin, Boston: DeGruyter Mouton.
- Stone, Megan Schildmier. 2009. Idioms and domains of interpretation. Paper presented at the Arizona Linguistics Circle 3, University of Arizona, Tucson, 1 November 2009.
- Stone, Megan Schildmier. 2013. A corpus-based approach to syntactic flexibility in verb-object idioms. Paper presented at the Arizona Linguistics Circle 7, University of Arizona, Tucson, 12 October 2013.
- Stone, Megan Schildmier. 2014. On the flexibility of verb-object idioms. Paper presented at the Linguistics Society of America 2014 Annual Meeting, Minneapolis, MN, 4 January 2014.
- Svenonius, Peter. 2005. Extending the extension condition to discontinuous idioms. *Linguistic Variation Yearbook* 5(1). 227–263.
- Swinney, David A. & Anne Cutler. 1979. The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior* 18(5). 523–534.
- Tabossi, Patrizia, Rachele Fanari & Kinou Wolf. 2008. Processing idiomatic expressions: Effects of semantic compositionality. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34(2). 313–327.
- Thibadeau, Robert, Marcel Adam Just & Patricia A. Carpenter. 1982. A model of the time course and content of reading. *Cognitive Science* 6(2). 157–203.

- Titone, Debra A. & Cynthia M. Connine. 1994a. Comprehension of idiomatic expressions: Effects of predictability and literality. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20(5). 1126–1138.
- Titone, Debra A. & Cynthia M. Connine. 1994b. Descriptive norms for 171 idiomatic expressions: Familiarity, compositionality, predictability, and literality. *Metaphor and Symbolic Activity* 9(4). 247–270.
- Townsend, David J. & Thomas G. Bever. 2001. *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.
- Travis, Lisa. 2000. Event structure in syntax. In Tenny, Carol and Pustejovsky, James (ed.), *Events as grammatical objects: The converging perspectives of lexical semantics and syntax*, 145–185. Stanford: CSLI Publications.
- Tukey, John W. 1977. *Exploratory data analysis*. Massachusetts: Addison-Wesley.
- Vendler, Zeno. 1957. Verbs and times. *The Philosophical Review* 66(2). 143–160.
- Vergnaud, Jean-Roger. 1974. *French relative clauses*. Cambridge, MA: MIT dissertation.
- Vergnaud, Jean-Roger. 1985. *Dépendances et niveaux de représentation en syntaxe*. Amsterdam: John Benjamins.
- Vincenzi, Marica De, Remo Job, Rosalia Di Matteo, Alessandro Angrilli, Barbara Penolazzi, Laura Ciccarelli & Francesco Vespignani. 2003. Differences in the perception and time course of syntactic and semantic violations. *Brain and Language* 85(2). 280–296.
- Wasow, Thomas, Ivan A. Sag & Geoffrey Nunberg. 1984. Idioms: An interim report. In Hattori, Shiro and Inoue, Kazuko (ed.), *Proceedings of the XIIIth International Congress of Linguistics*, 87–96. Tokyo: Nippon Toshi Center.
- Wearing, Catherine. 2012. Metaphor, Idiom, and Pretense. *Noûs* 46(3). 499–524.
- Weinreich, Uriel. 1969. Problems in the analysis of idioms: Lectures delivered before the Linguistic Institute of the Linguistic Society of America, University of Los Angeles, June 17 - August 12, 1966. In Jaan Puhvel (ed.), *Substance and structure of language*, 23–81. Berkeley/Los Angeles: University of California Press.



- Weskott, Thomas & Gisbert Fanselow. 2008. Variance and informativity in different measures of linguistic acceptability. In Abner, Natasha and Bishop, Jason (ed.), *Proceedings of the 27th West Coast Conference on Formal Linguistics (WCCFL)*, 431–439. Somerville, MA: Cascadilla Press.
- Wittenberg, Eva & Maria Mercedes Piñango. 2011. Processing light verb constructions. *The Mental Lexicon* 6(3). 393–413.
- Wood, Mary McGee. 1986. *A definition of idiom*. Bloomington, IN: Indiana University Linguistics Clubs.