COMPOUND CONCEPTUAL RELATIONS IN WORKING MEMORY

COMPOUND CONCEPTUAL RELATIONS IN WORKING MEMORY:
EFFECTS OF RELATION PRIMING IN IMMEDIATE SERIAL RECALL

BY

MICHAEL GREENCORN, B.A. (Hons)

A Thesis Submitted to the Department of Linguistics & Languages
and the School of Graduate Studies in Partial Fulfilment of the Requirements for the
Degree Master of Science

McMaster University

Master of Science (2017)
Linguistics & Languages
Cognitive Science of Language

McMaster University
Hamilton, Ontario
Canada

TITLE:	Compound Conceptual Relations in Working Memory: Effects of Relation Priming in Immediate Serial Recall
AUTHOR:	Michael Greencorn B.A. (Hons), Saint Mary's University
SUPERVISOR:	Dr. Elisabet Service
NUMBER OF PAGES:	*viii*, 77

Abstract

The conceptual relation theory postulates that English noun-noun compound words (e.g., snowman) have an underlying predicate structure that is not present in the surface form, but is recovered during compound processing (e.g., man made of snow). The relational nature of constituent binding in compound words marks them as a linguistic construction that is distinct from both the simplex words (monomorphemic) and other complex words (derived and inflected words) previously examined in the context of verbal working memory. In short-term memory research, a growing body of evidence suggests that semantic properties of words influence verbal recall; however, such effects have not been examined in the context of compound conceptual relations. The present study investigated the possible effects of compound conceptual relations in verbal working memory via an immediate serial recall task. The task was designed to examine whether sharing of an individual relation leads to facilitative or inhibitory effects for compounds associated with that relation and, more generally, whether this semantic property of compound words contributes to their recollection from short-term memory. Evidence from the serial recall experiment suggested an effect of compound relation priming in working memory. Relational similarity between recall list items appeared to inhibit recall performance. The thesis discusses how this may be the result of increased competition between compound constituents as a result of heightened constituent-level activation during word recall. This effect was not observed in relations that appeared to be overly general, suggesting that the effect is only present when compound words are matched according to salient, sufficiently specified relations.

Acknowledgements

The journey of writing a thesis is filled with contributions from a great many people. I would like to begin by thanking my supervisor, Dr. Elisabet Service, for her constant support, encouragement, and guidance throughout this process. Without our long, meandering conversations through the world of mind and memory this study would not have become what it is, and the journey would not have been half as fun. I am forever grateful to her for encouraging my curiosity, and for navigating me through my own self-doubts. I also extend heartfelt thanks to my thesis committee members, Dr. Victor Kuperman and Dr. Catherine Anderson, for being part of the thesis process, as well as my overall experience as a graduate student. I am lucky to have worked with them both.

Special thanks also must be made to Dr. Daniel Schmidtke who volunteered his time and expertise in the early stages of this study to help me find my footing. I am also extremely grateful to Victoria Stevenson of the Language, Memory and Brain (LMB) lab for assisting me in the running of my study. Without her help, I might still be endlessly travelling between home and the lab. I would also like to thank the other LMB lab members for providing me with valuable feedback and support.

I am very glad to have studied, worked, and laughed with many wonderful people: Richard, Kyle, Connie, Sara, Cassandra, Rober, Jitka, Gaisha, Kenny, Natalia, and everyone else who makes up this unique department. Special mentions go out to Zoë and Narcisse for the long talks about life beyond grad school; Fareeha for her constant support and commiseration; Laura for the laughs, talks, and delicious treats; Heather for her support and much-needed dinner escapes; and Chelsea for being my long-lost twin who managed to understand me even when I was too exhausted to speak—you all made every day better.

Perhaps the largest thanks goes to David, who has accompanied me through the ups and downs of the last two years and, in particular, the final gauntlet that was writing this thesis. Without his support, encouragement and late-night snack provisioning, I may not have made it through. Thank you for putting up with a grumpy old man muttering nonsense about snowmen and teapots.

Finally, I would like to dedicate this thesis to my parents who have always supported me throughout my life's occasionally surprising adventures. From a tiny fishing village on the coast of Nova Scotia to the other side of the world and back again,

they have always believed in me—even if they have not always understood what it is that I do. Thank you for everything that I am. This one’s for you.

Contents

List of Figures

List of Tables

Introduction

Short-term memory is often considered to be critically involved in—and have implications for—a variety of cognitive tasks, including language comprehension, mathematics, and visuospatial ordering (Walker & Hulme, 1999; Baddeley 1986; 1983; Hitch, 1980). As a result, it has been the subject of a vast number of studies over the last five decades. Studies of verbal short-term memory typically involve immediate serial recall tasks that ask participants to memorize a list of words then recall them in serial order immediately after presentation. Through the manipulation of the linguistic stimuli used in these tasks, researchers have revealed effects of various lexical and semantic properties of words on word recall performance. These observances have provided insight into the short- term processing, storage, and retrieval of linguistic information, and have helped to shape our understanding of how words are represented in the mind in both the short- and long-term. Theories regarding the role of phonological information in short-term word storage and retrieval have been prominent throughout the literature (e.g., Baddeley, 1983; Baddeley, 1986; Cowan et al., 1992); however, emerging evidence shows that semantic information (typically associated with long-term memory) also contributes to our ability to recall linguistic information (Poirier et al., 2015; Acheson et al., 2011; Poirier & Saint-Aubin, 1995; Huttenlocher & Newcombe, 1976). These findings have led to theories proposing that long-term memory plays either a central (e.g., Acheson & MacDonald, 2009; Roodenrys, 2009) or supporting role (e.g., Baddeley, 2000; 2012) in short-term memory operations on linguistic representations.

Studies of verbal short-term memory have largely focused on morphologically simplex (monomorphemic) words (e.g., *tea*, *sell*). Morphologically complex words are words composed of two or more meaningful linguistic units (morphemes), formed by the combination of a root word and some form of affix (e.g., *sell* + *ing*) or by combining two or more words together in a compound (e.g., *teacup*). The few existing studies of complex words in verbal short-term memory suggest that morphological complexity limits word recall, but that the extra load may be partially offset by support from increased semantic representation (Service & Maury, 2015; Németh et. al., 2011; Service & Tujulin, 2002). In the case of compound words, less is known about their representation in short-term memory and the effect of lexical or semantic properties on their recollection. A study by Wälchli (2016), reported that immediate serial recall of compound words appeared to rely heavily on the phonological encoding of constituents. The contribution of their semantic features was not investigated.

The relational nature of constituent binding in compound words marks them as a linguistic construction that is distinct from both the simplex words (monomorphemic) and other complex words (derived and inflected words) previously examined in the context of verbal working memory. For example, the combination of the noun *man* with a modifier noun *snow* produces a compounded concept, *snowman*, that is more than the simultaneous activation of two concepts. The conceptual relation theory (e.g., Downing, 1977; Levi, 1978) postulates that the creation of compound words (e.g., *snowman*) not only involves the concatenation of two words, but also requires the selection of a relationship that explains and rationalizes the connection between them. Under this theory, the combined concept *snowman* uses a MADE OF relationship to specify that the head "man" is MADE OF the modifier "snow". Empirical studies on compound conceptual relations (e.g., Gagné, 2001; Gagné, 2002; Estes & Jones, 2006; Popov & Hristova, 2015) have not only shown evidence for their psychological reality, but have also suggested that the interpretation of a compound can be facilitated by prior exposure to a different compound containing the same conceptual relation. In sensicality judgement tasks requiring a sensical/non-sensical response (e.g., Gagné, 2002; Estes & Jones, 2006), comprehension of a compound such as *snowman* (man MADE OF snow) is faster when preceded by a compound like *meatball* (ball MADE OF meat) than when preceded by a compound such as *bookshelf* (shelf FOR books) in a phenomenon dubbed "relation priming".

The study presented in this thesis investigated the possible effects of compound conceptual relations in verbal working memory via an immediate serial recall task. The task was designed to investigate whether repetition of an individual relation leads to facilitative priming for compounds associated with that relation and, more generally, whether this semantic property of compound words contributes to their recollection from short-term memory. I will first review the relevant background information across three chapters summarizing the literature on the topics of English compound words, compound conceptual relations, and words in verbal short-term memory. Two empirical studies are then presented. In the Immediate Serial Recall Experiment, participants were required to memorize lists of compound words and recall them in order immediately after presentation. Lists differed in whether the presented compounds used the same or different conceptual relations to probe the effects of conceptual relation similarity on compound word recall. In the Conceptual Relation Similarity Survey, participants were asked to compare compound words used in the recall task and rate the similarity of their conceptual relations. The similarity ratings from the survey were used in post-hoc analysis to validate the experimental design of the serial recall task. The findings are discussed, addressing the possible involvement of conceptual relations in how compound words are processed, stored, and retrieved from short-term memory.

Chapter 1

English Compound Words

Compound words such as *teacup*, *winter jacket*, and *frying pan* are common in English and other languages. Compounds typically serve the purpose of helping us distinguish between different types of things, or different manners of performing actions. For example, there are various types of *cup* that one may wish to purchase for different purposes: *teacups*, *coffee cups*, *measuring cups*, *egg cups*, etc. Similarly, one may wish to describe the type or manner of *walking* while unconscious using the term *sleepwalk*. The basic structure of such compound words is that they consist of a head word that provides the core meaning, and a modifier word that classifies the type or manner of the head. For example, *teacup* contains a head noun *cup*, which tells us that the thing in question is a cup, and a modifier *tea*, which indicates that it belongs to a class of cups used for tea. In English, a compound's syntactic category (noun, verb, etc.) is determined by its head; thus, a *teacup* is a noun based on its head *cup*, while *(to) sleepwalk* is a verb based on its head *walk*. This is illustrated in Table 1 below. The present study focuses on the structure of *nominal compounds*: those whose head word is a noun.

Table 1: components of English compound words

Compound	Modifier	Head	Syntactic Category
teacup	tea	cup	noun
(to) sleepwalk	sleep	walk	verb

Properties and Characteristics of Compound Words

The following sub-sections will briefly introduce some basic properties of compound words that are relevant to further discussion of their structure and the way readers comprehend them. Because these properties influence the way compound words are processed, they should be considered in any study investigating compound processing, including the study presented in this thesis.

Headedness

A prototypical compound word contains a head and a modifier. Headedness refers to the constituent (left or right; first or second) that serves as the semantic and syntactic head of the compound. In English, compound words are generally right headed, meaning the head is the second constituent (e.g., tea*cup*, sleep*walk*). Head position varies cross-linguistically, and languages may exhibit preference for left or right headedness, or even no apparent preference at all (Williams, 1981). In transparent and semi-transparent compounds—those with fully discernable meanings, such as *teacup*, or partially discernable meanings, such as *nickname*—the head plays an important role in defining the syntactic category of the compound (e.g., noun, verb), as well as providing the core meaning of the compound. Headedness in opaque compounds—those with meanings that are not easily understood, such as *humbug*—is less discernable. Humbug is neither a bug nor a type of hum, so it does not appear to receive semantic information from either constituent. Its syntactic category (noun), however, may stem from the word *bug*, suggesting that the second constituent may act as a syntactic head if not a semantic one. The presence of a head—or primary contributor to semantic and syntactic information—in opaque compounds is debatable; opaque compounds like *humbug* are likely to have had an originally transparent or semi-transparent meaning that has simply been lost over time, resulting in the loss of an obvious semantic or syntactic head (Libben et. al, 2003).

It has been shown that headedness has an influence in a variety of lexical and semantic properties. Libben et al. (2003) found that compounds with transparent heads (e.g., *teacup*, *nickname*) exhibited faster response times in lexical decision tasks than those with opaque heads (e.g., *jailbird*, *humbug*), and proposed that the transparency of the compound head plays a more significant role in processing than the transparency of the modifier. In a further study of Italian compound words, Arcara, Semenza, and Bambini (2014) compared the processing of head-initial and head-final compounds in Italian and found that there is a higher processing cost for head-final compounds. The authors propose that this is the result of a reanalysis: because the core semantic content occurs at the end of the word, the reader needs to reanalyze the whole compound to integrate the information gained from the head. In contrast, they argue that receiving the core semantic information first (i.e., in head initial compounds) does not require the same reanalysis.

All compounds in the present study are right-headed, thereby eliminating any effects of left-right head position; however, the finding that head transparency modulates compound processing times is an effect that that may carry into memory recall tasks. If this effect patterns similarly in memory recall, we may expect to find increased recall performance for compounds with higher head transparency ratings. This is discussed further under the heading *Semantic Transparency* (page 6).

Frequency

Word frequency describes how often a word occurs in a corpus of (usually written) language as an extrapolation of how frequent that word is in the language. Monomorphemic words (i.e., words made up of a single morpheme, such as *dog* or *restaurant*) that occur more frequently are generally found to be easier to process (Forster & Chambers, 1973; Inhoff & Rayner, 1986; Grainger, 1990) and to have more synonyms and word associations in a rich semantic network (Balota et al., 2004). And while these effects of word frequency are well documented in monomorphemic words, studies on the frequency effects of multimorphemic words—in particular, compound words—have painted a more complex picture (see Amenta & Crepaldi, 2012; Duñabeitia, Perea, & Carreiras, 2007 for reviews).

In the case of compound words, frequency measures can be obtained for the whole word (also known as the *surface frequency*, e.g., the frequency of the word *teacup*) and for each of its individual constituents (i.e., *tea* and *cup*). Evidence from a variety of empirical studies suggests that whole-word and constituent frequencies contribute to compound processing separately, including observed differences between the contributions of left and right constituent frequencies. Typically, compounds with high whole-word frequencies are processed faster than compounds with low whole-word frequencies, in a similar manner to monomorphemic words (Bertram & Hyönä, 2003; Amenta & Crepaldi, 2012). In an eye-tracking study mimicking natural reading, Hyönä and Pollatsek (1998) found that the frequency of the first constituent of a Finnish noun-noun compound influenced both a reader's initial fixation and their total gaze duration of the whole word (higher frequency predicted shorter fixations). The authors concluded that having a high frequency first constituent reduced the reading time of compound words. In a further study focusing on the second constituent, Pollatsek et al. (2000) found a similar, significant facilitative effect of second constituent frequency. This finding was replicated by Juhasz et al. (2003) who also found a facilitative effect of second constituent frequency; however, this study showed only a trend toward facilitation for first constituents. In an additional eye-tracking study, Andrews, Miller, and Rayner (2004) found differing frequency effects for both constituents. In this study, a higher frequency first constituent reduced readers' initial fixation duration and the total gaze duration for the whole compound word (as per Hyönä & Pollatsek, 1998), while a higher frequency second constituent reduced the initial fixation duration and the probability that a reader would have to regress for reanalysis. This finding suggests that while the frequencies of both constituents influence reading times (with higher frequencies equating to faster reading speeds), they contribute to the reading process in different ways.

Morphological Family Size and Family Frequency

Morphological family size refers to the number of words that a morphological constituent appears in (e.g., the morphological family size for the word *snow* is the count of all words that contain the morpheme *snow: snowed*, *snowing*, *snowball*, etc.). Studies of morphological family size have shown a consistent effect whereby words with more morphological relatives (i.e., a larger family size) are processed faster than words with a smaller family network, such that words that have frequent associations are processed more quickly. (Schreuder, 1997; Pylkkänen et al., 2004; Juhasz & Berkowitz, 2011). A similar effect is that of *morphological family frequency*, which is the sum of frequencies for the morphological family. Words with higher family frequencies are also processed faster than words with lower family frequencies (Baayen, Dijkstra, & Schreuder, 1997; Schreuder & Baayen, 1997). In noun-noun compound words, morphological family size and frequency are computed for left and right constituents independently (e.g., size and frequency of both *snow* and *man* in *snowman*). Positional family size and frequency are also calculated (e.g., the family size and frequency of *snow* as a modifier and as a head). Family size and frequency effects are typically attributed to the increased activation of relatives in a morphological network: words that belong to a larger morphological family (or a family with higher frequency) benefit from spreading activation of associated words in the mental lexicon which allows for faster processing and recognition times (de Jong, Schreuder, & Baayen, 2000).

Semantic Transparency

Noun-noun compound words are frequently discussed in terms of their semantic transparency: that is, how easily their meaning can be discerned from their constituent parts. Highly transparent or *endocentric* compounds such as *teacup* are easily understood through the analysis of their parts alone. Conversely, opaque or *exocentric* compounds such as *hogwash* have a seemingly idiosyncratic meaning that cannot be determined through analysis of their parts alone (i.e., *hogwash* has nothing to do with washing or hogs) and comprehension of their meaning must rely on some other knowledge. Moreover, the meaning contribution of individual constituents may have differing levels of transparency, thus allowing compounds to be classified into four transparency categories (Libben, 2003):

1. those with a transparent modifier and head (TT) (e.g., *teacup*)
2. those with a transparent modifier and opaque head (TO) (e.g., *jailbird*)
3. those with an opaque modifier and transparent head (OT) (e.g., *nickname*)
4. those with an opaque modifier and head (OO) (e.g., *hogwash*).

Semantic transparency is generally considered an important factor in compound word processing (Zwitserlood, 1994; Libben, 1998; Libben et al., 2003; El-Bialy, Gagné, & Spalding, 2013). One might assume, instinctually, that compound words that have meanings closely associated to—and easily derived from—their component parts would be easier to understand and, therefore, easier to process than compounds with opaque or partially opaque constituents. A review of the literature, however, shows that the effect of semantic transparency is still under debate, with disparate evidence emerging from different investigations (see Frisson, Niswander-Klement, &

Pollatsek, 2008, for a comprehensive review). Libben et al. (2003) found that the semantic transparency of a compound word is derived directly from the transparency of both of its constituents, although the transparency of the head constituent played the most significant role in lexical decision latencies. In this study, it was found that TT and OT compounds were recognized more quickly than TO and OO compounds, suggesting that the transparency of the head plays a more significant role in compound processing than the transparency of the modifier, and that having a transparent core meaning reduces processing cost. Conversely, Pollatsek and Hyönä (2005) found no evidence for semantic transparency effects in partially or fully opaque compounds (OT, TO, OO) in an eye-tracking study designed to investigate natural reading. This appears counter to the findings of Libben et al. (2003) in the case of OT compounds, and suggests that transparency is only a processing factor for fully transparent compounds. Moreover, Frisson et al. (2008) found no evidence of a transparency effect in English compounds during normal reading regardless of transparency composition (OT, TO, OO, TT) when compounds were written without a space (e.g., *mothball* vs *moth ball*). They argue that this suggests that unspaced compound words are processed as whole words and, therefore, constituent level transparency effects are not a factor; however, the authors note that the variation in findings across the literature may be evidence that transparency effects are partially task induced (2008, p.102), or that the eye-tracking measures utilized in their study were not sufficient to capture relevant effects.

In sum, the effects of semantic transparency in compound processing are not currently clear. Evidence suggests that there may be an effect of head transparency on compound visual processing, but that this effect may be both task-specific and limited to incidences of compounds presented as spaced words. In the present study, all compounds had relatively high transparency ratings and were presented as unspaced words. If the effects of word spacing found by Frisson et al. (2008) hold true in a memory recall design, then we would expect there to be no apparent effect of constituent transparency on recall performance. Evidence of an effect of constituent transparency may indicate that compound words are not stored as whole words in working memory, despite their orthographic presentation.

Section Summary

As highlighted previously, any study of compound words necessarily involves the consideration of a variety of interacting lexical and semantic factors that appear to interplay at both whole-word and constituent levels. These interactions are important for the understanding and discussion of how compound words are visually processed and represented within the mental lexicon how they are represented within working memory (Chapter 2) and in any analysis dependent on their processing (Chapter 4).

Processing and Representation of Morphologically Complex Words

Studies of morphologically complex words have largely focused on how such words are (1) processed, and (2) stored in the mind. Specifically, studies have largely focused on the question of whether morphologically complex words are processed as whole words (full-listing models), as a string of parts (morphological decomposition models), or a combination of the two (dual-route models) and whether they are similarly represented in the mind as whole words, individual components, or a combination thereof. An underlying assumption behind most theories of morphological decomposition is that there is a processing cost associated with decomposing and then reassembling constituent morphemes, thereby suggesting that whole-word access is more efficient, and that any combined route (i.e., involving both whole-word and constituent level processing and/or representations) would be slower or less efficient than whole-word access alone (Ji, Gagné, & Spalding, 2011). Experimental evidence suggests, however, that compound words are processed more quickly than monomorphemic words. Ji, Gagné, & Spalding (2011) found that both semantically transparent compounds (e.g., *rosebud*) and opaque compounds (e.g., *hogwash*) were processed more quickly than monomorphemic words matched in length and frequency. Similarly, Fiorentino and Poeppel (2007) found a processing advantage for compound words compared to monomorphemic words in lexical decision times and event-related potentials (ERP). Further, Juhasz (2006) found that compound words with a high frequency first constituent had shorter processing times than matched monomorphemic words in an eye-tracking task, while low frequency first constituent compounds did not differ in processing time from monomorphemic words. Together, the results of such studies suggest that the existence of multiple routes for processing—and, thereby, information access—via whole-words and constituents increases processing efficiency in multimorphemic words. This also suggests that compound words benefit from the access of whole-word and constituent level representations within the mental lexicon. Emerging accounts of compound word processing suggest that several sources of information are involved in a complex equation that considers constituent position, semantic transparency, and constituent family size (Libben, 1998; Libben, 2003; Kuperman, Bertram, & Baayen, 2008; Kuperman, Schreuder, Bertram, & Baayen, 2009).

Section Summary

The present study utilized the rapid serial visual presentation (RSVP) of compound words to participants: participants were required to read compound words as they were presented orthographically on a computer screen. Evidence from compound word processing studies indicates that both whole-word and constituent level representations play a role in compound word comprehension. This further suggests that the present examination of compound word representations and their interactions in working memory should consider both whole-word and constituent level factors.

Chapter 2

Compound Conceptual Relations and Meaning Derivation

That the constituents of compound words are joined by a relational structure appears to be evident to language users, as speakers can readily verbalize these relationships through paraphrase. For example, one might explain that a *teapot* is "a pot used FOR tea" or that a *snowball* is a "ball MADE OF snow". The fact that speakers define English compound words using relationships such as FOR and MADE OF suggests that these relationships are critical to the meaning of compounds, despite not appearing in the compounds themselves. This raises important questions regarding where such relational knowledge originates and how speakers integrate this knowledge with constituents to understand compound meaning. Psycholinguistic research on morphologically complex words has largely focused on derived and inflected words, while studies of compound words remain less common. In particular, theoretical and empirical studies of compound relationships are few and largely recent, but show important effects in compound word processing and the associations that are held between concepts (e.g., Boutonnet, McClain, & Thierry, 2014; Spalding & Gagné, 2011; Gagné, 2001; Gagné & Shoben, 1997; Levi, 1979; Downing, 1977.).

Ferdinand de Saussure's seminal work, *Course in General Linguistics* (*Cours de Linguistique Générale*, 1959), distinguishes two types of relationship that can exist between concepts: *syntagmatic* relations, those that exist only in context; and *paradigmatic* relations, those that are context independent and exist as the result of accumulated experience. For example, the concepts *apple* and *basket* have no intrinsic relationship, but when those concepts are combined to form *apple basket* (a basket for apples) they enter a context-specific relationship (i.e., a context wherein a basket is used for holding apples); however, the relation that exists between *apple* and *fruit* (an apple is a fruit) is constant regardless of context. This distinction suggests that the combination of concepts like *apple* and *basket*, or *snow* and *ball* to form compounded concepts requires the inclusion of a context-specific relationship that is otherwise not inherent between the two constituents. This relationship encodes information about the specific combination that goes beyond the static, denotational meaning of each constituent, and allows language users to understand the combination of concepts in context. While the interpretation of such relationships appears intuitive, this raises numerous questions: What are the possible relationships, and are they a finite class of structures? How are these relationships formed? How do speakers decide or know which relationships to use? Each of these questions reflects an ongoing investigation within the realm of compound word processing. In this chapter, each question is addressed briefly through a summary of the current literature.

Conceptual Relation Taxonomies

The goal of developing a taxonomy of conceptual relations is to describe an inventory of relations that is both general enough to apply to the infinite generativity of compounding, while also being specific enough to successfully encapsulate and convey the desired compound meaning. Despite the commonality of this goal, the number of relations presented in the various taxonomies for noun-noun compound words varies quite significantly (see Nastase et al. 2014 for review). Warren (1978) proposed 45 discrete compound relations derived from a study of the Brown Corpus (Kučera and Francis, 1967) that belong to 11 hierarchical categories, eventually collapsing into six relational "super categories". Similarly, Tratz and Hovy (2010) proposed an inventory of 43 relations within 10 high-level categorizations. In contrast, Levi (1978) devised a core list of just 12 relations to account for the semantic relations of complex nominals (excluding predicating nominals and nominalizations; see below), and Shoben (1991; Gagné & Shoben, 1997) expanded on Levi's inventory, delineating first 14 then 16 relations covering both predicating and nominalized compounds. The present study utilizes the compound conceptual relation categorizations defined by Shoben (1991; Gagné & Shoben, 1997) and modelled on Levi (1978). As Shoben's treatment does not diverge significantly from Levi, a summary of Levi's categorization is provided below.

In a seminal work on compound relations, Levi (1978) describes two classes of complex nominals that are formed through *adjective + noun* construction: those with *predicating adjectives* and those with *non-predicating adjectives*. Complex nominals with predicating adjectives contain an underlying predicating structure in the form of *Noun **is** Adjective*. For example, *electric clock* has an underlying predicating structure *clock **is** electric*. Complex nominals with non-predicating adjectives do not enter this copular structure, and thus cannot be expressed as *Noun **is** Adjective*. For example, *electrical engineer* logically does not represent an underlying structure *engineer **is** electrical*.. Levi demonstrates that noun-noun compound words (where the modifier is a noun rather than an adjective), pattern like nominals with non-predicating adjectives. For example, *teacup* does not represent an underlying structure *cup **is** tea*. Levi provides derivational analysis to show that non-predicating adjectives (i.e., the *electrical* type) are, in fact, underlyingly derived from nouns. In sum, predicating nominals are those that combine *adjective + noun* (*big bomb*), while non-predicating nominals are those that combine *noun + noun* (*atom bomb)*. The modifying noun of a non-predicating nominal may be derived into an adjectival form (*atomic bomb*), but remains non-predicating (**bomb is atomic*).

Throughout, Levi (1978) addresses the question of how the meanings of non-predicating nominals (NPNs) are structured and derived. She argues that instead of an underlying copular structure of *noun **is** modifier*, NPNs of the two types described above contain a deleted predicate from a finite list of "Recoverably Deletable Predicates" (RDPs). Rather than *cup* ***is*** *tea*, Levi argues that an NPN such as *teacup* contains the underlying form *cup* FOR *tea*, where the predicate FOR explains the relationship between the two constituents, but has been deleted in the surface form. In the course of processing an NPN like *teacup*, an individual is argued to recover the deleted predicate and reconstruct the full phrasal meaning. Levi identifies nine "primitive" RDPs that can be used to explain the semantic relationships between modifier and noun constituents: CAUSE, HAVE, MAKE, BE, USE, FOR, IN, ABOUT, and FROM. The CAUSE, HAVE, and MAKE relations are argued to have two related but distinct forms each: (1) Modifier CAUSE/HAVE/MAKE Head, and (2) Head CAUSE/HAVE/MAKE Modifier. This raises the total number of Levi's relations to 12. Although these relations are absent from the surface form, Levi argues that both the surface and underlying forms of complex nominals allow for the same semantic interpretations (i.e., the surface form is no more restrictive nor idiosyncratic than the underlying form, despite its simplicity), and that possible semantic ambiguities are resolved through the combination of sentential and situational context and pragmatics.

Levi provides a separate analysis for a final class of non-predicating nominals that that are formed through a verb nominalization process. This class of NPN is composed of a nominalized verb and a modifier that is either the subject of the underlying verb (e.g., *parental advisory* from 'parent advises'), or the object of the verb (e.g., *dream analysis* from 'analyse dream'). Levi proposes four types of nominalizations to account for the remaining semantic interpretations of NPNs, illustrated in Table 2 below.

Table 2: classes of verb-nominalization (adapted from Levi, 1978).

Nominalization Type	Example	
Act Nominalizations	*parental refusal*	act of parents refusing
Product Nominalizations	*human error*	that which is produced by (the act of) humans erring
Agent Nominalizations	*mail sorter*	x such that *x sorts mail*
Patient Nominalizations	*student invention*	y such that *students invent y*

The present study utilizes the 16 relations proposed by Shoben (1991; Gagné & Shoben 1997) which include derivatives of Levi's 12 RDPs and verb nominalizations. These are compiled with examples in Table 3 below.

Table 3: semantic relations, adapted from Shoben (1991; Gagné & Shoben 1997). H: head; M: modifier.

Relation	Example	Relation	Example
H ABOUT M	*newsflash*	M HAS H	*doorframe*

Relation	Example	Relation	Example
H BY M	*handclap*	H location is M	*farmyard*
H CAUSES M	*joyride*	H location is H	*neckline*
H CAUSED BY M	*sunbeam*	H MADE OF M	*snowman*
H DERIVED FROM M	*seafood*	H MAKES M	*flourmill*
H DURING M	*nightlife*	H IS M	*girlfriend*
H FOR M	*mealtime*	H USES M	*steamboat*
H HAS M	*bookshop*	H USED BY M	*witchcraft*

Further Notes on Relation Taxonomies

The number of relations required to accurately represent the seemingly infinite number of conceptual combinations is a matter of open debate, and many authors note that their own attempts to inventory relations are unlikely to be exhaustive (e.g., Gagné & Shoben, 1997; Downing, 1977). Downing (1977) argues that, in fact, compound relationships cannot be a finite class if an infinite number of novel interpretations can be generated. To illustrate this position, he provides the example *plate length* to describe "what your hair is when it drags in your food", arguing that the ability to generate an infinite number of such context-specific relations would require an infinite number of compound relations. Downing's position, however, appears to assume that a significant amount of idiosyncratic information is encoded in the relation itself, rather than being discerned through context. In contrast, the finite number of relations described by Levi (1978) are argued to be semantically "primitive" enough to allow for a multitude of distinct but related meanings based on context. For example, a sleeping pill is *a pill* FOR (ENABLING) *sleep*, while cold medication is *medication* FOR (ALLEVIATING) *a cold* (examples my own). Here, the relation FOR is general enough to describe the conceptual relation in both compounds, but allows for a distinct meaning in each that is discernable through context and pragmatics (i.e., medication is typically taken to cure—not cause—an illness). Levi argues that this is a common occurrence in sentences as well as compounds, and that context and pragmatics allow us to discern such meanings without the need for distinct, idiosyncratic relations. For example, a sentence such as "Jill baked a cake FOR Sue" could mean that Sue was the RECIPIENT of the cake, or that Sue was the BENEFICIARY of Jill baking the cake (on her behalf); however, the context in which the sentence is delivered will clear this ambiguity for the listener or reader. Levi argues that the generativity of compounds—and the capacity to easily interpret novel compound constructions—is suggestive of a finite list of relations that language users are capable of systematically applying and decoding by applying such pragmatic strategies.

Section Summary

The use of conceptual relations to explain and derive the meaning of complex nominals is widely agreed upon across the literature. Discrepancy arises when discussing the number, finiteness, and exact categorization of these relations, typically with respect to the degree of idiosyncratic information they must encode. The present study builds on the body of work exploring the conceptual relations proposed by Levi (1978) and expanded upon by Shoben (1991; Gagné & Shoben 1997). The topic of relation specificity is revisited under the heading *Relation Similarity* and in Chapter 4: *Discussion* as it relates to our analysis.

Theories of Conceptual Combination

Conceptual relation taxonomies and associated derivational models—such as those described by Levi (1978) and Warren (1978)—outline the proposed underlying structure of compound words from a syntactic and semantic perspective, and address the question of how their meaning is ultimately derived; however, such theories of derivation do not answer the question of how a reader or listener actively reconstructs such underlying forms. How a reader decides or 'uncovers' an appropriate conceptual relation is addressed instead in theories of the *conceptual combination* process. A prominent theory of conceptual combination is the Relational-Interpretation-Competitive-Evaluation (RICE) theory (Spalding, Gagné, Mullaly & Ji, 2010), which is the successor to the Competition Among Relations in Nominals (CARIN) theory (Gagné & Shoben, 1997). Both theories are described briefly below, and provide insight into the process of selecting a conceptual relation and its contribution to the larger process of compound word processing.

Competition Among Relations in Nominals (CARIN)

Gagné and Shoben (1997) argued that, while their list of 16 relations may not be entirely exhaustive, it allowed for the development of a predictive model of compound word processing that uses conceptual relations as a core component in the comprehension process. The Competition Among Relations in Nominals (CARIN) model demonstrated that the ease of interpreting a compound relation depended on the likelihood of that relation occurring between two constituents, weighed against the likelihood of all possible alternatives. For example, in Shoben (1991) the compound *mountain range* was found to have a longer processing time than other words containing the MADE OF relationship in a lexical decision task. Applying the CARIN model, Gagné and Shoben (1997) argued that this was because MADE OF is not a common relation for either the head (mountain) nor modifier (range) to instantiate during conceptual combination. Under CARIN, the competition generated by more commonly occurring alternatives for both constituents explained the increased processing time. In another study, Gagné (2001) found that priming a target compound with a compound that uses the same relation (e.g., *student accusation* — *student vote*, "y BY student") resulted in shorter sensicality judgement latencies, indicating decreased processing time for the target compound (note that all prime–target pairs shared the same modifier). Gagné concluded that this was likely due to increased availability of the target relation which facilitated the processing of the target modifier: a prediction under the CARIN model. In sum, the CARIN model was built upon evidence that competition exists between a constituent's possible relations, and that prior exposure to a relation could facilitate the initial relation selection process.

Relational-Interpretation-Competitive-Evaluation (RICE)

The competitive nature of conceptual relation selection identified in CARIN was expanded in the Relational-Interpretation-Competitive-Evaluation (RICE) theory of compound processing (Spalding, Gagné, Mullaly & Ji, 2010). RICE posits that the selection of a compound relation is a competitive process that unfolds in three distinct but overlapping stages of "suggest, evaluate, elaborate". First, relations that are commonly associated with the modifier are activated (suggested) and compete for selection. Second, a simultaneous competition occurs between the relations that can possibly link the modifier to the head (evaluate). The relation that best fits these criteria (appropriateness for both the modifier and the modifier-head connection) is selected as the interpretation. Third, the selected relation is used to derive the meaning of the full compound (elaborate). A key contribution of the RICE model is the ability for relation effects to occur at different stages of processing. Under RICE, the relation priming effects observed in studies such as Gagné (2001) are thought to occur during the evaluation stage, rather than the initial suggestion stage: thus, priming of a relation from a previous compound affects the ability to evaluate the final compound interpretation. This hypothesis predicts that the effects of relation priming are likely inhibitory rather than facilitatory: activation of an erroneous relation at the compound level would increase competition at the evaluation stage. This is discussed further under the heading *Compound Conceptual Relation Priming*.

Section Summary

Both the CARIN and later RICE theories of compound processing suggest a competitive selection process for compound relations that is sensitive to both the likelihood of a constituent entering a given relation (i.e., how frequently the constituent combines with other constituents while using that relation), the likelihood of that constituent entering other relations (i.e., the frequency with which the constituent enters into the desired relation relative to the frequency with which it enters into all other relations), and the plausibility of the chosen relation explaining the connection between two compound constituents. Further evidence (e.g., Gagné, 2002) suggests that this selection process can also be affected by prior exposure to conceptual relations in an effect known as relation priming. Findings from various relation priming studies are presented below.

Compound Conceptual Relation Priming

Evidence from several empirical studies of conceptual relations shows that the processing of compound words is influenced by recent exposure to a compound involving the same relation (e.g., Gagné, 2001; Maguire & Cater, 2004; Gagné et al., 2009). Specifically, it has been found that a compound is easier to process and interpret when preceded by a compound with the same relation relative to a compound with a different relation. Some of the earliest evidence for this priming phenomenon was established by Gagné (2001) who showed that a novel compound such as *student vote* (vote BY student) is easier to confirm as a sensible construction when preceded by *student accusation* (accusation BY student) than when preceded by *student car* (student HAS car). Similar effects have been reported for novel compounds in other languages, including French and Chinese (respectively: Maguire & Cater, 2004; Hongbo & Gagné; 2007) suggesting that this may be a common phenomenon across languages with compounding. Similarly, Gagné and Shoben (2002) found that a participant's understanding of an ambiguous compound (e.g., an *adolescent doctor* may be a doctor who IS adolescent, or a doctor FOR adolescents) is influenced by prior exposure to its possible relationships. In a task where participants were asked to define the meaning of ambiguous compounds, participants more readily provided the interpretation *a doctor* FOR *adolescents* when the compound was preceded by *animal doctor* (a doctor FOR animals), and the interpretation *a doctor who* IS *adolescent* when preceded by *male doctor* (a doctor who IS male). These studies indicate that conceptual relation selection and general compound processing are influenced by prior exposure to conceptual relations.

While the above studies confirm an effect of exposure to conceptual relations, they do not fully explain the nature of the effect observed: does relation priming facilitate compound processing when a prime and target share a relation (as predicted by the CARIN model), or does it inhibit processing when they use different relations (as predicted by the RICE model)? Using a lexical decision task (including known compounds and pseudo-compounds formed by combining two real words) Spalding and Gagné (2011) found evidence that relation priming effects are due to slower processing in a different-relation condition rather than faster processing in a same-relation condition. In this experiment, modifier-only primes (e.g., *snow—snowball*) were introduced as a form of baseline, under the premise that a single word would not instantiate a conceptual relation. To balance for lexical and semantic repetition, the same-relation and different-relation conditions also used primes and targets that shared the same first constituent (e.g., *snowman—snowball*, *snowshovel—snowball* respectively). Results showed that response times for same-relation and modifier-only primes did not differ, suggesting that the repetition of a relation does not facilitate target processing. Conversely, different-relation primes elicited slower response times than same-relation and modifier-only primes. Spalding and Gagné concluded that presenting a different-relation prime triggers an erroneous interpretation that conflicts with the established meaning of the compound, and must then be ruled out by the participant. In the *snowshovel—snowball* example, first processing *shovel* FOR *snow* prompts the incorrect interpretation *ball* FOR *snow* which must then be reconciled. This reconciliation process requires additional effort, which slows down the overall processing of the compound and the participant's ability to settle on the established interpretation. This computational incongruence does not occur when there is no relation primed (modifier-only primes) nor when the primed relation matches the relation of the target compound. As a result, what appears to be a facilitatory effect of same-relation priming is attributed to the absence of interference from a different-relation prime. This aligns with the predictions of the RICE model of conceptual combination and the general idea of competition effects playing a significant role in compound processing.

Section Summary

Compound processing is affected by prior exposure to compound conceptual relations. Critically, it has been shown that prior exposure to a conceptual relation different from the currently desired relation can inhibit processing time by introducing competition. This finding evidences the idea that conceptual relations are involved in the larger act of processing a compound word, and suggests that the influence of conceptual relations should be considered in studies of compound word processing. In this thesis, we examined the influence of conceptual relations in the short-term binding of compound words to a list for immediate serial recall. Specifically, we investigated the effect of repeated exposure to same compared to different relations in a list to look for similarity effects of relation priming. This is discussed further in Chapter 3 and Chapter 4.

Relation Similarity

Early studies of relation priming utilized paradigms with constituents shared between prime and target; however, the question of whether such repetition is required has been a topic of debate. Gagné (2001, 2002) and Gagné et al. (2009) found that relation priming for compounds was only obtainable when the prime and target shared an identical or semantically related constituent and that constituent occupied the same position (i.e., modifier or head). Conversely, Estes (2003) and Estes and Jones (2006) obtained relation priming effects without lexical repetition, and demonstrated that these findings were not explained by possible lexical or semantic similarity between prime and target constituents. Instead, Estes and Jones concluded that a target and prime must share a conceptual relation that is sufficiently similar if relation priming is to occur, and that the relations paired in previous experiments (e.g., Gagné 2001; 2002) were not similar enough to elicit such effects in absence of lexical repetition. Using similarity norms obtained through an offline comparison task of prime-target pairs used in previous studies, they showed that pairs with high relation similarity ratings had faster processing times than pairs with low relation similarity ratings. For example, the pair *steel scissors—straw hat* (head MADE OF modifier) was given a high relation similarity rating and was processed faster than the pair *tire rim—family cow* (head OF modifier) which was given a low relation similarity rating. In this example, Estes and Jones argue that, while the OF relation can be used to describe both *tire rim* and *family cow*, it does so in an overly general way: it does not distinguish the distinct PART-WHOLE relation of *tire rim* from the POSSESSION relation of family cow (p. 100). They conclude that the low relational similarity ratings given by participants for previous experimental prime-target pairs explained the failure to observe relation priming effects in the absence of lexical similarity in those studies. In a similar lexical decision task, Popov and Hristova (2015) found evidence for relation priming in Bulgarian using lexically and semantically dissimilar prime–target pairs (e.g., *planet core—fruit pit*) that were balanced for relational similarity following Estes and Jones (2006). This finding supports the observations of Estes and Jones that prior exposure to a compound perceived to be relationally similar to the target compound yields faster processing times than prior exposure to a relationally dissimilar prime.

Section Summary

The seemingly arbitrary categorization of compound words by their respective conceptual relations has been criticized by many researchers. It has been noted that multiple logical interpretations may exist for a single compound word, and that interpretation is likely to differ between individuals. Further, evidence suggests that relations that are overly general (i.e., encompass a variety of interpretations) are not guaranteed to instantiate the same meaning across compounds and, therefore, will not reliably demonstrate relation priming effects. The findings of Estes and Jones (2006) suggest that findings of relation priming depend on perceived relational similarity, which likely requires categorization by not overly generalized relations. This finding prompted the Conceptual Relation Similarity Survey, which was modelled on the relational similarity survey conducted by Estes and Jones (2006). The results of this survey were factored into the analysis presented in Chapter 4 and subsequent discussion.

Chapter 3

Verbal Working Memory and Compound Words

Most theories divide human memory into long- and short-term varieties; we generally recognize that our memories and tacit knowledge reside in some form of long-term memory, while our short-term memory is employed in day-to-day tasks of counting, speaking, and general remembering that do not require long-term storage. Various models, among them the working memory framework by Baddeley and Hitch (1974; Baddeley, 1983; 2000; 2012), have proposed that short-term memory is a discrete resource, whereas other conceptualizations argue that it is a subset of focused activations within long-term memory (e.g., Acheson et al., 2011; Cowan, 1999). In both cases, it is generally agreed that short term memory plays an important role in language processing and the storage and retrieval of linguistic information. The present study uses an immediate serial recall task traditionally thought to probe short-term memory and indicate how words are processed, stored, and retrieved.

Short-term memory over the last five decades has mainly been studied under the view that it is involved in—and has implications for—a variety of cognitive tasks, including language comprehension, mathematics, and visuospatial ordering (Walker & Hulme, 1999; Baddeley 1986; 1983; Hitch, 1980). Most influentially, Baddeley and Hitch (1974; Baddeley, 1983; 2000; 2012) redefined short-term memory as *working memory*: memory that provides "the temporary storage of information in connection with the performance of other cognitive tasks such as reading, problem-solving or learning" (Baddeley, 1983, p. 73). They proposed a model of working memory that incorporates four discrete but interconnected systems: the central executive, the phonological (or articulatory) loop, the visuo-spatial sketchpad, and (later) the episodic buffer. Described briefly, the central-executive is a master attentional system that coordinates information from the other three (slave) systems; the phonological loop (described in more detail below) is the critical component in language processing that stores linguistic information as a phonological code; the visuo-spatial sketch pad similarly stores visuo-spatial information; and the episodic buffer provides a modality-general workspace, connecting the phonological loop and visuo-spatial sketchpad to our long-term memory and the knowledge stored therein (Baddeley, 1983; 2000). The four components of this model work in unison to provide multi-modal methods of event encoding and recollection, thereby maximizing the likelihood of successful storage and retrieval.

Compound processing in this thesis is investigated in a task relying on verbal short-term memory, which is responsible for the short-term storage of both verbal and visual (written) linguistic information. Studies of verbal short-term memory often involve immediate serial recall tasks. In these tasks, a list of linguistic stimuli (typically between five and seven) are presented to participants who must then recall them immediately after presentation, in the order that they appeared. Measures of immediate recall include general item recall (how many words a participant remembers correctly) and serial or ordered recall (how many words a participant remembered correctly in the right order). Manipulation of the linguistic stimuli in these tasks has revealed numerous lexical and semantic factors that significantly affect recall performance. Factors such as word frequency and concreteness have shown positive effects on recall performance, while word length and morphological complexity have been found to impair recall. These findings and others are discussed in more detail in subsequent sections of this chapter. In this thesis, an immediate serial recall paradigm was employed to investigate the possible effects of conceptual relationships on compound words in verbal short-term memory.

Words in Working Memory

The most influential model of short-term verbal memory is the phonological loop model proposed by Baddeley and Hitch (1974; Baddeley, 1983). In this model, aural and visual word stimuli are encoded in working memory as a phonological code that is subject to decay over a short period of time. It is believed that traces of these phonological codes decay fully within approximately two seconds of being formed, whereupon the word is forgotten and irretrievable (Baddeley 1986, 1990). The model proposes that this decay can be halted or reversed through subvocal rehearsal (repeating a word silently to yourself), which serves as a refreshing mechanism to restore the phonological trace; as long as the trace is refreshed, it remains in working memory. This pattern of refresh and decay forms the phonological loop.

Under the phonological loop model, the capacity of an individual's verbal working memory (i.e., verbal memory span, or the maximum number of words a person can recall correctly) is proposed to be the maximum number of items that can be subvocally rehearsed within the two second decay window. When the sum of this rehearsal time surpasses two seconds, items later in the list may decay and be forgotten (e.g., Cowan et al., 1992; Estes, 1973). Further evidence for the importance of rehearsal and phonological trace-refreshing comes from studies of articulatory suppression. In these studies, researchers disrupt participants' capacity for subvocal rehearsal by having them vocally articulate an unrelated word or sound repeatedly (e.g., participants are asked to say "*ba*" repeatedly during list memorization). The use of articulatory suppression has shown a significantly detrimental effect on recall (Estes, 1973; Levy, 1971; Murray, 1967) and is argued to evidence the importance of subvocal rehearsal in maintaining verbal short-term memory traces.

A more recent approach to short-term verbal memory (and working memory in general) is the interference model (Oberauer et al., 2016; Oberauer et al. 2012; Oberauer & Kliegl, 2006; Saito & Miyake, 2004). Under this model working memory is not influenced by time-based decay, nor are linguistic representations assumed to decay on their own; instead, working memory capacity is limited by three distinct forms of interference generated by competing representations in memory. The first, *confusion*, occurs between representations at the time of retrieval and is said to be the result of non-target representations in memory having equal or higher levels of activation than the target representation. When two or more representations are equally activated, it is more difficult for an individual to remember which representation is the desired one. In the case that an erroneous representation has a stronger representation than the target, the individual may recall the erroneous one. The second form of interference, *superposition*, is proposed to occur when multiple representations are 'written' over one another in a neural network. As patterns of activation build in a neural network, the ability to disentangle such activations decreases, such that individual activations become increasingly difficult to discern and no longer serve as distinguishing information. The third, *feature overwriting*, is similar to superposition, but involves the overwriting of a specific feature or set of features that is shared between representations. In this case, the feature of one representation is overwritten so that it equals the feature of another representation, thereby increasing the similarity between the two representations and the likelihood of mistaking the two (Oberauer, 2016). Under the interference model, it is assumed that representations in working memory do not decay but, rather, their retrieval becomes increasingly hindered by the addition of various forms of interference until they are eventually irretrievable. Under this model, if time plays a factor it is only because interference increases as time goes on, and not because of an inherent expiry of representations. Both decay and interference models have mainly studied phonological forgetting from short-term memory. It is less well known why morphological and semantic information is lost from short-term storage.

Section Summary

The question of how linguistic information is encoded in short-term memory and how it degrades is longstanding and ongoing. Under the phonological loop model, the primary contributor to immediate serial recall is phonological representation and the capacity for representation refreshing; however, the model also makes room for contributions from long-term memory, as discussed in following sections. While this model assumes time-based decay, this does not make it entirely at odds with the interference model. For example, while memory decay may be the result of mounting interference and competition in an interference model, there is still room for subvocal rehearsal to refresh representations and reduce interference. Similarly, the assumed time-based decay in the phonological loop model may be explained by the accumulation of interference—including phonological interference—over time, as presented in interference models. Additionally, the question of whether short-term representations exist in a discrete form of short-term memory or as a subset of focused activations in long-term memory remains open; however, the location of these activations does not inherently preclude either model or their individual components.

The Effect of Semantic Variables and Long-Term Memory on Word Recall

A mounting body of evidence suggests that verbal memory not only relies on the short-term storage of phonological information and memoranda refreshing processes, but also on the coactivation of semantic and other non-phonological information typically associated with long-term memory. Studies by Poirier and Saint-Aubin (1995) and Huttenlocher and Newcombe (1976) found that words are easier to recall when grouped with other words belonging to the same syntactic category (e.g., nouns with nouns, verbs with verbs) than when they are mixed. This finding suggests an effect of simultaneous access of syntactic/semantic information stored in long-term memory. Tehan and Humphreys (1988) similarly demonstrated that recall for content words (e.g., nouns and adjectives) was better than the recall of function words (e.g., determiners and prepositions). The authors concluded that words with more semantic information (i.e., more robust meanings) are easier to recall than words that have less inherent meaning. This too indicates a contribution of semantic information to memory recall, and the findings align with a further study by Walker and Hulme (1999) which found that, across three experimental paradigms, concrete words exhibited higher rates of recall than abstract words. Similarly, Bourassa and Besner (1994) found that the imageability of a given word has a significant effect on its likelihood of recall, wherein more imageable words are more likely to be recalled than less imageable words.

Section Summary

Taken together, these findings suggest that verbal working memory involves the coactivation and retrieval of semantic and other lexical information stored in long-term memory in addition to the encoding and subvocal rehearsal of phonological representations and other trace-refreshing processes. Although the exact details of the various models differ in what causes short-term representations to become unrecallable, they appear to converge around the idea that lexico-semantic properties of words—including concreteness, semantic category, and syntactic category—can contribute to word recall performance (Baddeley and Ecob, 1970). In Experiment 2 of this thesis, the conceptual relations between the constituents of compound words (here assumed to be a semantic feature) are examined for possible effects on immediate recall performance.

Similarity Effects in Working Memory: Phonological and Semantic

Effects of phonological similarity between list items in memory recall tasks are well documented (Conrad, 1964). The results generally indicate that phonological similarity facilitates general word recall, but inhibits order recall (i.e., people are more likely to remember similar sounding words, but are also more likely to confuse their order). For example, Nimmo and Roodenrys (2006) found that phonological similarity (sharing rhyme or initial onset and vowel) resulted in poorer serial recall performance (order memory), but enhanced performance for item recall. The authors concluded that the facilitatory effect in item recall resulted from phonological similarity restricting one's memory search to words cued by the shared component and increasing activation at the shared phonemic level, thereby increasing the overall odds of recalling an item. The negative effect on serial recall, however, was attributed to the sharing of phonological features which restricts the amount of distinguishing phonological information available as a recall cue for each specific position in the list. Because recall items are less distinct, participants found it more difficult to discern the order of items accurately. Various additional studies have found similar phonological similarity effects on both serial and item recall performance (Baddeley, 1986; Baddeley et al., 1984; Besner & Davelaar, 1982; Conrad, 1964).

Historically, semantic similarity effects were widely cited as being in line with phonological similarity effects: semantic similarity increasing item recall, but hindering order recall (Baddeley, 1966; Crowder, 1979; Nairne, 1990; Nairne & Neumann, 1993). More recently, the generality of this finding has come under debate. Poirier and Saint-Aubin (1995) found that similarity of word class (noun, verb, etc.) showed the anticipated facilitatory effect on item recall, but had no observable effect on order recall. This finding was replicated in Saint- Aubin and Poirier (1999) and Saint-Aubin, Ouellette, and Poirier (2005), which led the authors to conclude that semantic properties of words do not influence serial recall processes. Conversely, a later study by Acheson et al. (2011) used a dual-task involving simultaneous picture judgement and delayed serial recall and found that order errors were more common in the recall task when the picture judgement task introduced interference in the form of semantic processing (asking participants to judge semantic category) than when it introduced interference in the form of spatial processing (asking participants to judge spatial orientation).This effect disappeared when the study was repeated using non-words. The authors concluded that lexical-semantic representations can affect serial ordering in short-term memory, particularly in the presence of semantic interference. In a follow-up study, Poirier et al. (2015) replicated this finding in three distinct serial recall experiments that contrasted semantically-similar and −dissimilar lists of memoranda. Across all three experiments, the authors found that in lists where the first three words were semantically similar to the fifth (and the fourth and sixth were dissimilar) the fifth word was more likely to be recalled out of position (earlier in the list) than in lists where no words were similar. They proposed that the first three items activate the fifth item via semantic similarity, making its level of activation similar to their own despite its temporal distance. As a result, those four items share similar levels of activation and are more likely to be confused during order recall. In the case of item recall, semantic similarity was found to be facilitatory, as in previous studies. Considering these findings, Poirier et al. proposed a model of order recall that proposes that (1) semantic information affects order recall in short-term memory and (2) order recall is, more generally, influenced by the degree of activation of relevant lexico-semantic information in long-term memory.

A Note on the Duration of Semantic Similarity Effects

In a study that manipulated stimulus presentation times, Howard and Kahana (2002) found that semantic similarity effects in the recall of serially presented memoranda are greatest when memoranda are presented rapidly, without interruption, and that the more memoranda are separated temporally, the weaker the resulting effect. This suggests that semantic representations do not benefit from subvocal rehearsal in the same way that phonological code does, and suggests that studies of semantic similarity effects should be conducted with short inter-stimulus time intervals. This observation is significant to the present experiment, which was designed to employ a 1200 ms inter-stimulus presentation rate to facilitate maximal similarity effects.

Section Summary

Although evidence has differed historically, current positions of phonological and semantic similarity indicate that both provide facilitatory effects on item recall, but hinder performance in order recall. While the exact explanation for these effects may vary in different models of verbal short-term memory and recall, there is some agreement that these effects are, at least partially, due to influences on the degree of activation of phonological and semantic representations. As the activation of a representation is enhanced, its probability of recall is also heightened; however, the likelihood of recalling an item in the correct serial position is determined by the number of items with similar or greater activation than the target.

Additional Observed Effects on Word Recall Performance

Word recall performance has also been observed to be affected by word frequency, primacy and recency list positions, and practice. Word frequency has a profound effect on word recall. Words with a higher frequency of occurrence (which may also be thought to reflect word familiarity) are more likely to be recalled correctly than words with low frequency, due to stronger representations and richer associative networks (Poirier & Saint-Aubin, 1996; Engle, Nations & Cantor, 1990; Tehan & Humphreys, 1988). Most models of short-term serial recall include some form of primacy gradient (see Hurlstone, Hitch, & Baddeley, 2014 for review). A primacy gradient implies that the strength of representation encoding decreases as memoranda are presented in a list, such that items presented early receive stronger activation and are more likely to be recalled than items later in the list. In the phonological loop model, early memoranda are thought to benefit from increased opportunity for subvocal rehearsal. Interference models suggest that early list items are subject to less interference and, therefore, benefit from stronger encoding than later items. Under models assuming long-term memory activations, the primacy gradient occurs across those activations. In serial recall of near-span lists a (generally small) recency effect favouring the last item or items in a list is also commonly reported, although this effect appears to be susceptible to variation in stimuli and testing conditions (also Hurlstone et al., 2014).

Section Summary

Taken together, evidence suggests that verbal short-term memory is affected by a variety of lexical, semantic, and temporal factors that can enhance or inhibit recall performance accordingly. Although different models of verbal memory have different approaches to explaining these effects, their contributions to recall performance are of consequence to designs employing immediate serial recall and, therefore, have been considered in the selection of recall stimuli for the present study and in the analysis of our findings. This is discussed further in Chapter 4.

WM and Morphologically Complex Words

Limited research has been conducted on morphologically complex words in working memory as morphology has not been considered in any of the current working memory models. However, recent studies have begun to uncover the effects of morphological complexity on the short-term storage and retrieval of linguistic information (e.g., Service & Maury, 2015; Németh et al., 2011; Reinitz & Hannigan, 2004; Service & Tujulin, 2002). In particular, empirical studies of morphologically complex words in immediate and delayed recall tasks have shown evidence that increased morphological complexity decreases word recall performance, notably without confound of word length effects. Németh et. al. (2011) compared immediate serial recall performance for morphologically complex and simplex Hungarian words matched for phonological length (all words were two syllables), phonological structure (consonant-vowel ordering), frequency, and concreteness. Results of the study revealed a significant recall advantage for morphologically simplex words (i.e., monomorphemic words), and indicated that verbal short-term memory span is negatively affected by morphological complexity. The study also investigated possible differences in morphological composition, and found that recall performance was better for derived words (e.g., *boy+hood)* than for inflected words (e.g., *boy+s*). The authors propose that the findings can be explained by an increased representational load of morphologically complex words. Each morpheme represents a “chunk” of semantic or grammatical information that must be represented in working memory, and the inclusion of multiple representations per word decreases the available working memory resources (as per a resource-based model). Although not discussed by the authors, the finding that derived words exhibit better recall outcomes than inflected words is compatible with the conclusions of Poirier and Saint-Aubin (1995) and Walker and Hulme (1999) who suggest that stronger semantic representations facilitate memory recall. The findings of Németh et al. align with those of Service and Tujulin (2002) and Service and Maury (2015) who found that recall performance in

Finnish fit a similar gradient, wherein monomorphemic words had the highest performance, inflected words the lowest, and performance for derived words was situated in between. Similar to above, Service and Maury (2015) suggest that Finnish derivational suffixes provide additional semantic representations, and that Finnish derived words benefit from heightened activation of semantic information encoded in long-term memory.

At this time, few studies of have closely examined compound words in working memory. Reinitz and Hannigan (2004) examined compound recall performance across delayed recognition and recall tasks to investigate the possible effects of stimulus presentation methods on memory performance. Across multiple experiments, participants memorized word pairs that were presented either simultaneously or sequentially with the goal of observing possible effects on constituent recombination errors. Participants were then asked to complete either a recognition task (identifying recalled words from a list including experimental stimuli and distractors formed from constituent recombinations), or a recall task (writing down as many words as they could remember). Reinitz and Hannigan found that participants consistently made recombination errors in the recognition task regardless of presentation method, while the recall task only showed within-list recombination errors, and only for word pairs presented simultaneously. When participants were shown word pairs sequentially, they did not produce recombination errors in recall. The authors concluded that the deliberate recall process generally shields participants from recombination errors: as the participant recalls the first constituent, the representation of the second constituent becomes activated through association more strongly than other representations. Serial presentation of word stimuli enhances this effect by isolating each representation in memory, thereby reducing the likelihood of recombination. In the case of the recognition task, the presence of false recombinations (distractors) produces stronger associations for erroneous constituents than would otherwise be present, thereby increasing the likelihood of false alarms for recombination distractors. This is similar to the effect found with simultaneous presentation of stimuli in the recall task. Wälchli (2016) conducted a comparative study of compound words in immediate serial recall and complex span tasks to investigate whether short- and long-term memory may be differentially sensitive to the representation and processing of compound words. The complex span task involves the memorization and recall of a list of memoranda in the presence of sentence processing demands (participants were asked to read and process sentences between the presentation of memoranda items). The complex span dual task is believed to require different encoding strategies as sentence processing is too time consuming for effective phonological refreshing. The study found differing patterns of recall errors for the two tasks: individual constituent errors reflecting the decomposition of compound words were more common in traditional immediate recall than in complex span, while general omissions were more common in complex span. Wälchli concluded that this was evidence of immediate serial recall relying heavily on phonological encoding of individual constituents, which would increase the chance of retrieving only one constituent, while complex span appeared to rely on whole-word activations, which were possibly facilitated by longer-lasting semantic activation, compatible with evidence for the slower decay of semantic effects (e.g., Campoy et al., 2014).

Section Summary

Studies of complex words in verbal short-term memory suggest that morphological complexity hinders word recall, but that this detriment may be partially offset by support from richer semantic representations. This may be especially true for compound words, which typically consist of two words with clear meanings (even if those meanings do not directly relate to the compound itself). In the present study, we utilized an immediate serial recall paradigm employing serial visual presentation which, based on Reinitz and Hannigan (2004), should further offset the inhibitory effects of morphological complexity by providing an environment that isolates individual compound words during memory encoding. Conversely, Wälchli (2016) found that immediate serial recall appeared to rely more heavily on phonological encoding of constituents, and that semantic encoding of whole-words seemed to be a more effective strategy in complex span tasks. Taken together, the question of what effect semantic features of compound words have on immediate serial list recall has not been previously investigated. The current study will investigate the effect of compound conceptual relations (herein considered a semantic feature of compound words) on immediate serial recall and working memory processes more generally.

Chapter 4

The Present Study

The present study investigated the possible effects of compound relation priming in verbal working memory via an immediate serial recall task. In this task, participants were required to memorize lists of compound words then recall them out loud and in order immediately after presentation. Lists differed in whether the constituents in the presented compounds were bound by the same or different conceptual relations. In previous studies of simplex words, semantic similarity in lists of memoranda has been shown to increase overall recall performance, but also to reduce order accuracy: participants are more likely to remember groups of similar words, but are less likely to recall them in the correct order. Poirier et al. (2015) attribute both effects to enhanced activation of list representations: increased activations make items easier to recall, but equalization of their activation makes remembering their order more difficult. Studies of complex words (particularly compound words) in working memory are few, and the effect of semantic similarity on the recall of complex words is not well understood. If semantic similarity affects complex word recall similarly to simplex words by providing a shared memory cue, then we expect the present study to show an increase in overall recall for compound words presented in same-relation lists. As multiple compounds sharing the same conceptual relationship are presented, their representational activations should be enhanced by this shared semantic feature, making them easier to recall than lists of compounds that do not benefit from this enhanced activation. At the same time, evidence suggests that semantic similarity inhibits order recall in simplex words, which may also be predicted in the same-relation environment; however, the unique complexities of ordering multiple (first and second) constituents in the case of compound word recall make it difficult to compare simplex and compound word predictions with respect to serial ordering.

Theories of conceptual combination and compound processing proposed by Gagné and Shoben (1997) and Spalding et al. (2010) argue that the selection of a conceptual relation during compound word processing is an inherently competitive process that operates at whole-word and constituent levels. Under the RICE model (Spalding et al., 2010), associations exist between constituents and the relations with which they are frequently paired. As one processes a compound, relations associated with each constituent are activated and compete for selection. In a priming paradigm, a primed relation competes more strongly for selection in this process and generates additional interference when it does not match the desired relation. If this is mirrored in working memory, we expect to see, as above, reduced recall performance for different-relation lists where participants are more likely to experience competition affecting semantic aspects of compound word processing during presentation and recall.

However, the contributions and possible competition between complex word constituents in short-term memory are not well understood. The proposal that conceptual relations serve as a shared means of representational activation assumes that conceptual relations enhance activations at a whole-word level, or that they do not introduce competition between individual constituents. Evidence of a negative effect on recall of same-relation lists could result from the activations being facilitated at the constituent level and/or that shared conceptual relations introduce competitive effects between constituents from different items. The inherently associative relationship between conceptual relations and compound constituents differs from the more general category membership explored in previous studies of semantic categorical similarity of simplex words. While semantic categorical membership facilitates activation within the given category network, the primary association is between individual members of the category and the category itself. Conceptual relations, however, instantiate a network of constituents with which they frequently co-occur (and vice-versa), and these constituents must be recalled in combination. An increase in constituent activation (both in short-term memory and in the long-term store) may increase competition during compound recombination, thereby reducing recall performance.

To help validate the experimental conditions in the recall task, this study also included a participant survey regarding the perceived similarity of compound conceptual relations. A survey modelled on Estes and Jones (2006) was issued to participants and asked them to compare compound words from the recall lists and rate the similarity of their conceptual relations. If the same- and different-relation conditions are well formed, we expect the same-relation condition to have significantly higher similarity than the different-relation condition.

Immediate Serial Recall Experiment

Experiments investigating the visual processing of compound words have found evidence of an effect of compound conceptual relations on processing time. Conceptual relation priming paradigms (Gagné, 2001. 2002; Gagné & Shoben, 2002; Gagné & Spalding; 2004; 2009) have shown that compound words are processed faster when preceded by a compound with the same conceptual relation between its constituents than when preceded by a compound with a different relation. Evidence suggests that this effect on processing is the result of increased competition in a different-relation environment rather than facilitation in a same-relation environment (Spalding & Gagné, 2011).

The serial recall experiment was designed to investigate the effect of conceptual relation similarity in working memory by analysing individual performance in immediate serial recall of lists of compound words. This experiment involved the rapid serial visual presentation of lists of memoranda consisting of either same-relation compound words (experimental lists) or different-relation compound words (control), followed by their immediate oral recall. Semantic similarity effects in working memory have been shown to facilitate item recall through increased activation of memory representations (Baddeley, 1966; Crowder, 1979; Nairne, 1990; Nairne & Neumann, 1993), and to impair order memory by equalizing activations of competing representations and restricting distinguishing information (Acheson et al., 2011; Poirier et al., 2015). Based on simplex word recall, we anticipated that item recall for the experimental lists would be facilitated and order recall might be inhibited relative to control lists of items with different conceptual relations. An observed difference in recall performance between these two relation similarity conditions would suggest that conceptual relations contribute to semantic representations that affect compound word recall processes.

Methods

Participants

44 participants (39 female; 5 male) were recruited through a convenience sampling of McMaster University undergraduate students enrolled in the Linguistics Research Participation System. Five participants identified as having a reading or learning disability and their data were removed from the analysis based on exclusion criteria (resulting in the final *n* = 39; F = 34; M = 5). Participation was also restricted to native English speakers, who ranged in age from 19 to 27 (M = 21.4). All participants had normal or corrected to normal vision. Participants were granted course credit for participation as per terms outlined by their individual course instructors.

Materials

80 unique English noun-noun compound words (e.g., *snowman*) were selected from a corpus compiled by Schmidtke, Kuperman, Gagné, & Spalding (2016) with equal distribution among four attributed conceptual relationships (i.e., 20 compounds per relationship): H MADE OF M, H FOR M, M HAS H, and H HAS M. To increase the reliability of the attributed relationship type, we selected compounds that had relation entropy scores less than or equal to the mean relation entropy of all compounds in the database (≤2.23). The mean relation entropy of the selected compounds was 1.67 (SD = 0.44). Entropy data was also derived from Schmidtke, Kuperman, Gagné, & Spalding (2016). Selection was further refined by number of letters (Range = 7–10*;* M = 8.31) and concreteness (M = 4.56; SD = 0.53), while frequency was balanced across lists (Range = 12–2421; M = 482).

Two types of recall list were created to form the two experimental conditions, with each list consisting of four compounds. Lists in the first condition were made with four compounds that used the same conceptual relationship (hereafter the SAME relation condition). Lists in the second condition were made with one compound representing each of the four conceptual relationships (hereafter the DIFFERENT relation condition). Examples are provided in Table 4 below (see *Appendices* for all stimulus lists). All 80 compounds were used once per condition, resulting in 20 lists per condition. An additional 12 compounds were used to form three practice lists: one list belonging to the SAME condition and two lists belonging to the DIFFERENT condition. These compounds did not appear outside of the practice lists and are excluded from the analysis.

Table 4: sample lists from SAME and DIFFERENT relation conditions.

Same Relation		Different Relation	
Compound	Relation	Compound	Relation
newsboy	H HAS M	*backache*	M HAS H
clipboard	H HAS M	*cornmeal*	H MADE OF M
songbook	H HAS M	*handrail*	H FOR M
minefield	H HAS M	*sandbox*	H HAS M

Participants were presented with all 40 lists from both conditions (160 compounds) so that each compound was presented to every participant twice. Lists were pseudo-randomized into four orders to help control for order effects. Each order was presented to an equal number of participants (11 participants per order). The order and structure of practice lists were identical for all participants.

Procedure

The experiment was programmed using SuperLab 5 (Cedrus Corporation, San Pedro, California) with each compound word presented in the center of the screen using Rapid Serial Visual Presentation (RSVP). All compounds were presented as single words (no spaces) in fixed-width font. Lists were preceded by a fixation cross and followed by a recall prompt. The experiment was conducted on-premises at the McMaster University *Language, Memory and Brain Laboratory* and took approximately 10 minutes per participant to complete. Participants were seated in front of a computer screen and instructed to read the four recall words silently as they appeared, then, when prompted on screen, to orally recall them in the order they had appeared. Recall lists were presented at a rate of 1200ms per word and the recall prompt appeared 1200ms after the presentation of the fourth word. The recall prompt remained on screen until the participant manually continued to the next list via spacebar press after completing the recall task. Participant responses were recorded manually by the researcher. Three (3) practice lists and forty (40) experimental lists were presented in total.

Participants were instructed to say *'blank'* when they were not able to recall a compound and, in the case that they could only recall part of a compound, to say the part they remembered in combination with '*something'* (e.g., *something-field*; *snow-something*). Participants were also told that each compound would appear twice during the experiment, and that they were permitted to take a short break between lists.

Variables and Statistical Considerations

The dependent variable for the current study is participant recall performance. Two measures of recall performance were captured: (1) absolute order score, the correct item recalled in the correct serial position, and (2) item score, an item recalled correctly from the current list, but not necessarily in the correct position. Scores were recorded as 1 for *correct* and 0 for *incorrect* or *omitted* (*'blanks'*). To tabulate and analyse various forms of recall errors, order and item scores were also calculated for each constituent. An example of the scoring breakdown is provided in Table 5 below.

Table 5: example order and item recall scoring.

Target		Recalled		Order	Item	Order		Item	
C1	C2	C1	C2	Word	Word	C1	C2	C1	C2
mouse	trap	mouse	trap	1	1	1	1	1	1
dust	pan	dust	something	0	0	1	0	1	0
life	guard	face	mask	0	1	0	0	1	1
face	mask	life	guard	0	1	0	0	1	1

Generalized linear mixed effects multiple logistic regression models (Baayen, Davidson, & Bates, 2008; Baayen, 2008) were run using the lme4 package for R (version 3.3.1, R Core Development Team, 2016) to account for the binary distribution of the dependent variable (i.e., correct/incorrect). The main question of this study was whether conceptual relation similarity environment (SAME/DIFFERENT) affects memory recall performance for compound words. Compound frequency and length effects were modelled with norms obtained from the COCA and CELEX databases, respectively. Practice, fatigue and other effects developing over trials were modelled as a trial number effect. Using the likelihood ratio test, fixed effects and interactions that were not found to significantly improve performance of the model were removed. Participants and recall words were included as random effects to account for variance between individual performance and recall items. The model was fitted with the H FOR M relation category and DIFFERENT relation environment set as reference levels. The reported figures show the fixed effects that were kept in the final model.

Table 6: fixed effects of the generalized linear mixed-effect model for recall accuracy by lexical variables and relation similarity environment. The standard deviation estimate for the random effect of Compound is 0.29. The standard deviation estimate for the random effect of Participant is 0.88. Number of trials = 6,240.

	Estimate	Std. Error	*z* value	*p* value
Intercept	0.360	0.177	2.035	0.042
Trial Number	0.078	0.028	2.804	0.005
Compound Frequency	0.105	0.046	2.300	0.021
Number of Phonemes	−0.101	0.049	−2.061	0.039
H HAS M	0.099	0.150	0.660	0.510
H MADE OF M	0.239	0.150	1.592	0.111
M HAS H	−0.120	0.155	−0.774	0.439
SAME Environment	0.235	0.111	2.115	0.034
H HAS M: SAME Environment	−0.483	0.157	−3.084	0.002
H MADE OF M: SAME Environment	−0.583	0.157	−3.708	0.000
M HAS H: SAME Environment	−0.023	0.157	−0.145	0.884

Results

Mean item recall (*M* = 2.58, *SD* = 0.55) was significantly higher than mean serial order recall (*M* = 2.32, *SD* = 0.71); $t(38) = -6.02$; $p < 0.001$. As serial scoring is a more conservative measure of immediate recall, serial scores are reported here; however, the effects reported below were observed across models fitted for both serial and item scores, except for the main effect of relation environment (SAME/DIFFERENT), which was significant in the serial scoring but missed significance in the item scoring ($p = .094$). Results anticipated based on previous studies of the immediate serial recall of monomorphemic and compound words are presented first.

The final model fitted to serial order scores found a significant effect of trial number ($\beta = 0.08, SE = 0.03, z = 2.80, p = 0.005$) indicating that participants were more likely to recall a compound correctly as the experiment progressed. A significant effect of compound frequency ($\beta = 0.11$, $SE = 0.05$, $z = 2.30$, $p = 0.021$) where higher compound frequency predicted higher recall performance, replicated previous findings that show high frequency simplex and complex words are more likely to be recalled than low frequency words (e.g., Engle, Nations & Cantor, 1990; Németh et al., 2011). Finally, a significant effect of compound length measured in phonemes was found ($\beta = -0.10, SE = 0.05, z = -2.06, p = 0.039$) and indicated that longer compound length predicted lower recall performance. This too was predicted for immediate serial recall (Baddeley, Thomson, & Buchanan, 1975; Németh et al., 2011).

Analysis of the key variable of interest, compound relation environment (SAME/DIFFERENT), found a significant effect of environment on recall performance ($\beta = 0.24$, $SE = 0.11$, $z = 2.12$, $p < 0.034$) which was qualified by a significant interaction between relationship category and environment for compounds using the H HAS M ($\beta = -0.48$, $SE = 0.16$, $z = -3.08$, $p = 0.002$) and H MADE OF M ($\beta = -0.58$, $SE = 0.16$, $z = -3.71$, $p < 0.001$) relations. This harmful similarity effect indicated that compounds bound by these relations were significantly less likely to be recalled correctly in a same-relation environment. This is visually represented in Figure 2Figure 2: recall accuracy as a function of relation category

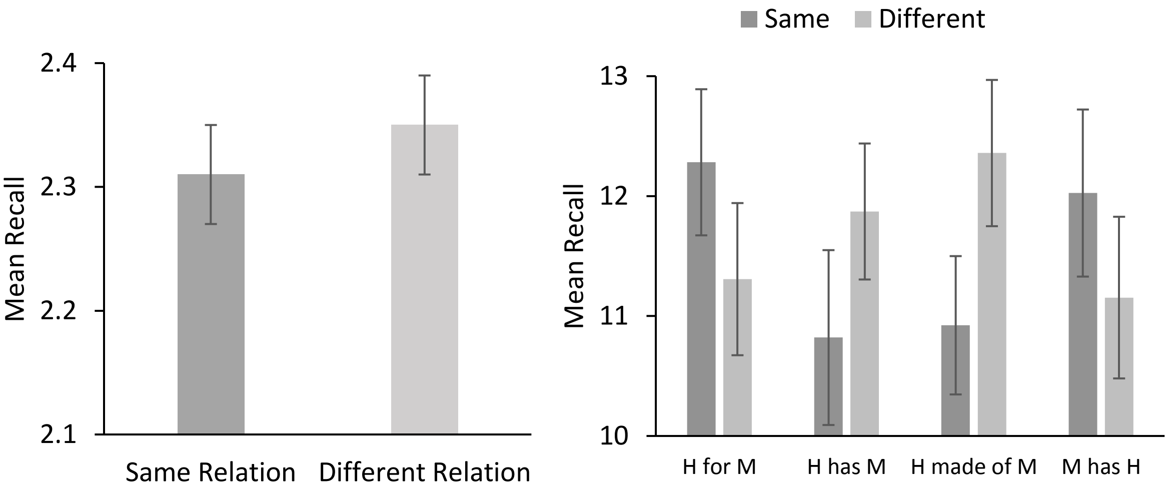

Figure 1: mean recall performance as a factor of relation similarity environment (left) and conceptual relation category (right).

and environment

following.

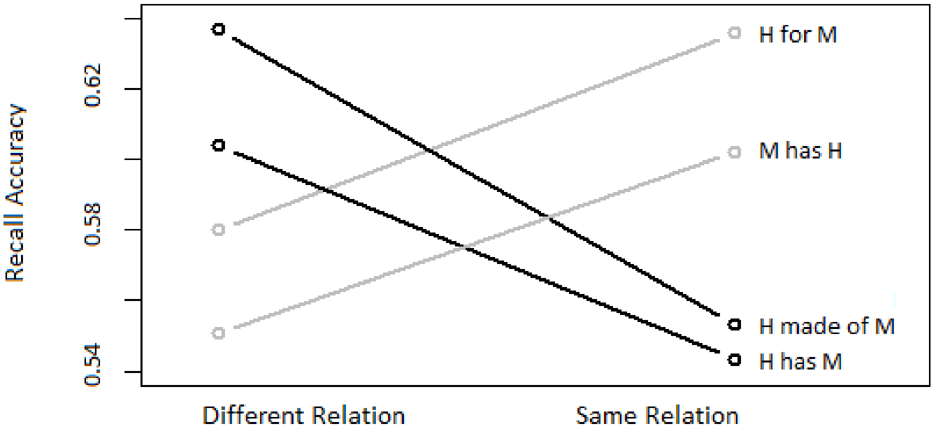

Figure 2: recall accuracy as a function of relation category and environment

Regression Analysis of Lexical and Semantic Variables

To examine the possible effects of lexical and semantic factors not considered by the experimental design, the model was refitted to include the following lexical variables obtained from Schmidtke, Kuperman, Gagné & Spalding (2016): left and right constituent frequencies, left and right constituent positional family size, left and right constituent positional family frequency, and computational measures of Latent Semantic Analysis (LSA) (Landauer & Dumais, 1997) as indicators of semantic transparency. Constituent frequencies and family-based estimates were originally derived from the 51 million-token SUBTLEX-US corpus (Brysbaert & New, 2009) and the 18 million-token English component of the CELEX lexical database (Baayen, Piepenbrock & Van Rijn, 1995) respectively. Constituent positional family size and family frequency are the number of compounds that share a given constituent in the same position, and the summed frequency of those compounds, respectively (e.g., how many compounds include the constituent *snow* in the left position, and the summed frequency of those compounds). LSA scores reflect estimates of semantic distance between paired words based on the co-occurrence of those words in a given semantic space (Landauer & Dumais, 1997) and have been previously employed as measures of semantic transparency (Schmidtke, Kuperman, Gagné & Spalding; 2016; Marelli et. al, 2015; Pham & Baayen, 2013). Measures where a higher score suggests greater semantic similarity for three semantic relationships were employed for three word-pairs, where a higher score suggests greater semantic similarity: left constituent and compound (e.g., *snow* and *snowman*), right constituent and compound (e.g., *man* and *snowman*), and left constituent and right constituent (e.g., *snow* and *man*).

Examination of mean left constituent frequency across conceptual relation categories showed that the two categories that demonstrated an effect of relation environment in the main analysis (H HAS M, H MADE OF M) had significantly lower mean left frequencies than those categories that did not show an effect (H FOR M, M HAS H). To investigate a possible confound between left constituent frequency and relation category, the model was fitted to include left constituent frequency and SAME environment as an interaction. The remaining lexical and semantic factors were fitted as main effects.

Table 7: fixed effects of the generalized linear mixed-effect model for recall accuracy by additional lexico-semantic variables and relation similarity environment. The H FOR M relation and DIFFERENT environment were set as reference levels. The standard deviation estimate for the random effect of compound is 0.26. The standard deviation estimate for the random effect of participant is 0.88. Number of trials = 6,240.

	Estimate	Std. Error	*z* value	*p* value
Intercept	0.378	0.184	2.054	0.040
Trial Number	0.073	0.030	2.477	0.013
Number of Phonemes	−0.142	0.049	−2.923	0.003
Compound Frequency	0.159	0.050	3.166	0.002
Left Frequency	−0.144	0.065	−2.207	0.027
Right Frequency	−0.093	0.056	−1.660	0.097
Left Family Size	−0.051	0.057	−0.886	0.375
Right Family Size	−0.121	0.101	−1.201	0.230
Left Family Frequency	0.021	0.054	0.384	0.701
Right Family Frequency	0.138	0.096	1.429	0.153
Left-Right LSA	0.123	0.051	2.389	0.017
Left-Compound LSA	−0.001	0.051	−0.025	0.980
Right-Compound LSA	−0.160	0.054	−2.945	0.003
SAME Environment	0.161	0.128	1.256	0.209
H HAS M	0.277	0.180	1.539	0.124
H MADE OF M	0.105	0.165	0.635	0.525
M HAS H	−0.291	0.174	−1.671	0.095
Left Frequency: SAME Environment	0.159	0.065	2.455	0.014
H HAS M: SAME Environment	−0.383	0.175	−2.190	0.029
H MADE OF M: SAME Environment	−0.468	0.180	−2.607	0.009
M HAS H: SAME Environment	0.015	0.172	0.087	0.931

Results of the refitted model showed a significant main effect of left-right LSA score ($\beta = 0.12$, $SE = 0.05$, $z = 2.39$, $p = 0.017$) which indicated that higher left-right semantic similarity predicted higher recall performance. The model also showed a significant effect of right-compound semantic similarity ($\beta = -0.16$; $SE = 0.05$; $z = -2.95$; $p = 0.003$) indicating that higher similarity between the right constituent and compound predicted lower recall performance. A significant main effect of left constituent frequency ($\beta = -0.14$, $SE = 0.07$, $z = -2.21$, $p = 0.027$) was observed;

however, this was qualified by a significant interaction between left constituent frequency and SAME environment ($\beta = 0.16$, $SE = 0.07$ $z = 2.46$, $p = 0.014$). This interaction indicated that higher left constituent frequency predicted higher recall performance in the SAME environment. Inclusion of these additional lexical and semantic variables accounted for a significant portion of variance, but did not eliminate the observed interaction between relationship category and environment for compounds using the H HAS M and H MADE OF M conceptual relations. Results of the model suggest that the observed differences for compounds bound by these relations are not explained by the presented lexical and semantic metrics alone.

Discussion

Effect of conceptual relation environment on immediate recall of compound words

Results from a linear mixed effects model fitted to serial recall performance found a significant interaction between relation environment (SAME/DIFFERENT) and relation category. This interaction indicated that compounds using H HAS M and H MADE OF M conceptual relations were significantly less likely to be recalled correctly in a same-relation environment. The effect differed from observed effects in lexical decision, sensicality judgement, and compound definition tasks that have shown facilitatory effects on compound processing in same-relation environments relative to different-relation environments (Gagné, 2001; Gagné & Shoben, 2002; Maguire & Cater, 2004; Hongbo & Gagné; 2007). It also differed from the findings of semantic similarity in working memory (Acheson et al., 2011; Poirier et al., 2015) which show that semantic similarity between list items increases item recall performance. Additionally, studies of semantic similarity in recall have shown detrimental effects on order recall performance; however, results of the current study did not reveal an effect of relation environment on the number of order errors produced. Conversely, this observation aligns with the possibility of conceptual relations activating constituent-level representations and thereby inducing additional competition in the compound word reassembly process during recall. We provide an analysis of this view below.

Under the RICE model, the effects of relation priming found in lexical decision and sensicality judgement tasks are explained by competing relation activations: a relation that has been encountered previously has a stronger activation (relative to an unencountered or temporally distant relation) and will compete more strongly for selection in the processing of a target compound. When prime and target do not share the same relation, the different relation activations compete for selection, thereby increasing processing time. In the case that a prime and target do share a relation, priming does not result in competition and the target is processed at item-specific baseline level.

In a memory recall environment, the heightened activation of a conceptual relation through repeated exposure may result in a form of competition at the time of recall that is not present in these other tasks. Taking the view that compounds are decomposed during processing, we argue that their reassembly during recall may be hindered by increased competition between constituents that have been activated via the spreading activation of a shared relation. We propose that as a relation becomes increasingly activated, so too do constituents that are commonly associated with that relation; while the increased activation of a single relation restricts competition between relations, the increased activation of constituents within its semantic network introduces competition in constituent selection during recall. This proposal aligns with proposed interference models of working memory which associate distinctness of phonological and semantic representations of items with more efficient parsing of strings into memoranda and selection of items for retrieval (Oberauer et al., 2016; Oberauer et al. 2012). Under such models, the sharing of phonological and semantic features can lead to a variety of recollection errors due to impaired ability to distinguish between target and other active representations (Oberauer et al., 2016). We propose that sharing a conceptual relation results in a form of under-specification that interferes with the compound recombination process at recall. As participants identify and recall the first constituent (modifier) of a compound word, this becomes a retrieval cue, and they must search through the remaining constituent representations to select the appropriate second constituent (head). The heightened activation of a single conceptual relation makes it a possible competing retrieval cue, as it results in the spreading activation among possible combinatorial constituents within its semantic network. This increases the number of possible head constituents one must search through (and ultimately reject) during recall.

Limitations of the Relation-Condition Interaction

Post-hoc analysis showed that compound words bound by the H HAS M and H MADE OF M relations had significantly lower left constituent frequencies than those bound by H FOR M and M HAS H compounds. This raises the question of whether the observed interaction between conceptual relation category and SAME environment condition could have resulted from a modifier frequency effect. Regression analysis showed that left constituent (modifier) frequency had a negative main effect on recall, which was modified by a positive interaction with SAME environment condition. This facilitatory interaction does not follow previous research on compound word recall, which has suggested an inhibitory effect of left constituent frequency. Wälchli (2016) found that a higher left constituent frequency predicted a lower overall rate of memory recall, and that a larger left family size predicted a higher rate of head error in a similar study of the immediate serial recall of compound words. The study concluded that compounds with larger left family sizes (e.g. *snowflake*, *snowman*, *snowshoe*, etc.) and higher left constituent frequencies use the modifier as a retrieval cue, but that the larger left family size results in greater activation across the morphological family network, thereby introducing competition in the head selection process. The present observation that compound words with high left constituent frequency are recalled better when they occur with other high left constituent frequency compounds does not align with the findings of

Wälchli (2016), nor lend itself to an obvious explanation. Available evidence suggests that the co-occurrence of high left constituent frequency compounds would introduce additional competition in recall, resulting in reduced recall performance. Further, the observed interaction between left frequency and SAME environment did not abolish the interaction between relation category and SAME environment, suggesting that they affected separate pools of variance.

We propose that the observed interaction between left constituent frequency and conceptual relation category is both a logical product of high frequency words entering a greater variety of relational associations, and the result of underspecified relational categories. Infrequent words (i.e., low frequency modifiers) will enter a smaller number of conceptual relations as there are generally less contexts in which they are likely to occur. This results in highly specific relation assignment for infrequent words. Conversely, high frequency words (i.e., high frequency modifiers) are likely to enter a variety of conceptual relationships that span nuanced distinctions in meaning. This opens a greater possibility for the miscategorization of compound conceptual relations when a compound contains a high frequency modifier. Following the criticism of Estes and Jones (2006), we argue that the H FOR M and M HAS H relational categories are overly generalized, and that the members of these categories are likely bound by a variety of overlapping but distinct semantic relationships. We suggest that this lack of homogeneity accounts for the failure to observe an interaction between these relation categories and the SAME relation environment. This possibility is explored further in the Conceptual Relation Similarity Survey.

Effect of Lexical Variables on Recall Performance

Examination of additional lexical variables revealed a significant effect of the semantic similarity between the left and right constituents (modifier and head) which indicated that higher left-right semantic similarity accurately predicted greater recall performance. Although frequently interpreted as a measure of compound transparency (e.g., Fruchter & Marantz, 2015; Pham & Baayen, 2013; Gagné & Spalding, 2009; Rastle, Davis, & New, 2004), LSA scores are derived through measures of word co-occurrence within a given semantic space and, therefore, also provide a more general measure of word co-occurrence. We propose that higher levels of co-occurrence for two constituents in the same semantic context results in increased co-activation of those constituents in a neural network: two words that frequently co-occur are more likely to activate one another when encountered. Increased co-activation during memory recall increases the likelihood that two constituents will be recalled together.

Analysis of the model also revealed that increased semantic similarity between the right-constituent and compound predicted lower recall performance. Although the co-occurrence of left and right constituents appears to facilitate their joint retrieval during recall, increased semantic similarity between the right constituent and compound hindered performance. We argue that, unlike constituent-to-constituent comparisons, the constituent-to-compound comparison more accurately represents a measure of transparency or semantic contribution on the part of the constituent in question. As highly transparent compounds are argued to more heavily rely on constituent-level memory encoding (i.e., stored primarily as parts with internal structure, see: Fiorentino & Poeppel, 2007), we suggest that they are more strongly represented at the constituent level in working memory and are more susceptible to constituent interference as a result.

Importantly, expansion of the model to include additional lexical variables did not eliminate the observed differences in recall performance across the experimental conditions for compounds using the H HAS M and H MADE OF M conceptual relations. We propose that these effects are the result of interference induced by a same-relation environment during working memory processes and word recall, as explained above.

Conceptual Relation Similarity Survey

A criticism of previous relation priming research has been that the categorization of compound words by conceptual relation type is often subjective. That is, two people may paraphrase the same compound using different conceptual relations. For example, while someone might paraphrase the compound *sandbox* as "*a box that* HAS *sand*", another might say "*a box* FOR *sand*", "*sand* LOCATED IN *a box*", or another appropriate representation. Similar criticism has been made of frequent under-specificity in conceptual relation categories themselves, and their tendency to encompass a broad spectrum of meanings. A common example is the FOR relation proposed by Levi (1978) which can simultaneously apply to compounds like *sleeping pills* and *malaria pills*, despite conveying opposing meanings: sleeping pills are *pills* FOR ENABLING *sleep*, while malaria pills are *pills* FOR PREVENTING *malaria* (e.g., Downing, 1977; Spalding, 1991; Estes & Jones, 2006). Estes and Jones (2006) found that when participants' perceptions of relation similarity did not align with previous experimental designs, no effects of relation priming were observed. They concluded that investigations of relation priming should incorporate similarity norms to ensure that same- and different-relation primes are perceived to be similar and dissimilar, respectively.

To investigate the perceived similarity of the recall lists used in Experiment 2 of this thesis, participants were provided with a survey to rate the degree of relational similarity between pairs of words used in the experiment. Experiment 2 presented recall lists belonging to proposed same- and different-relation conditions: wherein compound words shared the same conceptual relation or different conceptual relations, respectively. If the experimental conditions were well formed, we expected compounds from recall lists in the SAME condition to have significantly higher similarity ratings than compounds from recall lists in the DIFFERENT condition.

Methods

Participants

The same participants from the Serial Recall Experiment completed the Conceptual Relation Similarity Survey during the same session and after completing the recall task. Data for one participant was not captured due to technical error (N = 38; F = 33; M = 5).

Materials

240 compound word pairs (e.g., *facemask | lifeguard*) were compiled from the 40 recall lists used in the serial recall experiment. Recall lists contained four compound

words each. Specifically, each recall list from the experiment was used to form six word-pairs so that all words in a list were paired once with all other words in that list (40 lists x 6 permutations per list = 240 word-pairs). The order within pairs was pseudo-randomized so that a word did not always appear in the same position. See Table 8 below for example and *Appendices* for the full list of word pairs.

Table 8: sample word pairs from relation similarity survey.

Recall List	Randomized Word Pairs
facemask	facemask / lifeguard
lifeguard	dustpan / facemask
dustpan	facemask / mousetrap
mousetrap	lifeguard / dustpan
	mousetrap / lifeguard
	dustpan / mousetrap

The 240 word-pairs were divided into two lists of 120 pairs each. Each participant was presented with one list, and each list was presented to an equal number of participants. Both lists were auto-randomized by the administering software, LimeSurvey (LimeSurvey GmbH, Hamburg, Germany), so that each participant received a unique order of presentation.

Procedure

The survey was conducted using the online software LimeSurvey. Participants completed the survey on-premises at the McMaster University *Language, Memory and Brain Laboratory*. Participants were allocated 40 minutes to complete the survey, though most finished between 20 and 30 minutes. Instructions were provided both orally and on screen, and participants were permitted to answer each question at their own pace. Prior to beginning the survey, participants entered a unique identifier to pair their responses with their recall data from the serial recall experiment.

The same question was present at the top of the screen for all trials: *To what degree are the conceptual relationships of the compounds below similar or dissimilar?* One of the word pairs was presented below at random, and participants selected a response from a seven-point Likert scale ranging from Highly Dissimilar (1) to Highly Similar (7). Intermediary points were not labelled. Participants clicked 'Next' to proceed through the questions and were permitted to revisit a question using a 'Back' button.

Results

A paired samples t-test was conducted to compare mean similarity ratings for recall lists in the SAME and DIFFERENT conditions, respectively (20 lists per condition). The test showed significantly higher similarity ratings for SAME (M=4.37, SD=0.61) compared to DIFFERENT (M=3.32, SD=0.30) conditions; t(19) = 6.64, $p < 0.001$. Survey results were adjusted (z-transformed) to account for individual variation in using the scale, but the transformed scores yielded similar results: SAME (M = 0.19, SD = 0.34), DIFFERENT (M = −0.31, SD = 0.12); t(19) = 5.95, $p < 0.001$. These findings indicate that participants perceived relations in the SAME condition to be significantly more similar than relations in the DIFFERENT condition overall. That is, the conceptual relationships of compounds classified as the same (e.g., haystack / tinfoil = H MADE OF M) were found to be more alike than the conceptual relationships of compounds classified as different (e.g., cornmeal / sandbox = H MADE OF M and H HAS M, respectively). These findings are visualized below in Figure 3. As a more accurate measure, normalized ratings are presented throughout the discussion.

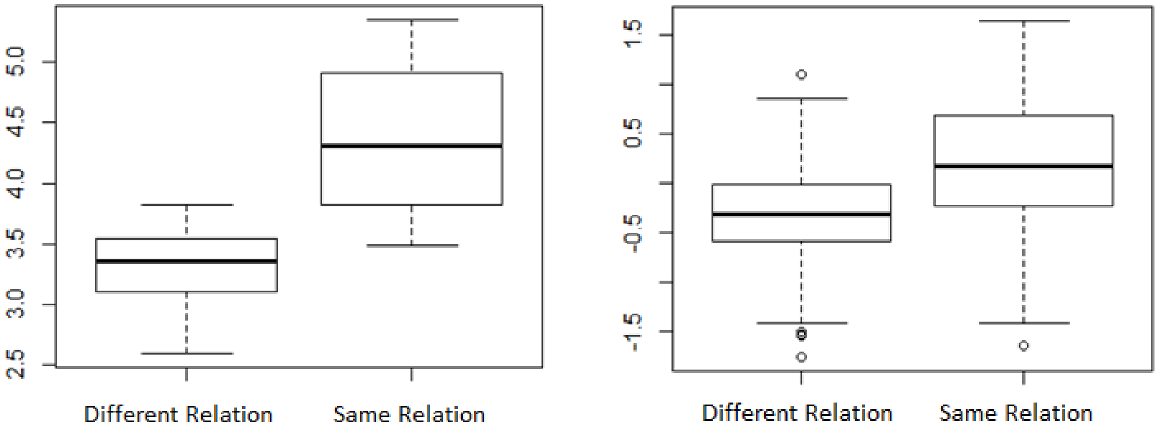

Figure 3: box plots of mean relation similarity scores as a function of condition. left: raw scores. right: normalized.

Relation Similarity by Relation (SAME condition)

A linear mixed effects regression model was run using the lme4 package for R (version 3.3.1, R Core Development Team, 2016) to investigate the effect of relationship category on similarity ratings of word pairs in the SAME relation environment. The model included participants and word pairs as random effects with the H FOR M relation set as a reference level. The model revealed a significant effect of relation category on SAME relation list similarity, indicating that word pairs representing the H HAS M relation had significantly lower similarity ratings (β = −0.303, SE = 0.11, z = −2.71, p = 0.022) and word pairs representing the H MADE OF M relation had significantly higher similarity ratings (β = 0.27, SE = 0.11, z = 2.48, p = 0.015) than the reference relation H FOR M.

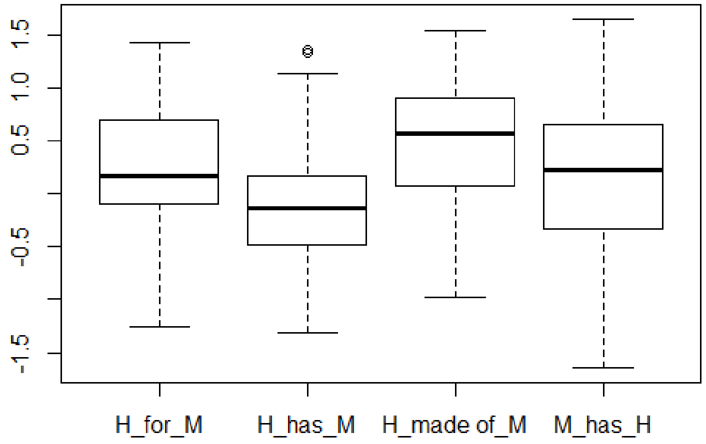

Figure 4: relation similarity rating by category (SAME condition)

Discussion

The survey results revealed that the relational similarity of SAME and DIFFERENT condition lists differed significantly in the manner intended (compound relations in SAME lists were perceived to be significantly more alike than compound relations in DIFFERENT lists) indicating that the SAME/DIFFERENT conditions were significantly contrastive. However, the difference between conditions was not as robust as reported in Estes & Jones (2006, Experiment 2).

Analysis of by-relation similarity in the SAME condition showed a significant effect of relation category on similarity rating, indicating that H HAS M and H MADE OF M words had significantly lower and significantly higher similarity ratings than the H FOR M relation, respectively. This may suggest that the accuracy of the attributed relation categories differs, or that participants' intuitions aligned with some attributed category memberships more readily than others. For example, participants appeared to agree that compounds identified as using the H MADE OF M relation had similar relationships (M = 0.48), while they did not agree that compounds identified as using the H HAS M relation had similar relationships (M = −0.12). This may suggest that compounds identified as using the H HAS M relation are incorrectly classified, or that the H HAS M relation is an overly general classification that encompasses distinct semantic relationships. Conversely, the higher similarity rating for H MADE OF M words suggests that words in this category have been more accurately classified, or that the MADE OF relationship is a more concise categorization that is less susceptible to individual interpretation.

Limitations of the Relation-Condition Interaction

Results of the immediate recall task show that performance across the four relation categories did not pattern uniformly. In particular, performance for M HAS H and H FOR M words did not show an interaction with similarity environment. We provide a possible interpretation of this observation below.

Examination of survey data suggests that M HAS H word-pair ratings exhibited a high degree of variability within and between participant responses. This suggests that participants' interpretations of compound relations in M HAS H lists did not align as readily as in other relation groupings. We interpret this to mean that participants do not readily agree with the categorization of M HAS H compounds, and that these lists are not relationally homogenous. As a result, priming effects would be unlikely in M HAS H lists due to insufficient relational similarity (per Estes & Jones, 2006). The fact that these lists exhibited positive relation similarity scores may have been the result of the compounds using semantic relationships that are related enough to be described using a common paraphrase, but that are ultimately too distinct to elicit a priming effect.

In the case of H FOR M word pairs, between-participant survey responses did not appear to differ significantly. We argue, however, that the general proclivity of the H FOR M paraphrase may have resulted in over generalization during similarity scoring. Analysis of the original Schmidtke et al. (2016) corpus showed that the H FOR M relation (one of just 16 possible relations) accounted for %35 of the 610 compounds categorized. We argue that this disproportionate representation is due to the overgeneralization of the FOR relation (as per Estes & Jones, 2016) and the tendency to use the *for* paraphrase to account for numerous semantic relationships. A tendency to compare prepositional paraphrase rather than relational meaning during the survey may have artificially inflated the similarity scores. As a result, these lists may not be relationally homogenous. We speculate that participants found these compounds to be similar to one another based on prepositional paraphrase, but processed them differently due to their distinct semantic relationships.

Error Analysis

Analyses of the serial recall experiment were conducted to further investigate the possible effects of conceptual relations and conceptual relation similarity on recall performance. Error analyses at the constituent and combinatorial level (recombining compound constituents in new forms) also provide more generalized insight into the processing of compound words in working memory.

Table 9 details the distribution of participant recall errors across eight general categories. Chi-square analysis of error distributions across experimental conditions (SAME and DIFFERENT) revealed no effect of experimental condition on error distribution: $X^2(6) = 0.81$; $p = 0.99$.

Table 9: high level errors by experimental condition.

Error	Description	SAME	DIFFERENT	Total
Omission (“blank”)	Participant did not provide response, or said “blank”.	825	814	1639
Incorrect Target	*Recalled item was a compound from the experiment but not the target.*			
Previous List	Recalled compound was a target compound in the previous list.	41	37	78
Order Error	Recalled compound was a target in the current list, but was recalled in the incorrect position.	206	204	410
Experiment Compound	Recalled compound was a compound from the experiment, but not from the current or previous list.	35	31	66
Non-Item	*Recalled item was not a compound from the experiment.*			
Modifier Error	Modifier recalled incorrectly, head recalled correctly.	47	42	89
Head Error	Modifier recalled correctly, head recalled incorrectly.	82	87	169
Joint Error	Modifier and head both incorrect.	88	84	172
Total		1324	1299	2623

Further analysis was conducted on error distributions across conceptual relationship category (see Table 10). Chi-square analysis showed a significant difference in distributions across the four relation types, $X^2(18) = 44.38$; $p < 0.001$. Examination of cell residuals indicated that these differences were the result of: significantly fewer previous list errors (recall of a compound from the previous list) than predicted when the target word was a compound using the H HAS M relation; significantly more modifier errors than predicted when the target word was a compound using the H HAS M relation; and significantly fewer modifier errors than predicted when the target word was a compound using the H FOR M or M HAS H relation.

Table 10: high level errors by conceptual relation.

Error	H FOR M	H HAS M	H MADE OF M	M HAS H	Total
Omission ("blank")	413	426	389	411	1639
Previous Compound	21	9	23	25	78
Order Error	94	94	107	115	410
Experiment Compound	13	23	12	18	66
Modifier Error	12	40	25	12	89
Head Error	48	41	46	34	169
Joint Error	39	42	50	41	172
Total	640	675	652	656	2623

Table 11 below details the distribution of recombination errors within the four possible recombination categories and across experimental conditions. Recombination errors occur when participants combine two existing constituents to form an untargeted compound. For example, a participant may see the words *snowman* and *cornfield* then recombine constituents to form *snowfield*. Chi-square analysis of error distributions (omitting Previous List Recombinations, as N < 5) across experimental conditions (SAME and DIFFERENT) revealed no effect of condition on error distribution: $X^2(2) = 1.76$; $p = 0.414$.

Table 11: recombination errors by experimental condition.

Error	Description	SAME	DIFFERENT	Total
Within-List Recombination	Constituents from the target list recombined to form an untargeted compound.	36	31	67
Between-List Recombination	One constituent from the target list recombined with a constituent from the previous list to form an untargeted compound.	23	32	55

Error	Description	SAME	DIFFERENT	Total
Previous List Recombination	Two constituents from the previous list (not present in the current list) recombined to form an untargeted compound.	0	1	1
Experiment Recombination	Two constituents present in the experiment recombined to form an untargeted compound that does not classify as within-list, between-list, nor previous list recombination.	34	34	68
Total		93	98	191

Table 12 below details the distribution of recombination errors across the four conceptual relation categories examined. Chi-square analysis of error distributions across relation category (omitting Previous List Recombinations, as N < 5) revealed no effect of relation on error distribution: $X^2(6) = 7.76$; $p = 0.257$.

Table 12: recombination errors by conceptual relation.

Error	H FOR M	H HAS M	H MADE OF M	M HAS H	Total
Within-List Recombination	23	15	12	17	67
Between-List Recombination	11	16	18	10	55
Previous List Recombination	0	0	0	1	1
Experiment Recombination	14	21	17	16	68
Total	48	52	47	44	191

The following analysis examined all constituent errors, including recombinations where a constituent from the target word recombined with a non-target constituent to form a word that was present in the experiment, as well as recombinations that formed words not present in the experiment. Individual cell contributions in the contingency table reveal that participants made approximately 80% more head errors ($n = 186$) than modifier errors ($n = 103$). Chi-square analysis of error distributions across experimental conditions (SAME and DIFFERENT) revealed no effect of condition on error distribution: $X^2(1) = 0.07$; $p = 0.79$.

Table 13: general constituent errors by experimental condition.

Error	Description	SAME	DIFFERENT	Total
Modifier Error	Modifier recalled incorrectly, head recalled correctly.	54	49	103

Error	Description	SAME	DIFFERENT	Total
Head Error	Modifier recalled correctly, head recalled incorrectly.	93	93	186
Totals		147	142	289

Chi-square analysis of constituent error distributions across conceptual relation category showed a significant difference between relation types: $X^2(3) = 16.36$; $p < 0.001$. Examination of cell residuals indicates that when attempting to recall compounds with the H HAS M relation, participants were significantly more likely to produce a modifier error than predicted.

Table 14: general constituent errors by conceptual relation.

Error	H FOR M	H HAS M	H MADE OF M	M HAS H	Total
Modifier Error	15	46	28	14	103
Head Error	48	45	48	45	186
Totals	63	91	76	59	289

Regression Analysis: errors by lexical variables

To examine the possible effects of lexical factors on recall error distributions, generalized linear mixed effects models were fitted to include the lexical variables identified under the recall experiment: left and right constituent frequencies, left and right constituent positional family size, and Latent Semantic Analysis (LSA) scores. Participants and recall words were included as random effects. Models were fitted for each error type individually.

Omissions. Results of the model revealed a significant main effect of compound frequency on Omissions ($\beta = -0.10$, $SE = 0.05$, $z = -2.10$, $p = 0.036$) indicating that lower compound frequency predicted a higher likelihood of a participant omitting a response or saying "blank". A significant main effect of compound length (measured in phonemes) was also found ($\beta = 0.10$, $SE = 0.04$, $z = 2.30$, $p = 0.021$) and indicated that higher phoneme count predicated a higher likelihood of omission.

Previous Compound & Order Errors. No main effects of lexical variables were observed for Previous Compound Errors nor Order Errors.

Modifier Errors. A significant effect of the Left-Right LSA score was observed ($\beta = -0.48$, $SE = 0.19$, $z = -2.49$, $p = 0.013$) and indicated that increased semantic similarity between the target compound's left and right constituents predicted lower likelihood of modifier error.

Head Errors. Results of the model revealed a significant main effect of compound frequency on Head Errors ($\beta = -0.41$, $SE = 0.16$, $z = -2.67$, $p = 0.008$) and indicated that lower compound frequency predicted increased likelihood of making a head error. A significant effect of left constituent frequency was also observed (β = 0.42, *SE* = 0.17, *z* = 2.49, *p* = 0.013) and indicated that increased left constituent (modifier) frequency predicted higher likelihood of making a head error. Further, a significant effect of left family size (the sum of compounds containing the same left constituent) was observed (β = 0.36, *SE* = 0.15, *z* = 2.45, *p* = 0.015) and indicated that increased left family size predicted a higher likelihood of making a head error.

Within-List Recombinations: Results of the model revealed a significant main effect of left family size on within-list recombinations (β = 0.39; *SE* = 0.16; *z* = 2.48; *p* = 0.013) and indicated that higher left family size of the target word predicted an increased likelihood of within-list recombination. Further, a significant main effect of left-right semantic similarity was observed (β = −0.46; *SE* = 0.22; *z* = −2.08; *p* = 0.038) and indicated that higher left-right similarity of the target word predicted a decreased likelihood of within-list recombination.

Between-List Recombinations: Results of the model revealed a significant main effect of compound frequency on between-list recombinations (β = −0.38; *SE* = 0.18; *z* = −2.21; *p* = 0.027) and indicated that higher frequency of the target compound predicted a decreased likelihood of between-list recombination. A significant main effect of right constituent frequency was observed (β = −0.54; *SE* = 0.23; *z* = −2.30; *p* = 0.021) and indicated that higher right constituent frequency of the target compound predicted a decreased likelihood of between-list recombination. A significant main effect of compound length in phonemes was observed (β = 0.45; *SE* = 0.16; *z* = 2.79; *p* = 0.005) and indicated that longer phoneme count of the target compound predicted an increased likelihood of between-list recombination. A significant main effect of left-right constituent semantic similarity was observed (β = −0.52; *SE* = 1.19; *z* = 2.79; *p* = 0.005) and indicated that higher left-right similarity of the target compound predicted decreased likelihood of between-list recombination. A significant main effect of right-compound semantic similarity was also observed (β = 0.47; *SE* = 0.18; *z* = 2.62; *p* = 0.009) and indicated that higher right-compound similarity of the target compound predicted an increased likelihood of between-list recombination. Finally, an effect approaching significance of left-compound semantic similarity was observed (β = 0.31; *SE* = 0.16; *z* = 1.92; *p* = 0.055) and suggested that increased semantic similarity between the left constituent and compound predicted an increased likelihood of between-list recombination error.

Experiment Recombinations: Experiment recombinations are those involving the recombination of two constituents present in the experiment, but not classified under the other three recombination categories. Results of the model revealed a significant main effect of compound frequency on experiment constituent recombinations (β = −0.34; *SE* = 0.16; *z* = 2.18; *p* = 0.029) and indicated that higher frequency of the target compound predicted a decreased likelihood of experiment recombination. Further, a significant main effect of right-compound semantic similarity was observed

($\beta = 0.31$; $SE = 0.15$; $z = 2.14$; $p = 0.033$) and indicated that higher right-compound similarity of the target compound predicted an increased likelihood of experiment constituent recombination.

Discussion

Effects of Environment and Lexical Variables on Recall Errors

No main effect of experimental condition was found for error distributions, although observed counts show an approximately 40% higher occurrence of between-list recombinations in the different-relation condition ($n = 32$) relative to the same-relation condition ($n = 23$). This may have been due to the fact that the experiment did not control for the even distribution of different-relation lists, which sometimes resulted in the consecutive presentation different-relation lists. Such consecutive presentation would open the possibility for between-list compounds using the same relation to recombine. Although this was not accounted for in the analysis, it may explain the apparent disparity between conditions.

Omissions. Significant main effects of compound frequency and compound length (measured in phonemes) were found for general recall omissions (i.e., a participant failing to provide a response or saying "blank"). These results indicated that lower compound frequency and longer phonetic length predicted failure to recall a target correctly. These findings align with previous research on the immediate serial recall of both simplex and complex words (e.g., Engle, Nations & Cantor, 1990; Gregg, Freedman & Smith; 1989; Baddeley & Hitch, 1974).

Modifier Errors. A significant main effect of the left-right constituent semantic similarity showed that increased semantic similarity predicted lower likelihood of modifier error. As with the similar effect of left-right similarity on general recall performance, we argue that higher levels of co-occurrence for two constituents in the same semantic context results in increased co-activation and likelihood that two constituents will be recalled together.

Head Errors. Analysis of head errors suggested that higher compound frequency protected participants from making head errors, likely due in part to the strong semantic association between the head constituent and compound in transparent compound words (Libben et al., 2003). Significant main effects of left constituent frequency and left family size (the sum of compounds containing the same left constituent) found that increased left constituent frequency and family size predicted increased likelihood of making a head error. Large left constituent frequencies and family sizes both result in greater activation of the morphological family network in the mental lexicon (Kuperman et al., 2009). As participants recall the first constituent of a compound word, left frequency and family size effects spur the activation of the morphological family and increase competition for the head constituent. The increase in possible candidates for the head position may lead to increased interference and confusion, resulting in more head errors. This observation aligns with the findings of Wälchli (2016).

Effect of lexical variables on recombination errors

Within-List Recombinations. Results indicated that higher left family size of the target word predicted increased likelihood of within-list recombination. As indicated above, higher left family size may cause an increase in competing head constituents through spreading activations. Increased activation of head candidates within a given list reduces a participant's ability to discern the correct match and increases the likelihood of recalling an incorrect head. As with modifier errors, increased left-right constituent cooccurrence helped prevent within-list recombinations through heightened probability of recalling the modifier and head together.

Between-List Recombinations. Results suggest that multiple factors influence the probability of between-list recombinations, but that these factors do group together under similar patterns. The likelihood of a participant producing a between-list recombination was typically reduced by factors commonly associated with improved word recall. Higher compound (surface) frequency, head frequency, and left-right cooccurrence all lowered the occurrence of between-list recombination by increasing the likelihood of recalling the target compound. Similarly, shorter compound length also facilitated recall and inhibited between-list recombinations. Increased semantic similarity between left constituent and compound, and right constituent and compound (herein considered measures of compound transparency) increased the likelihood of between-list recombinations. We argue that high transparency compounds are more likely to undergo decomposition and be represented at the constituent level in memory. As a result, they become more susceptible to constituent interference and erroneous recombinations.

Experiment Recombinations: Recombinations of two constituents present in the experiment (but not from the current or previous lists) showed an inhibitory effect of compound frequency which decreased the likelihood of recombination. A further effect of right-compound similarity predicted increased likelihood of experimental recombination, perhaps because increased transparency suggests an increased likelihood of decomposition and possibility for constituent competition.

Errors by Conceptual Relation Category

Results showed significantly fewer previous compound errors (recall of a compound from the previous list) and significantly more modifier errors than predicted when the target word was a compound using the H HAS M relation. Currently, findings regarding previous compound errors—in the absence of shared constituents—are not readily interpretable and may be due primarily to sample size or statistical chance. Alternatively, previous compound errors may be influenced by relative compound frequencies across lists; comparisons of which have not been calculated for this analysis. The effect on modifier errors may be tied to the relation's inherent description of *possession*, and the plausibility of the head possessing a variety of attributes or components; however, in light of the observation that H HAS M categorization appears disputed by participants, error effects attributed to this relation are generally difficult to interpret at this time.

Results also indicated that participants made significantly fewer modifier errors than predicted when the target word was a compound using the H FOR M or M HAS H relation. Analysis of modifier frequency revealed that the mean frequencies of H FOR M and M HAS H compound modifiers were significantly higher than those of H HAS M and H MADE OF M compounds. Higher frequency would facilitate modifier recall and reduce the likelihood of making modifier errors for compounds bound by these relations.

General Discussion

Results of the conceptual relation similarity survey revealed that participants perceived the relations of list items in the same-relation condition to be significantly more similar (M = 0.19) than those of list items in the different-relation condition (M = −0.31). For example, the conceptual relationships of the same-relation pair *meatball* (ball MADE OF meat) and *snowfield* (field MADE OF snow) were found to be more alike (M = 0.80) than the conceptual relationships of the different-relation pair *meatball* (ball MADE OF meat) and *gunboat* (boat HAS gun) (M = −0.79). These results indicated that the experimental conditions (SAME/DIFFERENT) of the serial recall task established significantly differing relational similarity environments for item memorization and recall. As a result, this contrast provided opportunity to observe possible effects of conceptual relation similarity on recall performance.

Results of the serial recall task revealed an inhibitory effect of relation similarity environment on recall performance; however, this effect was only observed for half of the relation categories used in the study. Specifically, compounds bound by H HAS M and H MADE OF M conceptual relations were significantly less likely to be recalled correctly in a same-relation environment. No effect of relation similarity on recall performance was observed for the remaining relations, H FOR M and M HAS H. We argue that the absence of an effect for these latter relations is the result of an experimental confound: compounds attributed to these categories were not relationally homogenous and, therefore, were not sufficiently similar to form a same-relation environment. First, the H FOR M relation was found to be disproportionally represented in the original Schmidtke et al. (2016) corpus, accounting for 35% of all relationally classified compounds. We propose that this is likely due to the over-generalization of the H FOR M paraphrase. Although the greatest number of compounds were categorized under this paraphrase, they constitute a grouping of multiple discrete semantic relations (for example, the *promote/inhibit* distinction between *sleeping pills* and *malaria pills*). This finding aligns with the criticism that the Levi (1978) taxonomy is overly general and does not sufficiently capture distinct semantic relationships (e.g., Spalding, 1991; Estes & Jones, 2006). Similarly, evidence from the relational similarity survey suggests that participants agreed the least on the sameness of word pairs said to be bound by the M HAS H relation. Again, this indicates that lists identified as using the M HAS H relation were not relationally homogenous, and likely included multiple discrete semantic relationships grouped under a common paraphrase.

For compounds bound by H HAS M and H MADE OF M conceptual relations, recall performance was significantly lower in the same-relation environment. We propose that this effect is the result of increased competition between compound constituents during word recall. As a relation becomes increasingly activated through repeated exposure in a same-relation environment, so too do combinatorial constituents that are commonly associated with that relation. As the number of activated constituents increases across the relation's semantic network, participants must parse through an increasing number of candidates competing for selection at recall. Simultaneously, the activations of these constituents are strengthened through relational similarity, thereby reducing the gradience in activation levels that is normally used as a recall cue. Evidence from the current study and Wälchli (2016) suggests that the modifier is retrieved first and acts as a recall cue for the head. The recall of the first constituent is likely facilitated by phonological traces, while the recall of the second constituent appears to be dependent on a more complex process related to the reconstruction of the compound word. While the first constituent is recalled phonologically, the second constituent must be selected on the basis of both phonological trace and semantic match. The normalizing of constituent activations across the semantic network makes this head selection process more difficult. In sum, we argue that, although participants are able to recall first constituents based on a phonological trace alone, the recovery of a second constituent is dependent on a more robust process that is prone to failure in an environment that promotes competition in constituent selection. This theory is supported by the finding that participants made nearly twice as many head errors as modifier errors.

The findings of relational similarity in the present study appear to differ from observed relation priming effects in lexical decision and sensicality judgement tasks (e.g., Gagné & Shoben, 1997; Spalding & Gagné, 2011). Under the RICE model (Spalding et al., 2010), relation priming effects are the result of competitive inhibition from different-relation primes. When a target compound is primed with a compound bound by a different relation, the prime relation is activated and competes for selection during the relation processing of the target compound. In this way, the observed effects of relation priming are thought to be the result of competition between relations. In the present study, the presence of a same-relation appears to inhibit compound word recall in an immediate serial recall task. We argue that this effect is the result of competition at the constituent word level, as words that are strongly associated with a given conceptual relation become activated and compete for selection during recall. We suggest that the effect of relational similarity is task specific, and that in working memory, the concurrent activation of numerous constituent-level representations associated with a common relation results in an environment that is prone to competitive interference.

Analysis of lexical variables on recall performance and error occurrences revealed two key findings in line with Wälchli (2016). First, whole-word or surface properties that have previously shown effects on the recall of simplex and complex words pattern similarly in this study of compound words. Higher compound frequency and shorter compound length (in phonemes) predicted higher recall performance, which aligns with prior research of simplex, derived, and inflect words

(Poirier & Saint-Aubin, 1996; Engle, Nations & Cantor, 1990; Tehan & Humphreys, 1988). Similarly, measures of left-right constituent cooccurrence also accurately predicted recall performance, with higher levels of cooccurrence predicting higher recall. Second, factors associated with compound decomposition and individual constituent representations also appeared to affect recall performance, but in the opposite direction. Compounds with higher transparency measures were found to have lower recall performance. As highly transparent compounds are argued to more heavily rely on constituent-level memory encoding (i.e., stored primarily as parts with internal structure, see: Fiorentino & Poeppel, 2007), we suggest that they are more strongly represented at the constituent level in working memory and are more susceptible to constituent interference as a result. Similarly, higher left constituent frequency and left constituent family size were also predictors of lower recall performance. These measures are associated with heightened activation of related constituents across the semantic network and are proposed to increase competition between activated representations. These findings suggest that the more strongly a memory word is encoded as discrete constituents (rather than as a whole word) the more likely it is for that word to be forgotten or recalled incorrectly as a result of increased constituent level competition during word recall.

Limitations and Future Directions

Further Considerations for the Different-Relation Environment

The possible interactions between compound constituents and conceptual relations in the different-relation environment of the recall task are many and difficult to accurately control. In the different-relation environment, it was presumed that each represented conceptual relation had equal activation: i.e., the inclusion of four compounds bound by four distinct relations would result in four equally activated relations; however, individual constituents may lend themselves to various relations with differing levels of ease and proclivity. For example, the word *copper* likely selects for (and suggests) a relatively narrow range of conceptual relations, the most prominent being the MADE OF relation, as copper is commonly recognized as a composition material. A word like *back*, however, may select from and enter a wider variety of conceptual relations, such as *paperback* (H MADE OF M), *backache* (M HAS H), *hunchback* (H HAS M), etc. As a result, some compound constituents may be more susceptible to interference from the concurrent activation of multiple relations than others, while simultaneously contributing to the activation of multiple relations. Together, this could result in increased competition for a subset of constituents in a different-relation list and/or the disproportionate activation of a subset of conceptual relations. Such effects were not controlled in the present experiment and should be considered in future explorations of similar designs.

Relational Similarity Survey Design and Interpretation

Due to time constraints, results of the relational similarity survey were incorporated as a post-hoc analysis to validate the experimental conditions of the serial recall experiment. While the results of the survey indicated that the two similarity environments (same/different) were significantly contrastive, it is clear that participants do not agree with the categorization of all compounds. Future studies would benefit from developing recall stimulus lists on the basis of similarity norms, preferably obtained from the target participant pool. In addition, surveyed participants should be clearly instructed to avoid comparing prepositional paraphrase and to focus on the distinct semantic relationship encoded between the to constituents of a compound.

Conclusions

Evidence from the reported serial recall experiment suggests an effect of compound relation priming / relational similarity in working memory, wherein relational similarity between recall list items appears to inhibit recall performance. This effect was not observed in relations that appear to be overly generalized, suggesting that the effect is only present when compound words are matched with salient, sufficiently specified relations. The effect was most robust for compounds bound by the H MADE OF M relation, which received the highest relational similarity rating in a follow-up participant survey. This suggests that the effect observed for H MADE OF M compounds may represent relation similarity effects in working memory most accurately.

The observed relation similarity effects do not appear to pattern with other semantic similarity effects in studies of serial recall, which have been shown to facilitate recall performance (e.g., Baddeley, 1966; Crowder, 1979; Acheson et al., 2011; Poirier et al., 2015). We suggest that this is due to the strong constituent-level interactions conceptual relations instantiate. While categorical similarity appears to facilitate recall in simplex words, this is thought to be due to increased levels of activation at the whole-word level (e.g., Poirier et al., 2015). In the case of conceptual relations, similarity results in increased activation at the constituent level, which introduces interference in compound recombination during recall. This finding suggests, more generally, that interactions between constituent-level representations and constituent-level semantic and lexical properties are likely to produce different patterns of recall performance than observed in simplex words. This is in line with similar observations regarding the distinction between simplex and complex word visual processing and the contributions of constituent level lexical and semantic factors (e.g., Juhasz, 2006; Pollatsek, Hyönä, & Bertram, 2000; Andrews, Miller, & Rayner, 2004).

The reported findings suggest that the effects of relation priming / relational similarity are task specific, and interact with working memory processes differently from those involved in previous studies of relation priming. While studies of relation priming in lexical decision and sensicality judgement tasks (e.g., Gagné, 2001; 2002; Estes & Jones, 2006) found that different-relation primes inhibit compound word processing (Spalding & Gagné, 2011), the current study observed opposite effects. The immediate recall task differs from more generalized processing tasks (e.g., lexical decision, sensicality judgements) in that there is an increased representational load associated with the simultaneous storage and retrieval of multiple stimuli. In working memory tasks, the concurrent activation of numerous constituent-level representations results in an environment that is especially prone to competitive interference. In the relation priming studies reported by Gagné (e.g., 2001; 2002) and Spalding and Gagné (2011), competition effects are proposed to be the result of concurrent relational interpretations which must be rejected through reanalysis of the presented stimuli. Such tasks do not introduce the constituent-level competition observed in working memory.

Analysis of lexical variables on recall performance and error production also yielded interesting results toward the general study of compound words in working memory. Specifically, evidence suggests that lexical and semantic factors that operate on a whole-word level (e.g., frequency, length, constituent co-occurrence) appear to facilitate the recall of compound words, while lexical and semantic factors that emphasize constituent-level representations (e.g., transparency, left constituent frequency, left family size) appear to hinder recall performance. This would suggest that a key factor in successfully recalling compound words is the ability to recall both constituents together, and factors that reflect heightened association between these constituents predict increased recall performance. Conversely, factors that reflect discrete constituent representations (e.g., transparency) or constituent competition (e.g., family size) predict lower recall performance. Together, this suggests that compound words are primarily represented as individual constituents in working memory, and that the strength of association between a compound's constituent parts is a critical predictor of its recall probability.

Although not a primary investigation of this study, results are in line with interference theories of working memory (Oberauer et al., 2016; Oberauer et al. 2012; Oberauer & Kliegl, 2006; Saito & Miyake, 2004) in that normalized activation levels across compound constituents appears to decrease gradience used as a recall cue and increase competition between constituents during recall. However, competition effects may be explained more generally as increasing the time required for processing during compound recombination, thereby allowing for time-based decay to erode phonological traces (Baddeley & Hitch, 1974; Baddeley, 1983; 2000; 2012). This would be especially true for the recollection of head constituents, and would also explain the observation that head errors were significantly more common than modifier errors. In sum, observations from the present study are equally attributable to both interference and phonological models of verbal short-term memory.

References

Abboud, H., Schultz, H., & Zeitlin, V. (2008). SuperLab 5.0 [computer software]. San Pedro, CA: Cedrus Corporation.

Acheson, D. J., MacDonald, M. C. (2009). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin, 135*, 50–68.

Acheson, D. J., MacDonald, M. C., & Postle, B. R. (2011). The Effect of Concurrent Semantic Categorization on Delayed Serial Recall. J*ournal of Experimental Psychology: Learning, Memory, and Cognition, 37(1)*, 44–59. http://doi.org/10.1037/a0021205

Amenta, S., & Crepaldi, D. (2012). Morphological processing as we know it: An analytical review of morphological effects in visual word identification. *Frontiers in Psychology, 3*, 1–12. http://doi.org/10.3389/fpsyg.2012.00232

Andrews, S., Miller, B., & Rayner, K. (2004). Eye Movements and Morphological Segmentation of Compound Words: There is a Mouse in Mousetrap. *European Journal of Cognitive Psychology, 16(1–2)*, 285–311. http://doi.org/10.1080/09541440340000123

Arcara, G., Semenza, C., & Bambini, V. (2014). Word structure and decomposition effects in reading. *Cognitive Neuropsychology, 31(1–2)*, 184–218. http://doi.org/10.1080/02643294.2014.903915

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language, 37(1)*, 94-117.

Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1995). The CELEX database. Nijmegen: Center for Lexical Information, Max Planck Institute for Psycholinguistics.

Baddeley, A. D., Hitch, G. J. (1974). Working memory. In G.A. Bower (Ed), *The Psychology of Learning and Motivation: Advances in Research and Theory*, 47–89. New York: Academic Press.

Baddeley, A. D., Lewis, V., Vallar, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology*, 36A, 233–52.

Baddeley, A. D. (1976). *The Psychology of Memory.* New York: Basic Books.

Baddeley, A. D. (1986). *Working Memory*. Oxford, UK: Oxford University Press.

Baddeley, A. D. (1983). Working memory. *Philosophical Transactions of the Royal Society of London, 302B*, 311-324.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4(11)*: 417-423.

Baddeley, A. D., & Ecob, J. R. (1970). Simultaneous acoustic and semantic coding in short-term memory. *Nature, 227(5255)*, 288-289.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word Length and the Structure of Short-Term Memory. *Journal of Verbal Learning and Verbal Behavior, 14(6)*, 575–589.

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic, and formal similarity. *Quarterly Journal of Experimental Psychology, 18A*, 362-365.

Baddeley, A. D. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology*, *63*(1), 1–29. http://doi.org/10.1146/annurev-psych-120710-100422

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition for single syllable words. *Journal of Experimental Psychology: General, 133*, 283–316.

Bertram, R., & Hyönä, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from eye movements when reading short and long Finnish compounds. *Journal of Memory and Language, 48*, 615–634.

Besner, D., & Davelaar, E. (1982). Basic processes in reading: Two phonological codes. *Canadian Journal of Psychology, 36*, 701-711.

Bourassa, D., & Besner, D. (1994). Beyond the articulatory loop: A semantic contribution to serial order recall of subspan lists. *Psychonomic Bulletin and Review, I*, 122-125.

Boutonnet, B., McClain, R., & Thierry, G. (2014). Compound words prompt arbitrary semantic associations in conceptual memory. *Frontiers in Psychology, 5*, 1–7. http://doi.org/10.3389/fpsyg.2014.00222

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41(4)*, 977-990.

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition, 28(3)*, 297-332.

Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology, 55*, 75-84.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake, & P. Shah (Eds.), *Models of Working Memory* (62–101). Cambridge, UK: Cambridge University Press.

Cowan, N., Day, L., Saults, J. S., Keller, T. A., Johnson, T. D., & Flores, L. (1992). The role of verbal output time in the effects of word length on immediate memory. *Journal of Memory and Language, 31*, 1-17.

Crowder, R.G. (1978). Memory for phonologically uniform lists. *Journal of Verbal Learning and Verbal Behaviour, 17*, 73-89.

Crowder, R.G. (1979). Similarity and order in memory. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory,* Vol. 13. New York: Academic Press.

De Jong IV, N. H., Schreuder, R., & Harald Baayen, R. (2000). The morphological family size effect and morphology. *Language and cognitive processes, 15(4-5)*, 329-365.

Downing, P. (1977). On the creation and use of English compound nouns. *Language, 53*, 810-842.

Duñabeitia, J. A., Perea, M., & Carreiras, M. (2007). The role of the frequency of constituents in compound words: evidence from Basque and Spanish. *Psychonomic Bulletin & Review, 14(6)*, 1171–1176. http://doi.org/10.3758/BF03193108

El-Bialy, R., Gagné, C. L., & Spalding, T. L. (2013). Processing of English compounds is sensitive to the constituents' semantic transparency. *The Mental Lexicon, 8(1)*, 75–95. http://doi.org/10.1075/ml.8.1.04elb

Engle, R.W., Nations, J.K., & Cantor, J. (1990). Is "working memory capacity" just another name for word knowledge? *Journal of Educational Psychology, 82*, 799-804.

Estes, W. K. (1973). Phonemic coding and rehearsal in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 12*, 360-372.

Estes, Z. (2003). A tale of two similarities: Comparison and integration in conceptual combination. *Cognitive Science, 27(6)*, 911–921. http://doi.org/10.1016/j.cogsci.2003.01.001

Estes, Z., & Jones, L. L. (2006). Priming via relational similarity: A copper horse is faster when seen through a glass eye. *Journal of Memory and Language, 55(1)*, 89–101. http://doi.org/10.1016/j.jml.2006.01.004

Fiorentino, R., & Poeppel, D. (2007). Compound words and structure in the lexicon. *Language and Cognitive Processes, 22(7)*, 953–1000. http://doi.org/10.1080/01690960701190215

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of verbal learning and verbal behavior, 12(6)*, 627-635.

Frisson, S., Niswander-Klement, E., & Pollatsek, A. (2008). The role of semantic transparency in the processing of English compound words. *British Journal of Psychology, 99(1)*, 87–107. http://doi.org/10.1348/000712607X181304

Fruchter, J., & Marantz, A. (2015). Decomposition, lookup, and recombination: MEG evidence for the full decomposition model of complex visual word recognition. *Brain and Language, 143*, 81–96.

Gagné, C. L. (2000). Relation-Based Combinations Versus Property-Based Combinations: A Test of the CARIN Theory and the Dual-Process Theory of Conceptual Combination. *Journal of Memory and Language, 42(3)*, 365–389. http://doi.org/10.1006/jmla.1999.2683

Gagné, C. L. (2001). Relation and lexical priming during the interpretation of noun–noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27(1)*, 236–254. http://doi.org/10.1037/0278-7393.27.1.236

Gagné, C. L. (2002). Lexical and relational influences on the processing of novel compounds. *Brain and Language, 81(1–3)*, 723–735. http://doi.org/10.1006/brln.2001.2559

Gagné, C. L., & Shoben, E. J. (1997). Influence of thematic relations on the comprehension of modifier–noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23(1)*, 71–87. http://doi.org/10.1037/0278-7393.23.1.71

Gagné, C. L., & Shoben, E. J. (2002). Priming relations in ambiguous noun-noun combinations. *Memory and Cognition, 30(4)*, 637–46. http://doi.org/10.3758/BF03194965

Gagné, C. L., & Spalding, T. L. (2004). Effect of discourse context and modifier relation frequency on conceptual combination. *Journal of Memory and Language, 50(4)*, 444–445. http://doi.org/10.1016/j.jml.2004.01.003

Gagné, C. L., & Spalding, T. L. (2004). Effect of relation availability on the interpretation and access of familiar noun-noun compounds. *Brain and Language, 90(1–3)*, 478–486. http://doi.org/10.1016/S0093-934X(03)00459-0

Gagné, C. L., & Spalding, T. L. (2009). Constituent integration during the processing of compound words: Does it involve the use of relational structures? *Journal of Memory and Language, 60(1)*, 20–35. http://doi.org/10.1016/j.jml.2008.07.003

Gagné, C. L., Spalding, T. L., & Ji, H. (2005). Re-examining evidence for the use of independent relational representations during conceptual combination. *Journal of Memory and Language, 53(3)*, 445–455. http://doi.org/10.1016/j.jml.2005.03.006

Gagné, C. L., Spalding, T. L., Figueredo, L., & Mullaly, A. C. (2009). Does snow man prime plastic snow? The effect of constituent position in using relational information during the interpretation of modifier-noun phrases. *The Mental Lexicon, 4(1)*, 41–76. http://doi.org/10.1075/ml.4.1.03gag

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language, 29(2)*, 228–244. http://doi.org/10.1016/0749-596X(90)90074-A

Gregg, V.H., Freedman, C.M., & Smith, D.K. (1989). Word frequency, articulatory suppression and memory span. *British Journal of Psychology, 80*, 363-374.

Campoy, G., Castellà, J., Provencio, V., Hitch, G. J., & Baddeley, A. D. (2014). Automatic semantic encoding in verbal short-term memory: Evidence from the concreteness effect. *The Quarterly Journal of Experimental Psychology*, *68(4)*, 759-78. http://doi.org/10.1080/17470218.2014.966248

Hitch, G. J. (1980). Developing the concept of working memory. In G. Claxton (Ed.), *Cognitive psychology: New directions* (154-196). London: Routledge & Kegan Paul.

Hongbo, J., & Gagné, C. L. (2007). Lexical and relational influences on the processing of Chinese modifier-noun compounds. *The Mental Lexicon, 2(3)*, 387–417.

Howard, M. W., & Kahana, M. J. (2002). When Does Semantic Similarity Help Episodic Retrieval? *Journal of Memory and Language, 46(1)*, 85–98. http://doi.org/10.1006/jmla.2001.2798

Hurlstone, M. J., Hitch, G. J., Baddeley, A. D., Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2013). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological Bulletin, 140(2)*, 339-373. http://doi.org/10.1037/a0034221

Huttenlocher, J., & Newcombe, N. (1976). Semantic effects on ordered recall. *Journal of Verbal Learning and Verbal Behavior, 15*, 387-399.

Hyönä, J., & Pollatsek, A. (1998). Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1612–1627. doi:10.1037/0096-1523.24.6.1612

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics, 40(6)*, 431–439. http://doi.org/10.3758/BF03208203

Ji, H., Gagné, C. L., & Spalding, T. L. (2011). Benefits and costs of lexical decomposition and semantic integration during the processing of transparent and opaque English compounds. *Journal of Memory and Language, 65(4)*, 406–430. http://doi.org/10.1016/j.jml.2011.07.003

Juhasz, B. J. (2007). The influence of semantic transparency on eye movements during English compound word recognition. In R. von Gompel, W. Murray, & M. Fischer (Eds.), *Eye movements: A window on mind and brain* (373–389). Boston: Elsevier

Juhasz, B. J. (2008). The processing of compound words in English: Effects of word length on eye movements during reading. *Language and Cognitive Processes, 23*, 1057–1088.

Juhasz, B. J., & Berkowitz, R. N. (2011). Effects of morphological families on English compound word recognition: A multitask investigation. *Language and Cognitive Processes, 26*, 653–682. http://doi.org/10. 1080/01690965.2010.498668

Juhasz, B. J., Starr, M. S., Inhoff, A. W., & Placke, L. (2003). The effects of morphology on the processing of compound words: Evidence from naming, lexical decisions and eye fixations. *British Journal of Psychology, 94(2)*, 223–244. http://doi.org/10.1348/000712603321661903

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99(1)*, 122–149. http://doi.org/10.1037/0033-295X.99.1.122

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Dartmouth Publishing Group.

Kuperman, V., Bertram, R., & Baayen, R. H. (2008). Morphological dynamics in compound processing. *Language and Cognitive Processes, 23(7–8)*, 1089–1132. http://doi.org/10.1080/01690960802193688

Kuperman, V., Schreuder, R., Bertram, R., & Baayen, R. H. (2009). Reading polymorphemic Dutch compounds: toward a multiple route model of lexical processing. *Journal of Experimental Psychology. Human Perception and Performance, 35(3)*, 876–895. http://doi.org/10.1037/a0013484

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104(2)*, 211–240.

Levi, J. N. (1978). *The syntax and semantics of complex nominals*. New York: Academic Press.

Levy, B. A. (1971). The role of articulation in auditory and visual short-term memory. *Journal of Verbal Learning and Verbal Behavior, 10*, 123-132.

Libben, G. (1993). Are morphological structures computed during word recognition? *Journal of Psycholinguistic Research, 22(5)*, 535–544. http://doi.org/10.1007/BF01068253

Libben, G. (1998). Semantic transparency in the processing of compounds: consequences for representation, processing, and impairment. *Brain and Language, 61(1)*, 30–44. http://doi.org/10.1006/brln.1997.1876

Libben, G., Gibson, M., Yoon, Y. B., & Sandra, D. (2003). Compound fracture: The role of semantic transparency and morphological headedness. *Brain and Language, 84(1)*, 50–64. http://doi.org/10.1016/S0093-934X(02)00520-5

LimeSurvey GmbH (2015). *LimeSurvey: An Open Source survey tool.* Hamburg, Germany. URL: http://www.limesurvey.org.

Maguire, P., & Cater, A. (2004). Is Conceptual Combination Influenced by Word Order? *The Companion Volume to the Proceedings of 42st Annual Meeting of the Association for Computational Linguistics*, 134–137.

Marelli, M., Zonca, G., Contardi, A., & Luzzatti, C. (2014). The representation of compound headedness in the mental lexicon: A picture naming study in aphasia. *Cognitive Neuropsychology, 31(1-2)*, 26–39. http://doi.org/10.1080/02643294.2013.860024

Murray, D.J. (1967). The role of speech responses in short-term memory. *Canadian Journal of Psychology, 21*, 263-276.

Nairne, J.S. (1990). A feature model of immediate memory. *Memory and Cognition, 18*, 251-269.

Nairne, J.S., & Neumann, C. (1993). Enhancing effects of similarity on long-term memory for order. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 329-337.

Nastase, V., Nakov, P., Séaghdha, D. Ó., & Szpakowicz, S. (2014). Semantic Relations Between Nominals. *Synthesis Lectures on Human Language Technologies, 6(1)*, 1–119. http://doi.org/10.2200/S00489ED1V01Y201303HLT019

Németh, D., Ivády, R. E., Guida, A., Miháltz, M., Peckham, D., Krajcsi, A., & Pléh, C. (2011). The effect of morphological complexity on short-term memory capacity. *Acta Linguistica Hungarica, 58(1-2)*, 85-107. http://doi.org/10.1556/ALing.58.2011.1-2.5

Nimmo, L. M., & Roodenrys, S. (2006). The influence of phoneme position overlap on the phonemic similarity effect in nonword recall. *Quarterly Journal of Experimental Psychology, 59(3)*, 577–96. http://doi.org/10.1080/02724980443000845

Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language, 55(4)*, 601-626.

Oberauer, K., Farrell, S., Jarrold, C., & Lewandowsky, S. (2016). What limits working memory capacity? *Psychological Bulletin, 142(7)*, 758–799. http://doi.org/10.1037/bul0000046

Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review, 19(5)*, 779–819. http://doi.org/10.3758/s13423-012-0272-4

Pham, H., & Baayen, R. H. (2013). Semantic relations and compound transparency: A regression study in CARIN theory. *Psihologija, 46(4)*, 455–478. http://doi.org/10.2298/PSI1304455P

Poirier, M., & Saint-Aubin, J. (1995). Memory for Related and Unrelated Words: Further Evidence on the Influence of Semantic Factors in Immediate Serial Recall. *The Quarterly Journal of Experimental Psychology Section A, 48(2)*, 384–404.

Poirier, M., & Saint-Aubin, J. (1995). Memory for Related and Unrelated Words: Further Evidence on the Influence of Semantic Factors in Immediate Serial Recall. *The Quarterly Journal of Experimental Psychology Section A, 48(2)*, 384–404. http://doi.org/10.1080/14640749508401396

Poirier, M., & Saint-Aubin, J. (1996). Immediate serial recall, word frequency, item identity and item position. *Canadian Journal of Experimental Psychology, 50*, 408–412.

Poirier, M., Saint-Aubin, J., Mair, A., Tehan, G., & Tolan, A. (2015). Order recall in verbal short-term memory: The role of semantic networks. *Memory and Cognition, 43*, 489–499. http://doi.org/10.3758/s13421-014-0470-6

Pollatsek, A., & Hyönä, J. (2005). The role of semantic transparency in the processing of Finnish compound words. *Language and Cognitive Processes, 20*, 261–290.

Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human Perception and Performance, 26(2)*, 820–833. http://doi.org/10.1037/0096-1523.26.2.820

Popov, V., & Hristova, P. (2015). Unintentional and efficient relational priming. *Memory and Cognition, 43(6)*, 866–878. http://doi.org/10.3758/s13421-015-0514-6

Pylkkänen, L., Feintuch, S., Hopkins, E., & Marantz, A. (2004). Neural correlates of the effects of morphological family frequency and family size: an MEG study. *Cognition, 91(3)*, B35-B45.

Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review, 11(6)*, 1090–1098. http://doi.org/10.3758/BF03196742

Reinitz, M. T., & Hannigan, S. L. (2004). False memories for compound words: Role of working memory. *Memory & Cognition, 32(3)*, 463–473. http://doi.org/10.3758/BF03195839

Roodenrys, S. J. (2009). Explaining phonological neighbourhood effects in short-term memory. In A. Thorn & M. Page (Eds.), I*nteractions between short-term and long-term memory in the verbal domain* (177–197). Hove: Psychology Press.

Rubin, Gary S; Becker, Curtis A; Freeman, R. H. (1979). Morphological Structure and Its Effect on Visual Word Recognition. *Journal of Verbal Learning and Verbal Behavior, 18(6)*, 757–767.

Saito, S., Miyake, A. (2004). On the nature of forgetting and the processing-storage relationship in reading span performance. *Journal of Memory and Language, 50*, 425–43.

Schmidtke, D., Kuperman, V., Gagné, C. L., & Spalding, T. L. (2016). Competition between conceptual relations affects compound recognition: the role of entropy. *Psychonomic Bulletin & Review, 23(2)*, 556–570. http://doi.org/10.3758/s13423-015-0926-0

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological pro- cessing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (131–154). Hillsdale, NJ: Erlbaum.

Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of memory and language, 37(1)*, 118-139.

Service, E., & Maury, S. (2015). Differential recall of derived and inflected word forms in working memory: examining the role of morphological information in simple and complex working memory tasks. *Frontiers in Human Neuroscience, 8*, 1–16. http://doi.org/10.3389/fnhum.2014.01064

Service, E., & Tujulin, A. M. (2002). Recall of morphologically complex forms is affected by memory task but not dyslexia. *Brain and Language, 81(1–3)*, 42–54. http://doi.org/10.1006/brln.2001.2505

Shoben, E. J. (1991). Predicating and Non-Predicating Combinations. In P. Schwanenflugel (Ed.), *The Psychology of Word Meanings* (117–135). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Spalding, T. L., & Gagné, C. L. (2011). Relation priming in established compounds: facilitation? *Memory & Cognition, 39(8)*, 1472–86. http://doi.org/10.3758/s13421-011-0112-1

Spalding, T. L., Gagné, C. L., Mullaly, A. C., & Ji, H. (2010). Relation-based interpretations of noun-noun phrases: A new theoretical approach. In S. Olson (ed.), *Linguistische Berichte Sonderheft 17: New impulses in word-formation* (283-315). Buske, Hamburg.

Tratz, S., & Hovy, E. (2010). A taxonomy, dataset, and classifier for automatic noun compound interpretation. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics* (678–687). Uppsala, Sweden.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *Quarterly Journal of Experimental Psychology, 57A*, 745–765.

Taft, M., & Forster, K. I. (1975). Lexical Storage and Retrieval of Prefixed Words. *Journal of Verbal Learning and Verbal Behavior, 14(6)*, 638–647.

Taft, M., & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior, 15(6)*, 607–620. http://doi.org/10.1016/0022-5371(76)90054-2

Tehan, G., & Humphreys, M.S. (1988). Articulatory loop explanations of memory span and pronunciation rate correspondences: A cautionary note. *Bulletin of the Psychonomic Society, 26*, 293-296.

Wälchli, Z. (2016). *Recall of compound words in simple and complex span tasks*. (Master's thesis). McMaster University, Hamilton, Ontario. Retrieved from http://hdl.handle.net/11375/20616.

Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25(5)*, 1256–1271. http://doi.org/10.1037/0278-7393.25.5.1256

Warren, B. (1978). *Semantic Patterns of Noun-Noun Compounds*. Göteborg: Universitatis Gothoburgensis.

Williams, E. (1981). On the notions" lexically related" and "head of a word". *Linguistic Inquiry, 12(2)*, 245-274.

Zwitserlood, P. (1994). The role of semantic transparency in the processing and representation of Dutch compounds. *Language and Cognitive Processes, 9(3)*, 341–368. http://doi.org/10.1080/01690969408402123

Appendices

Table 15: different relation recall lists

compound	relation	compound	relation	compound	relation
heartache	M HAS H	handrail	H FOR M	stoplight	H FOR M
facemask	H FOR M	floorplan	M HAS H	icepack	H MADE OF M
bunkhouse	H HAS M	paperback	H MADE OF M	mailman	H HAS M
applesauce	H MADE OF M	cornfield	H HAS M	shoelace	M HAS H
goldfish	H HAS M	honeycomb	H MADE OF M	toenail	M HAS H
scrapheap	H MADE OF M	songbook	H HAS M	beefsteak	H MADE OF M
earache	M HAS H	panhandle	M HAS H	alehouse	H HAS M
mousetrap	H FOR M	tradeshow	H FOR M	lifeguard	H FOR M
bloodlust	H FOR M	hailstorm	H HAS M	clipboard	H HAS M
doornail	M HAS H	pinhead	M HAS H	headband	H FOR M
speedboat	H HAS M	sweatband	H FOR M	bedpost	M HAS H
limestone	H MADE OF M	fruitcake	H MADE OF M	sandbank	H MADE OF M
milkshake	H MADE OF M	racetrack	H FOR M	mothball	H FOR M
barnyard	M HAS H	tinfoil	H MADE OF M	powerboat	H HAS M
nightgown	H FOR M	doorframe	M HAS H	doorknob	M HAS H
sidecar	H HAS M	newsboy	H HAS M	haystack	H MADE OF M
sandbox	H HAS M	paperclip	H FOR M	lunchroom	H FOR M
backache	M HAS H	stoneware	H MADE OF M	drugstore	H HAS M
cornmeal	H MADE OF M	weekday	M HAS H	brickwork	H MADE OF M
drumstick	H FOR M	hunchback	H HAS M	shoeshine	M HAS H
sandpaper	H MADE OF M	shellfish	H HAS M	gunboat	H HAS M
trainload	M HAS H	mealtime	H FOR M	firewood	H FOR M
bookshop	H HAS M	eyelash	M HAS H	meatball	H MADE OF M
dustpan	H FOR M	snowfield	H MADE OF M	kneecap	M HAS H
shipload	M HAS H	wolfpack	H MADE OF M		
woodwork	H MADE OF M	eggshell	M HAS H		
birdcage	H FOR M	footwear	H FOR M		
minefield	H HAS M	tollgate	H HAS M		

Table 16: same relation recall lists

compound	relation	compound	relation	compound	relation
drumstick	H FOR M	powerboat	H HAS M	cornmeal	H MADE OF M
tradeshow	H FOR M	mailman	H HAS M	brickwork	H MADE OF M
nightgown	H FOR M	tollgate	H HAS M	fruitcake	H MADE OF M
headband	H FOR M	bookshop	H HAS M	sandpaper	H MADE OF M
lunchroom	H FOR M	newsboy	H HAS M	stoneware	H MADE OF M
paperclip	H FOR M	clipboard	H HAS M	applesauce	H MADE OF M
firewood	H FOR M	songbook	H HAS M	tinfoil	H MADE OF M
handrail	H FOR M	minefield	H HAS M	paperback	H MADE OF M
mealtime	H FOR M	sandbox	H HAS M	woodwork	H MADE OF M
bloodlust	H FOR M	gunboat	H HAS M	milkshake	H MADE OF M
racetrack	H FOR M	shellfish	H HAS M	haystack	H MADE OF M
sweatband	H FOR M	bunkhouse	H HAS M	scrapheap	H MADE OF M
footwear	H FOR M	hunchback	H HAS M	wolfpack	H MADE OF M
mothball	H FOR M	sidecar	H HAS M	limestone	H MADE OF M
birdcage	H FOR M	drugstore	H HAS M	meatball	H MADE OF M
stoplight	H FOR M	cornfield	H HAS M	snowfield	H MADE OF M
facemask	H FOR M	hailstorm	H HAS M	icepack	H MADE OF M
lifeguard	H FOR M	alehouse	H HAS M	honeycomb	H MADE OF M
dustpan	H FOR M	speedboat	H HAS M	beefsteak	H MADE OF M
mousetrap	H FOR M	goldfish	H HAS M	sandbank	H MADE OF M
doorframe	M HAS H	bedpost	M HAS H	floorplan	M HAS H
heartache	M HAS H	shoelace	M HAS H	eyelash	M HAS H
toenail	M HAS H	doorknob	M HAS H	pinhead	M HAS H
shipload	M HAS H	kneecap	M HAS H	weekday	M HAS H
doornail	M HAS H	backache	M HAS H		
barnyard	M HAS H	eggshell	M HAS H		
earache	M HAS H	panhandle	M HAS H		
trainload	M HAS H	shoeshine	M HAS H		

Table 17: relational similarity survey word pairs

No.	Compound 1	Compound 2	Relation 1	Relation 2
1	heartache	facemask	M HAS H	H FOR M
2	heartache	bunkhouse	M HAS H	H HAS M
3	applesauce	heartache	H MADE OF M	M HAS H
4	bunkhouse	facemask	H HAS M	H FOR M
5	applesauce	facemask	H MADE OF M	H FOR M
6	applesauce	bunkhouse	H MADE OF M	H HAS M
7	brickwork	cornmeal	H MADE OF M	H MADE OF M
8	fruitcake	cornmeal	H MADE OF M	H MADE OF M
9	cornmeal	sandpaper	H MADE OF M	H MADE OF M
10	brickwork	fruitcake	H MADE OF M	H MADE OF M
11	brickwork	sandpaper	H MADE OF M	H MADE OF M
12	fruitcake	sandpaper	H MADE OF M	H MADE OF M
13	drumstick	tradeshow	H FOR M	H FOR M
14	drumstick	nightgown	H FOR M	H FOR M
15	headband	drumstick	H FOR M	H FOR M
16	nightgown	tradeshow	H FOR M	H FOR M
17	headband	tradeshow	H FOR M	H FOR M
18	headband	nightgown	H FOR M	H FOR M
19	scrapheap	goldfish	H MADE OF M	H HAS M
20	earache	goldfish	M HAS H	H HAS M
21	goldfish	mousetrap	H HAS M	H FOR M
22	scrapheap	earache	H MADE OF M	M HAS H
23	scrapheap	mousetrap	H MADE OF M	H FOR M
24	earache	mousetrap	M HAS H	H FOR M
25	stoneware	applesauce	H MADE OF M	H MADE OF M
26	stoneware	tinfoil	H MADE OF M	H MADE OF M
27	stoneware	paperback	H MADE OF M	H MADE OF M
28	applesauce	tinfoil	H MADE OF M	H MADE OF M
29	applesauce	paperback	H MADE OF M	H MADE OF M
30	tinfoil	paperback	H MADE OF M	H MADE OF M
31	bloodlust	doornail	H FOR M	M HAS H
32	speedboat	bloodlust	H HAS M	H FOR M
33	limestone	bloodlust	H MADE OF M	H FOR M
34	speedboat	doornail	H HAS M	M HAS H
35	limestone	doornail	H MADE OF M	M HAS H
36	limestone	speedboat	H MADE OF M	H HAS M

37	barnyard	milkshake	M HAS H	H MADE OF M
38	milkshake	nightgown	H MADE OF M	H FOR M
39	milkshake	sidecar	H MADE OF M	H HAS M
40	barnyard	nightgown	M HAS H	H FOR M
41	barnyard	sidecar	M HAS H	H HAS M
42	nightgown	sidecar	H FOR M	H HAS M
43	newsboy	clipboard	H HAS M	H HAS M
44	songbook	newsboy	H HAS M	H HAS M
45	minefield	newsboy	H HAS M	H HAS M
46	songbook	clipboard	H HAS M	H HAS M
47	minefield	clipboard	H HAS M	H HAS M
48	minefield	songbook	H HAS M	H HAS M
49	backache	sandbox	M HAS H	H HAS M
50	sandbox	cornmeal	H HAS M	H MADE OF M
51	sandbox	drumstick	H HAS M	H FOR M
52	backache	cornmeal	M HAS H	H MADE OF M
53	backache	drumstick	M HAS H	H FOR M
54	cornmeal	drumstick	H MADE OF M	H FOR M
55	woodwork	milkshake	H MADE OF M	H MADE OF M
56	haystack	woodwork	H MADE OF M	H MADE OF M
57	scrapheap	woodwork	H MADE OF M	H MADE OF M
58	haystack	milkshake	H MADE OF M	H MADE OF M
59	scrapheap	milkshake	H MADE OF M	H MADE OF M
60	scrapheap	haystack	H MADE OF M	H MADE OF M
61	icepack	stoplight	H MADE OF M	H FOR M
62	mailman	stoplight	H HAS M	H FOR M
63	stoplight	shoelace	H FOR M	M HAS H
64	icepack	mailman	H MADE OF M	H HAS M
65	shoelace	icepack	M HAS H	H MADE OF M
66	shoelace	mailman	M HAS H	H HAS M
67	paperclip	lunchroom	H FOR M	H FOR M
68	firewood	lunchroom	H FOR M	H FOR M
69	lunchroom	handrail	H FOR M	H FOR M
70	paperclip	firewood	H FOR M	H FOR M
71	handrail	paperclip	H FOR M	H FOR M
72	handrail	firewood	H FOR M	H FOR M
73	bloodlust	mealtime	H FOR M	H FOR M
74	racetrack	mealtime	H FOR M	H FOR M
75	mealtime	sweatband	H FOR M	H FOR M

76	bloodlust	racetrack	H FOR M	H FOR M
77	bloodlust	sweatband	H FOR M	H FOR M
78	racetrack	sweatband	H FOR M	H FOR M
79	toenail	beefsteak	M HAS H	H MADE OF M
80	alehouse	toenail	H HAS M	M HAS H
81	lifeguard	toenail	H FOR M	M HAS H
82	alehouse	beefsteak	H HAS M	H MADE OF M
83	beefsteak	lifeguard	H MADE OF M	H FOR M
84	alehouse	lifeguard	H HAS M	H FOR M
85	mothball	powerboat	H FOR M	H HAS M
86	doorknob	mothball	M HAS H	H FOR M
87	haystack	mothball	H MADE OF M	H FOR M
88	doorknob	powerboat	M HAS H	H HAS M
89	powerboat	haystack	H HAS M	H MADE OF M
90	doorknob	haystack	M HAS H	H MADE OF M
91	bedpost	shoelace	M HAS H	M HAS H
92	doorknob	bedpost	M HAS H	M HAS H
93	kneecap	bedpost	M HAS H	M HAS H
94	doorknob	shoelace	M HAS H	M HAS H
95	shoelace	kneecap	M HAS H	M HAS H
96	doorknob	kneecap	M HAS H	M HAS H
97	clipboard	headband	H HAS M	H FOR M
98	bedpost	clipboard	M HAS H	H HAS M
99	sandbank	clipboard	H MADE OF M	H HAS M
100	bedpost	headband	M HAS H	H FOR M
101	headband	sandbank	H FOR M	H MADE OF M
102	bedpost	sandbank	M HAS H	H MADE OF M
103	doornail	barnyard	M HAS H	M HAS H
104	earache	doornail	M HAS H	M HAS H
105	trainload	doornail	M HAS H	M HAS H
106	earache	barnyard	M HAS H	M HAS H
107	barnyard	trainload	M HAS H	M HAS H
108	earache	trainload	M HAS H	M HAS H
109	lunchroom	drugstore	H FOR M	H HAS M
110	brickwork	lunchroom	H MADE OF M	H FOR M
111	shoeshine	lunchroom	M HAS H	H FOR M
112	brickwork	drugstore	H MADE OF M	H HAS M
113	drugstore	shoeshine	H HAS M	M HAS H
114	brickwork	shoeshine	H MADE OF M	M HAS H

115	handrail	floorplan	H FOR M	M HAS H
116	paperback	handrail	H MADE OF M	H FOR M
117	cornfield	handrail	H HAS M	H FOR M
118	paperback	floorplan	H MADE OF M	M HAS H
119	floorplan	cornfield	M HAS H	H HAS M
120	cornfield	paperback	H HAS M	H MADE OF M
121	gunboat	sandbox	H HAS M	H HAS M
122	shellfish	sandbox	H HAS M	H HAS M
123	sandbox	bunkhouse	H HAS M	H HAS M
124	gunboat	shellfish	H HAS M	H HAS M
125	gunboat	bunkhouse	H HAS M	H HAS M
126	shellfish	bunkhouse	H HAS M	H HAS M
127	doorframe	heartache	M HAS H	M HAS H
128	toenail	doorframe	M HAS H	M HAS H
129	shipload	doorframe	M HAS H	M HAS H
130	toenail	heartache	M HAS H	M HAS H
131	heartache	shipload	M HAS H	M HAS H
132	toenail	shipload	M HAS H	M HAS H
133	honeycomb	songbook	H MADE OF M	H HAS M
134	panhandle	honeycomb	M HAS H	H MADE OF M
135	tradeshow	honeycomb	H FOR M	H MADE OF M
136	panhandle	songbook	M HAS H	H HAS M
137	songbook	tradeshow	H HAS M	H FOR M
138	panhandle	tradeshow	M HAS H	H FOR M
139	powerboat	mailman	H HAS M	H HAS M
140	tollgate	powerboat	H HAS M	H HAS M
141	bookshop	powerboat	H HAS M	H HAS M
142	tollgate	mailman	H HAS M	H HAS M
143	mailman	bookshop	H HAS M	H HAS M
144	tollgate	bookshop	H HAS M	H HAS M
145	hailstorm	pinhead	H HAS M	M HAS H
146	sweatband	hailstorm	H FOR M	H HAS M
147	fruitcake	hailstorm	H MADE OF M	H HAS M
148	sweatband	pinhead	H FOR M	M HAS H
149	pinhead	fruitcake	M HAS H	H MADE OF M
150	sweatband	fruitcake	H FOR M	H MADE OF M
151	floorplan	eyelash	M HAS H	M HAS H
152	pinhead	floorplan	M HAS H	M HAS H
153	weekday	floorplan	M HAS H	M HAS H

154	pinhead	eyelash	M HAS H	M HAS H
155	eyelash	weekday	M HAS H	M HAS H
156	pinhead	weekday	M HAS H	M HAS H
157	racetrack	tinfoil	H FOR M	H MADE OF M
158	racetrack	doorframe	H FOR M	M HAS H
159	racetrack	newsboy	H FOR M	H HAS M
160	tinfoil	doorframe	H MADE OF M	M HAS H
161	tinfoil	newsboy	H MADE OF M	H HAS M
162	doorframe	newsboy	M HAS H	H HAS M
163	paperclip	stoneware	H FOR M	H MADE OF M
164	paperclip	weekday	H FOR M	M HAS H
165	hunchback	paperclip	H HAS M	H FOR M
166	weekday	stoneware	M HAS H	H MADE OF M
167	stoneware	hunchback	H MADE OF M	H HAS M
168	weekday	hunchback	M HAS H	H HAS M
169	wolfpack	limestone	H MADE OF M	H MADE OF M
170	meatball	wolfpack	H MADE OF M	H MADE OF M
171	snowfield	wolfpack	H MADE OF M	H MADE OF M
172	meatball	limestone	H MADE OF M	H MADE OF M
173	limestone	snowfield	H MADE OF M	H MADE OF M
174	meatball	snowfield	H MADE OF M	H MADE OF M
175	sandpaper	trainload	H MADE OF M	M HAS H
176	bookshop	sandpaper	H HAS M	H MADE OF M
177	dustpan	sandpaper	H FOR M	H MADE OF M
178	bookshop	trainload	H HAS M	M HAS H
179	trainload	dustpan	M HAS H	H FOR M
180	bookshop	dustpan	H HAS M	H FOR M
181	footwear	mothball	H FOR M	H FOR M
182	footwear	birdcage	H FOR M	H FOR M
183	stoplight	footwear	H FOR M	H FOR M
184	birdcage	mothball	H FOR M	H FOR M
185	mothball	stoplight	H FOR M	H FOR M
186	birdcage	stoplight	H FOR M	H FOR M
187	hunchback	sidecar	H HAS M	H HAS M
188	drugstore	hunchback	H HAS M	H HAS M
189	cornfield	hunchback	H HAS M	H HAS M
190	sidecar	drugstore	H HAS M	H HAS M
191	sidecar	cornfield	H HAS M	H HAS M
192	drugstore	cornfield	H HAS M	H HAS M

193	shellfish	mealtime	H HAS M	H FOR M
194	eyelash	shellfish	M HAS H	H HAS M
195	snowfield	shellfish	H MADE OF M	H HAS M
196	eyelash	mealtime	M HAS H	H FOR M
197	mealtime	snowfield	H FOR M	H MADE OF M
198	eyelash	snowfield	M HAS H	H MADE OF M
199	icepack	honeycomb	H MADE OF M	H MADE OF M
200	icepack	beefsteak	H MADE OF M	H MADE OF M
201	sandbank	icepack	H MADE OF M	H MADE OF M
202	beefsteak	honeycomb	H MADE OF M	H MADE OF M
203	honeycomb	sandbank	H MADE OF M	H MADE OF M
204	beefsteak	sandbank	H MADE OF M	H MADE OF M
205	firewood	gunboat	H FOR M	H HAS M
206	meatball	gunboat	H MADE OF M	H HAS M
207	gunboat	kneecap	H HAS M	M HAS H
208	firewood	meatball	H FOR M	H MADE OF M
209	firewood	kneecap	H FOR M	M HAS H
210	meatball	kneecap	H MADE OF M	M HAS H
211	backache	eggshell	M HAS H	M HAS H
212	backache	panhandle	M HAS H	M HAS H
213	shoeshine	backache	M HAS H	M HAS H
214	panhandle	eggshell	M HAS H	M HAS H
215	eggshell	shoeshine	M HAS H	M HAS H
216	panhandle	shoeshine	M HAS H	M HAS H
217	hailstorm	alehouse	H HAS M	H HAS M
218	speedboat	hailstorm	H HAS M	H HAS M
219	goldfish	hailstorm	H HAS M	H HAS M
220	alehouse	speedboat	H HAS M	H HAS M
221	alehouse	goldfish	H HAS M	H HAS M
222	speedboat	goldfish	H HAS M	H HAS M
223	shipload	woodwork	M HAS H	H MADE OF M
224	shipload	birdcage	M HAS H	H FOR M
225	minefield	shipload	H HAS M	M HAS H
226	birdcage	woodwork	H FOR M	H MADE OF M
227	woodwork	minefield	H MADE OF M	H HAS M
228	birdcage	minefield	H FOR M	H HAS M
229	wolfpack	eggshell	H MADE OF M	M HAS H
230	wolfpack	footwear	H MADE OF M	H FOR M
231	tollgate	wolfpack	H HAS M	H MADE OF M

232	footwear	eggshell	H FOR M	M HAS H
233	eggshell	tollgate	M HAS H	H HAS M
234	footwear	tollgate	H FOR M	H HAS M
235	facemask	lifeguard	H FOR M	H FOR M
236	dustpan	facemask	H FOR M	H FOR M
237	mousetrap	facemask	H FOR M	H FOR M
238	dustpan	lifeguard	H FOR M	H FOR M
239	lifeguard	mousetrap	H FOR M	H FOR M
240	dustpan	mousetrap	H FOR M	H FOR M