2023

Investigating Semantic Effects in Adjective-Noun Conceptual Combination

Tara McAuley
University of Windsor

Investigating Semantic Effects in Adjective-Noun Conceptual Combination

By

Tara McAuley

A Dissertation
Submitted to the Faculty of Graduate Studies
through the Department of Psychology
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy at the
University of Windsor

Windsor, Ontario, Canada

Investigating Semantic Effects in Adjective-Noun Conceptual Combination

by

Tara McAuley

APPROVED BY:

____
C. Gagné, External Examiner
University of Alberta

____
K. Poling
Department of Integrative Biology

____
D. Jackson
Department of Psychology

____
K. Romero
Department of Psychology

____
L. Buchanan, Advisor
Department of Psychology

September 16, 2022

DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

Conceptual combination is an active meaning construction process involved in the production and comprehension of complex concepts (e.g., SLEEP TREE, STONY FACE). Distributional and schema-based theories of conceptual combination have proposed various cognitive mechanisms with a primary focus on the processing of noun-noun complex concepts (e.g., SLEEP TREE). The manipulation of variables related to the constituent (e.g., relational frequency) and phrase (e.g., typicality) composition has provided insightful advances into the conceptual representation and processing of complex concepts. Within this context, semantic variables related to semantic richness and concreteness of complex concepts have not been examined in the conceptual combination literature despite having been thoroughly investigated with respect to the processing of simple concepts.

The primary objective of the current study is to investigate the processing of adjective-noun combinations (e.g., STONY FACE) by manipulating semantic variables related to the constituent (i.e., semantic neighbourhood density or SND) and phrase (e.g., concreteness) structure. The adjective-noun stimulus set was constructed with participant ratings using a novel quantitative measure to capture a varying degree of novelty (Experiment 1a) and concreteness (Experiment 1b). In the remaining experiments, the processing of adjective-noun combinations was examined with methodology capturing online processing with tasks of differential semantic engagement (Experiments 2-4) as well as an offline interpretation task (Experiment 5). Collectively, the findings of the current study inform our understanding of the conceptual representation and comprehension of adjective-noun phrases.

The results of the online processing experiments demonstrated orthographic and semantic effects, which were observed in a graded fashion based on the level of semantic processing the task required. In the shallowest double lexical decision task with non-pronounceable non-words (Experiment 2), only orthographic effects pertaining to the visual word form of adjective-noun phrases were found (i.e., combined letter length, mean orthographic frequency). In Experiment 3, where non-words were

pronounceable and required a deeper level of semantic processing compared to Experiment 2, a partial meaningfulness effect was observed, as high meaningful adjective-noun pairs had faster response latencies compared to low meaningful adjective-noun pairs, though no differences were observed for the intermediate meaningful group. A concreteness effect, in which concrete word pairs are processed faster relative to abstract word pairs, was also observed in Experiment 3, particularly for low meaningful adjective-noun phrases. Complete main effects of meaningfulness and concreteness were observed in Experiment 4, the deepest semantic processing task that required participants to make judgments about whether adjective-noun pairs made sense as a pair, essentially recruiting conceptual combination under pressured time constraints. SND effects were also prominent in Experiment 4 and yielded asymmetrical modifier and noun effects based on the meaningfulness and concreteness of the phrase. In Experiment 5, participants were asked to provide an explicit interpretation of novel (low meaningful) adjective-noun phrases, and four themes of interpretation types were identified, including slot-filling, noun elaboration, abstraction, and adjective-reversal. The proportion of unique interpretations and interpretation types differed based on the semantic composition of the adjective-noun phrases.

The results were taken as further support for language-based models of conceptual representations, based on the SND effects observed in Experiment 4 and 5, as SND is a quantitative variable derived from a language-based co-occurrence model (Durda & Buchanan, 2008). Kintsch's (2000) computational model of constructing sentence meaning was applied as a mechanism of constructing meaning for adjective-noun phrases using Experiment 4 and 5 findings, based on previous results in adjective-noun metaphors (Al-Azary et al., 2021). This model can account for a variety of points made by other theorists of conceptual combination, including recruitment in both familiar and novel phrases, an important role of the modifier, an interaction between modifier and noun constituents, competition among different potential processing routes, and recruitment of prior background knowledge.

TABLE OF CONTENTS

LIST OF DEFINITIONS

Abstractness - denoting a concept that is immaterial, conceptual, or nonspecific and not tied to a physical object

Attribute – a property dimension of a conceptual representation, such as colour, shape, and taste. Also called a **slot**

Complex concept – a conceptual representation that is a product of conceptual combination denoted by two linguistic units with or without an interspace (e.g., MARBLE DREAM, BUTTERFLY)

Compound word – a complex concept without an interspace between two constituents (e.g., BUTTERFLY)

Combination – a complex concept with an interspace between two constituents which may have a noun-noun composition (e.g., MARBLE DREAM) or an adjective-noun composition (e.g., FANCY DREAM)

Conceptual combination – An active compositional cognitive process involved in combining two simple concepts to construct a coherent conceptual representation of a complex concept that a single concept cannot sufficiently capture

Conceptual representation, or **concept** – information about word meanings that is formed, stored, and retrieved from the mind, used interchangeably with **semantic representation**

Concreteness - denoting a concept that is based in physical reality with sensory experiences.

Constituent – One of two simple concepts within a complex concept

Dimension-based models, or **slot-filling approaches** – in conceptual combination, these models propose that head nouns are schemas that can be decomposed into attribute-value pairs (e.g., taste-sweet) that become altered by the preceding modifier

Embodied approach – incorporates perceptual neural substrates into the processing of semantic information, with the content of conceptual representations involving grounding in prior sensory-motor experiences

Extensional set – meaning is defined by all referents that can be included within the set (e.g., a bachelor is a list of all unmarried men in the world)

Filler – the specific entity that occupies an attribute dimension, such as red, round, and sweet. Also called a **value**.

Intensional set – meaning is defined by establishing necessary and sufficient properties that warrant inclusion in the set (e.g., a bachelor is an unmarried man)

Language-based model – symbolic models that propose that conceptual representations are formed and organized through the linguistic context in which words co-occur

Lexical representation – information about word forms that is stored and retrieved from the mind

Lexical variable – A variable related to a word's lexical representation or word form (e.g., letter length, orthographic frequency)

Linguistic head – The second constituent in a complex concept, also called a **head noun**

Linguistic modifier – The first constituent in a complex concept, also called a **modifier** that may be an adjective or a noun

Meaningfulness – the ease of meaning construction when encountering a complex concept

Object-based model, or **schema-based model –** symbolic models that propose that conceptual representations store and organize information based on an object's physical characteristics, including feature-based and network models

Psycholinguistic model – model of the underlying nature of language processes

Noun phrase- used to encompass all complex concepts with head nouns including noun compounds and combinations

Property-mapping – in conceptual combination, an interpretation strategy that involves mapping a salient property of the modifier onto the head noun (e.g., a SKUNK BOX is a "smelly box")

Psychological model – model of the underlying nature of cognitive processes

Referent – the person, thing, or idea that is being referred to by a word or phrase

Relational linking – in conceptual combination, an interpretation strategy that forms a relation between modifier and head nouns (e.g., a SKUNK BOX is a BOX "full of" SKUNKS)

Semantic neighbourhood density – a measure of semantic richness that reflects the variability in the distribution of semantic neighbours in a target's word semantic neighbourhood. Semantic neighbourhoods may be described as "dense" if a target word is tightly associated with its' semantic neighbours, or it may be described as "sparse" if a target word is loosely dispersed from its' semantic neighbours

Semantic representation – information about word meanings that is formed, stored, and retrieved from the mind, used interchangeably with **conceptual representation** or **concept**

Semantic richness – a multidimensional construct that captures the amount of variability in the information related to a word's meaning

Semantic variable – A variable related to a words semantic representation or a word's meaning (e.g., semantic neighbourhood density, concreteness, meaningfulness)

Simple concept – A conceptual representation within the mind denoted by a single linguistic unit

Slot – a property dimension of a conceptual representation, such as colour, shape, and taste. Also called an **attribute**.

Symbolic approach – an approach that suggests that semantic processing involves translation, in which external words are mapped onto an internal conceptual representation, and this processing is distinct from lower level perceptual sensory-motor processing

Value – the specific entity that occupies an attribute dimension, such as red, round, and sweet. Also called a **filler**.

LIST OF ABBREVIATIONS

CARIN – Competition Among Relations in Nominals

DLDT – Double Lexical Decision Task

ECCo – Embodied Conceptual Combination

ERP – Event-related potential

fMRI- functional Magnetic Resonance Imaging

HAL – Hyperspace Analogue to Language

LASS - Language and Situated Simulation theory

LOTH – Language of Thought Hypothesis

OF – Orthographic frequency

RICE – Relational Interpretation Competitive Evaluation

RT – Reaction time

SMM- Selective Modification Model

SND- Semantic neighbourhood density

WAT – Words as Social Tools

WINDSORS – Windsor Improved Norms of Distance and Similarity of Representations of Semantics

CHAPTER 1

INTRODUCTION

Conceptual representations are created and modified based on the accumulation of our experiences and knowledge of the world, and they serve to classify and label to help us understand our surroundings. According to the Language of Thought Hypothesis (LOTH; Fodor, 1975), conceptual representations have linguistic structure, such that each word and its underlying meaning correspond to a concept (Murphy, 1988). Because concepts form the basis of word meaning they must permit the compositionality inherent in language (Hampton, 1997). This compositional process is called conceptual combination, and it is considered to be a fundamental cognitive process that accesses and merges basic concepts to form a meaningful complex concept (Ran & Duimering, 2009; Wisniewski, 1996). Conceptual combination allows for the innovative construction of elaborate and infinite representations by recycling familiar concepts to form novel concepts (Hampton, 1997; Thagard, 1984). Such construction ultimately permits language extension (Wisniewski, 1996). For example, the conceptual concept *fuzzy snake* may be unfamiliar, but we can understand what such a concept is and may even generate an image of what it might look like. Deciphering the mechanism(s) and process(es) involved in conceptual combination will thus provide insight into conceptual representations as well as both language comprehension and language production (Maguire et al., 2007).

Numerous theories have been proposed to explain how we represent semantic knowledge at a conceptual level. A general overview of these linguistic and psychological theories follows, with an initial emphasis on simple concepts (e.g., BERRY, SHADOW) followed by a consideration of more complex concepts (e.g., BLUEBERRY; DARK SHADOW). Although this dissertation is not meant to test a symbolic versus embodied approach to conceptual representation, both proposals are delineated in the context of semantic processing to provide a full picture of the state of the current literature, with an emphasis on the former (symbolic) distributional language-based theories. Collectively, theories of semantic representation inform the basis of theories of conceptual combination, which will be examined in subsequent sections. Current challenges and limitations in the field of conceptual combination include a paucity of research into the processing of adjective-noun combinations relative to noun-noun combinations. Further, well-established semantic richness effects that have been studied in simple concepts have not been investigated, and therefore implemented, in current psychological models of conceptual combination. Lastly, an examination of abstractness in the context of conceptual combination is a missing but crucial step in establishing a comprehensive model of conceptual combination. This dissertation serves to address these gaps in our current knowledge base in an extension of McAuley (2018).

Overview of Theories of Conceptual Representation and Processing

Prior to exploring conceptual combination, researchers began with the more fundamental question of how the mind interacts with the world (Hampton, 1997). Concepts are considered the mental representations that operate at this interface to allow us to infer meaning from our experiences (Hampton, 1997). Although concepts are highlighted as a unifying aspect critical to all domains of cognition (e.g., deductive and inductive reasoning, instantiating a goal; Ran & Duimering, 2009), the present overview will be circumscribed to the representation and processing of lexical concepts (i.e., the semantic unit that corresponds to a linguistic form) in semantic memory. Early research in this field began by investigating the functional role of conceptual representations in object classification, or how we label and categorize objects (Hampton, 1997). Classification theories will be briefly outlined below, as many early theories of conceptual combination emerged from such theories.

Aside from categorization, communication is a second broad functional role of lexical concepts (Rips et al., 2012), and this function has guided research on how word meanings and concepts (i.e., knowledge representation) are structured within semantic memory and constructed to create meaning. Conceptual combination is thought to serve a critical function in communicative contexts (Rips et al., 2012; Wisniewski, 1996) and likewise, in semantic processing. Providing an overview of theories of semantic representation is meant to set the stage for the present study. These foundational theories will be briefly described to highlight the organizing principles underlying semantic knowledge as well as the content of semantic representations, with an emphasis on the representations of concrete and abstract concepts, as this distinction is pertinent to the present study.

Classification Theories of Concepts

Many early theories of concepts embody the notion proposed by Rumelhart (1980) that all knowledge is embedded and organized within units (i.e., schemas; similar to frame theory proposed by Minsky, 1975). Schemas facilitate the acquisition of new knowledge by incorporating novel experiences and information into pre-existing knowledge structures or by creating new schemas (Rumelhart, 1980). Prototype models are one class of models derived from this framework. Rosch and Mervis (1975) proposed that conceptual prototypes comprised of discrete features form an ideal instance of a category and constitute the representational form of a concept. Items that share more attributes with other members within a category and fewer attributes with members in other categories exemplify the prototype of category. Rosch and Mervis (1975) found support for a shared family resemblance among exemplars of a category based on their underlying distribution of attributes and proposed that other instances encountered are categorized on the basis of similarity to the prototype. Empirical support for prototype models comes from exhibiting typicality effects in sentence verification and semantic categorization tasks. For example, Mervis and Rosch (1981) found that participants were faster to classify a ROBIN as a type of bird compared to an OSTRICH because the former is considered to more closely resemble a prototypical bird based on feature overlap (e.g., nests in trees, is able to fly).

Posner and Keele (1968) conducted a series of experiments on perceptual classification learning and proposed that these prototypical examples may be constructed over time through experiences with the exemplars of a category. In contrast to prototype theories such as Rosch and Mervis' (1975), exemplar theories like Posner and Keele's (1968) propose that concepts are represented by these average instances in memory (Medin & Schaffer, 1978; Smith & Medin, 1981). For example, Medin (1975) proposed a context model in which retrieval of stored exemplar information is the basis of object classification as opposed to feature-based information. Concretely this means that one's representation of a ROBIN is based on prior instances in which a robin was encountered by an individual (i.e., through media or real-life experiences). Empirical support for this model comes primarily from classification learning experiments (e.g., Medin & Schaffer, 1978; although see Murphy, 2016 for recent critiques) in which participant classification judgments were facilitated by prior exposure to exemplars.

Murphy and Medin (1985) argued that judgements based on perceptual similarity to a stored representation do not adequately constrain concepts to facilitate conceptual coherence (i.e., meaning). Rather, the influence of concepts and world knowledge should be viewed as bidirectional and closely intertwined: Concepts help us create meaning from the world and our understanding of the world (and its underlying principles -- referred to as "theories") imposes meaning on the concepts themselves. Thus, Murphy and Medin (1985) proposed that concepts are embedded in a deeper understanding of connections within their observed traits. For example, the concept of BIRD is represented within an extensive set of interconnections of theoretical structures that facilitate understanding of why birds have wings (e.g., *allows them to fly*), why they fly (e.g., *migrate, defense against predators, forage for food*), and why they nest in trees (e.g., *hidden from predators*). As such, each observed trait is connected through links that are explanatory, causal, and/or goal directed. This theory-based network of world knowledge is thought to be most relevant when categorizing novel objects or borderline cases that do not neatly fit into a category (Murphy & Medin, 1985).

The above theories form the basic building blocks for modern theories of conceptual combination. These classification theories are primarily limited to capturing how concrete concepts are categorized, because abstract concepts do not have observable features to form prototypes or exemplars in memory. Additionally, classification theories face numerous challenges when accounting for the compositionality of concepts (i.e., conceptual combination) that will be discussed in a later section. The representational structure of the individual constituent concepts is thought to heavily influence the interactive process of conceptual combination (Murphy & Medin, 1985). With this in mind, a more focused description of other relevant models that focus on semantic representation and processing follows.

Concept Representations in Semantic Memory

Language is thought to have an influential organizing effect on conceptual structure and function (Medin & Coley, 1998). Many conceptual representation models postulate differences in how semantic concepts are formed, stored, and retrieved in semantic memory. The primary focus in this section will be on models that make explicit reference to the underlying organizational principles of semantic memory, with structural distinctions based on semantic similarity and concreteness discussed. Within this discussion, the content of semantic representations will be outlined, with reference to the symbolic and embodied distinction. Although not central to this dissertation, a review of conceptual representation would be incomplete without discussing the embodied perspective in this context.

Semantic Similarity and Content. One critical underlying difference between models of conceptual representation is the description of semantic similarity (both definition and implementation) in the structure of semantic memory. Object-based, or schema-based, models propose that semantic similarity is defined by an objects' physical characteristics (e.g., colour, size, shape) whereas language-based models define semantic similarity through associations in linguistic contexts (Buchanan et al., 2001). These models are subsumed under symbolic models, which propose that the content of conceptual representations consists of internal symbolic representations that external words map onto, and meaning is derived through the relationships between abstract symbols (Meteyard et al., 2012). Importantly, symbolic models propose that perceptual and sensory inputs are transduced into symbols, reserving distinct and higher-level processing for language that is separate from lower-level perceptual processing. At the other end of the spectrum, embodied theories propose a perceptual basis for the content of conceptual representations, and intimately integrate perceptual level processing into a function of semantic processing (Meteyard et al., 2012).

Object-based models, which assume an underlying schematic structure, can be largely divided into feature-based models and network models. In their feature comparison model, Smith et al. (1974) proposed that lexical items are represented by their semantic features, which vary in the extent that they characterize a category. Defining features are those that are shared by all category members whereas characteristic features are those that are shared by most members of a category but are not necessary for category membership. Returning to our example of the concept BIRD, defining features may include *have wings* and *lay eggs*, whereas a characteristic feature may be *able to fly*. As such, features have associated weights that capture their relevance to the concept. When encountering a category-member semantic categorization task, Smith et al. (1974) described a two-stage comparison process in which all features (both defining and characteristic) are initially considered and compared for similarity. If this process yields uncertainty in the response, then a second process ensues that restricts the search to shared defining features based on their associated feature weights. This model was supported by their earlier research on typicality ratings of category-exemplar pairs (Rips et al., 1973) as well as the family resemblance/typicality findings described earlier (Mervis & Rosch, 1981; Rosch & Mervis, 1975).

Similarly, Tversky's (1977) contrast model posits that concepts of objects are captured by a collection of features, and a similarity assessment is pursued to categorize objects. This latter process is described as a comparison of featural overlap and consideration for the weighted difference between shared and distinctive features. Unlike the feature-comparison model, the contrast model considers the salience of features as well as context in the assessment of similarity. Tversky (1977) also acknowledged that representations contain world knowledge (as proposed by Murphy & Medin, 1985), but argued that certain tasks (e.g., similarity assessments) do not elicit the recruitment of this elaborate knowledge to perform the task. Murphy and Medin (1985), however, ultimately propose that similarity (as the basis for meaning) lacks precision and explanatory power because an attribute matching process has the potential to produce infinite possibilities. As such, feature similarity between concepts within a category is viewed as simply a by-product rather than a determinant of conceptual coherence (Murphy & Medin, 1985).

Akin to feature-based models, network models propose that semantic information is represented by features, although these features are thought to be stored within relationally linked nodes in which the path length between nodes reflects the similarity among concepts (e.g., Collins & Quillian, 1969). Additionally, they propose that many concepts do not require the storage of redundant general information. Instead, information about exemplars can be inferred by their connections to a higher-order category. For example, the node that represents the concept BIRD may include the feature *has wings* whereas the connected lower order node for HUMMINGBIRD would not. Thus, they propose that semantic knowledge is organized hierarchically in individual nodes with general concepts at the top of the hierarchy and specific concepts at the bottom. Their experiments supported this notion, as participants were faster to verify the statement “a hummingbird can sing” as opposed to “a hummingbird has wings” since the latter purportedly requires retrieval of information from the superset category node (Collins & Quillian, 1969). Therefore, the entire meaning of a concept is captured within a connectionist network of interlocking nodes, and each property differs in its associated weight towards the meaning of the concept. Compared to the feature comparison model, network models propose a reliance on diverse sources of evidence to make categorization decisions (e.g., superordinate links) and argue that members within a category cannot be unified by a set of defining features.

Collins and Loftus (1975) proposed an extended spreading activation framework to the network model. In this extension, concepts are organized by semantic similarity, which is determined by the aggregate of shared interconnected nodes, and spreading activation occurs when the stimulation of a target concept results in the activation of nearby, related concepts to facilitate interpretation. These models are supported by empirical evidence such as the semantic priming effect (Meyer & Schvaneveldt, 1971), in which response latencies to a target word (e.g., BUTTER) were faster when preceded by a semantically related word (e.g., BREAD) than an unrelated word (e.g., NURSE) in a primed lexical decision task.

In contrast to object-based models, other models that focus on the organizing principles of semantic knowledge emphasize the linguistic context in which words occur. Such language-based models postulate that meaning construction arises through our interactions with language use in the world, and by doing so acknowledge that there are individual differences in the structure of semantic space based on experiences. They do propose, however, that general organizational influences impose on the structure of semantic space and can be characterised and delineated (Buchanan et al., 2001). Association models, for example, determine the semantic associates of target words by using human judgements in free association tasks (Nelson, et al., 2000; Nelson et al., 2004) whereas lexical co-occurrence models quantify the characteristics of semantic space by aggregating large volumes of text and generating large databases to compute how frequently words co-occur in similar linguistic contexts (Buchanan et al., 2001). In contrast to classifying lexical concepts by feature overlap or category membership that is central to object-based models, in language-based models, concepts are classified by their statistical co-occurrence in language and modeled within an associative network.

This notion is instantiated in Hyperspace Analogue to Language (HAL; Lund & Burgess, 1996), a representational model of semantic memory that produces a lexical co-occurrence matrix by analyzing written text. Words are represented as vectors in a high-dimensional semantic space, and the distance between vectors quantifies their semantic similarity. As such, the meaning of a target word is captured by its relation to other associated words in similar linguistic contexts. For example, the meaning for CANDLE is constructed by its co-occurrence with related words such as FLAME, WICK, and LIGHT. These associated words are semantic neighbours of the target word CANDLE. The metrics from Durda and Buchanan's (2008) Windsor Improved Norms of Distance and Similarity of Representations of Semantics (WINDSORS) eliminated the frequency effects that confounded the HAL database (see Lutfallah et al., 2018 for semantic neighbourhood data).

Other models that propose that semantic relatedness is constrained by occurrence in linguistic contexts include Latent Semantic Analysis (LSA; Landauer & Dumais, 1997), High Dimensional Explorer (HiDEx; Shaoul & Westbury, 2006), Latent Dirichlet Allocation (LDA; Blei et al, 2003), Topic Model (Griffiths et al., 2007), and Bound Encoding of the Aggregate Language Environment (BEAGLE; Jones & Mewhort, 2007). Collectively, these models are coined semantic distributional models, as meaning is characterized as a function of the statistical distribution of words across written and spoken contexts (Andrews et al., 2014) and similar linguistic contexts tend to share semantically similar words (i.e., the distributional hypothesis; Harris, 1954). The plausibility of semantic distributional models as reflective of human semantic representation has been studied extensively (for an overview, refer to Günther et al., 2019; Jones et al., 2015) and has a large empirical basis from behavioral tasks including word recognition (e.g., Danguecan & Buchanan, 2016; Durda & Buchanan, 2008; Lund & Burgess, 2008), priming (Jones et al., 2006; Mandera et al., 2017), semantic classification (e.g., Pexman et al., 2008; Yap et al., 2011), semantic relatedness (e.g., Malhi & Buchanan, 2018), and metaphor processing (e.g., Al-Azary & Buchanan, 2017) tasks, to illustrate a few, as well as in clinical populations including patients with deep dyslexia (e.g., Al-Azary et al., 2019; Buchanan et al., 1996; Malhi et al., 2019).

Compared to object-based models, language-based distributional models have the advantage of being quantifiable by characterising semantic space through objective measures, such as semantic neighbour distance, semantic neighbourhood size, and semantic neighbourhood density (Buchanan et al., 2001). Similar to Collins and Loftus (1975), language-based models are aligned with a spreading activation mechanism that facilitates activation between semantic neighbours within a semantic neighbourhood. Language-based models can also account for semantic priming tasks in which priming occurs between associated words (e.g., SCRATCH and CAT) in the absence of featural overlap between the concepts (Buchanan et al., 2001). Further, Durda et al. (2009) demonstrated mapping between co-occurrence vectors and featural knowledge, suggesting that the latter information is subsumed within the former (see also Baroni & Lenci, 2010).

Both object and language-based models described above are considered symbolic in that the basis of semantic processing is a function of translation, where external lexical items are mapped onto an internal conceptual representation (Weiskopf, 2010) and perceptual and sensory processing is hypothesized to be qualitatively distinct and fundamentally separate from conceptual processing (Meteyard et al., 2012). In contrast to the symbolic view, real world perceptual experiences form the basis of understanding words in embodied theories. Within this framework, the content of conceptual representation is grounded within sensory and motor information, and perceptual neural substrates are thought to be simultaneously recruited when processing semantic information (Meteyard et al., 2012). Thus, to understand CANDLE, an embodied theorist would suggest that this concept is understood through our prior experiences with CANDLE, including seeing a lit candle and smelling its scent.

One prominent embodied theory, Barsalou's (1999) Perceptual Symbol Systems theory, posited the recruitment of initial bottom-up processing during concept acquisition, with direct perceptual experiences engaging sensorimotor regions of the brain. These sensory, proprioception, and introspection experiences elicit neural activation (i.e., perceptual symbols), which then form perceptual schematic representations that are stored in memory. As such, conceptual content consists of the partial recordings of the initial neural activation during direct perceptual experiences, and this activation is simulated in a top-down fashion in the absence of perceptual experience to facilitate semantic processing (Barsalou, 1999). Another embodied theory that proposed a similar mechanism at the sentence level is called the Indexical Hypothesis (Glenberg & Robertson, 1999). According to this theory, language comprehension is facilitated through the indexing between words and phrases to objects, pictures, or perceptual symbols, which in turn generates and meshes affordances (i.e., the different ways in which individuals are able to interact with an object) to constrain interpretation. Embodied theories are supported by research on body-object interactions at the word (e.g., Siakaluk et al., 2008) and sentence (e.g., Glenberg & Kaschak, 2002) level, functional neuroimaging studies (e.g., Esopenko et al., 2012), and in clinical populations such as patients with motor neuron disease (e.g., Bak et al., 2001).

Many theorists reject a purely symbolic or purely embodied approach, as both approaches are represented at extreme ends of a spectrum (e.g., Andrews et al., 2014; Meteyard et al., 2012), and many theories instead lie within this continuum (although see Goldinger et al., 2016 for recent criticisms of the embodiment view). For example, Meteyard et al. (2012) argue that semantics is not possible without some form of symbolism and context independent representation. Semantic processing during reading starts with an arbitrary format (e.g., the visual word) to refer to the same concepts consistently and systematically, whether the referent is concrete or abstract (Meteyard et al., 2012). Further, both approaches have extensive support in the literature as aforementioned.

To reconcile these seemingly oppositional theories, mixed interactive models have gained popularity, including the Language and Situated Simulation theory (LASS; Barsalou et al., 2008) and the Symbol Interdependency Hypothesis (Louwerse, 2007). Both theories assume a parallel but distinct time course of activation of linguistic and perceptual systems, wherein the linguistic system reaches peak activation prior to the perceptual system, suggesting that many linguistic tasks (e.g., lexical decision tasks) do not recruit the deeper processing elicited by the perceptual system. In line with this view, Louwerse and Jeuniaux (2010) as well as Malhi and Buchanan (2018) have found that when task demands evoke embodied relations (i.e., iconicity judgement task), an embodied factor (i.e., iconicity), but not a symbolic factor (e.g., semantic distance), was recruited to facilitate processing whereas the opposite was found for a task that evoked symbolic relations (i.e., semantic relatedness task).

Louwerse (2007) further proposed that symbols can convey perceptual experiences, and the language system can function as a short-cut to the perceptual system. Thus, language serves to encode relations in the world, and can also capture embodied relationships in this way and form intralinguistic relationships (Hutchinson & Louwerse, 2014; Louwerse, 2011; Louwerse and Jeuniaux, 2008). Riordan and Jones (2011) also highlighted the redundancy of information coded by perceptual and linguistic experiences and showed substantial evidence that language-based distributional statistical models can account for most of the information captured by sensory-motor feature-based data. Additionally, embodied theories have been largely circumscribed to understanding concrete concepts, a relative strength for symbolic models (Dove, 2014), although some have been extended to incorporate abstract concepts (e.g., Kousta et al., 2011). Concreteness is another relevant organizing principle of semantic knowledge, and therefore, conceptual representations.

Concreteness. Many empirical investigations in language and memory research demonstrate a concreteness effect, or an advantage for concrete words (e.g., COUCH) over abstract ones (e.g., JUSTICE; Kroll & Merves, 1986; Paivio, 1971), and these processing differences suggest unique conceptual representations dependent on the concreteness (or abstractness) of the concept. This notion is further supported in studies utilizing neuroscience techniques such as event-related potentials (ERPs) or functional magnetic resonance imaging (fMRI) (e.g., Huang and Federmeier, 2015; Huang et al., 2010; Pexman et al., 2007). Neuropsychological case studies in deep dyslexia (e.g., Katz & Goodglass, 1990; Plaut & Shallice, 1993) and case series in semantic dementia (e.g., Jefferies et al., 2009) also highlight differences in the representations of concrete and abstract knowledge.

Two primary theories suggest that concrete word representations benefit from more rich associations than abstract ones. In the Dual Coding Hypothesis, Paivio (1971) proposed two distinct but functionally related systems: a verbal, linguistic system, and a nonverbal, imagistic system. In his theory, concrete words are represented within both systems and benefit from a processing advantage due to a broader access to information, whereas abstract words are represented solely by the linguistic system. The second theory, the Context Availability Hypothesis (Schwanenflugal & Shoben, 1983), proposed that concrete and abstract concepts are both represented verbally. To facilitate access to meaning, associations with contextual knowledge (e.g., from discourse or semantic memory) are recruited, and concrete words have fewer and richer associations with contextual knowledge whereas abstract words have weaker associations that are more widely dispersed. These two models have been combined based on ERP findings and form the basis of the Context-extended Dual Coding Theory (Holcomb et al., 1999; West & Holcomb, 2000). ERP studies have classified two components related to concreteness: a larger N400 and a late component peaking between 700-800 milliseconds. The N400 effect is thought to elicit engagement from the verbal semantic system for both concrete and abstract words, although this is amplified in concrete words consistent with having denser interconnections within semantic memory (in accordance with context availability theory), whereas the later 700-800 component is thought to reflect recruitment of mental imagery for concrete words only (consistent with dual coding theory).

Although influential, the above theories largely focus on what abstract concepts lack. However, abstract concepts are arguably more complex than concrete concepts (Borghi & Binkofski, 2014; Schwanenflugel, 2013). For example, abstract concepts are characterised by variability in meaning within and between individuals and are often shaped by dynamic linguistic and social experiences (Borghi & Binkofski, 2014; Borghi et al., 2017). This is supported by fMRI evidence that visualizes the recruitment of more widely distributed brain regions for abstract concepts (e.g., Rodríguez-Ferreiro et al., 2011). Relatedly, abstract concepts require context to constrain their meaning, whereas concrete concepts are more reliably consistent across contexts (Crutch & Warrington, 2005). Abstract concepts are also shown to be acquired later in life (Schwanenflugel, 1991) lending additional support to the above points. Further, abstract concepts are themselves a heterogenous group on the basis of their content, as some denote reference to mental states (e.g., CURIOUS, MEDITATION), social scenarios (e.g., PARTY, FRIENDSHIP), and institutional properties (e.g., LAW, OWNERSHIP), to name a few. Their complexity and variability likely emerge from what abstract concepts do lack, which is a direct referent to an object (Borghi et al., 2017), whereas concrete concepts are defined by their intrinsic properties (Wiemer-Hastings & Xu, 2005). Rather than focusing on quantitative differences between abstract and concrete concepts, many researchers have turned to qualitative distinctions by proposing that abstract concepts are defined by complex associations with objects and linked to internal, introspective experiences including emotional states, situations, events, and beliefs.

For example, in the Different Representational Frameworks model, Crutch and Warrington (2005) propose that concrete and abstract concepts are organized within fundamentally discrete architectures and evoke qualitatively distinct conceptual relations. That is, concrete concepts are primarily organized by categorical relations based on semantic similarity (e.g., OVEN-FRIDGE) rather than semantic associations based on linguistic contexts (e.g., OVEN-APRON) whereas abstract concepts mainly consist of semantic associations (e.g., WAR-PEACE) as opposed to semantic similarity (e.g., WAR-CONFLICT). This notion was supported in clinical cases (e.g., Crutch & Warrington, 2005; Crutch et al., 2006) as well as healthy participants (e.g., Crutch et al., 2009; Duñabeitia et al., 2009), although recent mixed findings (e.g., Hamilton & Martin, 2010) have led Crutch and Jackson (2011) to propose that these conceptual relations are graded across concrete and abstract concepts, rather than binary.

Initially, abstract concepts presented a challenge to embodied theories (Dove, 2014). However, many emerging embodied theories have suggested that the grounding of abstract concepts is qualitatively different from concrete concepts. For example, the Theory of Embodied Abstract Semantics

(Vigliocco et al., 2009) proposes that experiential (i.e., sensory and motor for concrete; affective for abstract) and linguistic (i.e., verbal associations) information contribute to the representations of concrete and abstract concepts. In order to rule out the explanations proposed by the Dual Processing Theory and Context Availability Theory, Kousta et al. (2011) designed a linguistic stimulus set that controlled for imageability and context availability among other known influential linguistic variables. Using lexical decision tasks, they found an “abstractness effect”, or an advantage for abstract words compared to concrete words. The findings were attributed to differences in experiential information; that is, affective associations characterised abstract stimuli (Kousta et al., 2011). These findings led to proposal of the “affective embodiment account” (Kousta et al., 2011), in which concrete and abstract concepts are differentially grounded either externally through our interactions with the physical environment or internally through our experiences with emotional states, respectively. Other embodiment proposals that aim to describe the grounding of abstract concepts include Barsalou and Wiemer-Hastings (2005), who stated that individuals recruit introspective experiences and simulate concrete situations to represent abstract concepts. Similarly, Lakoff and Johnson (1980) proposed that individuals rely on metaphorical mappings from concrete to abstract concepts (i.e., Conceptual Metaphor Theory).

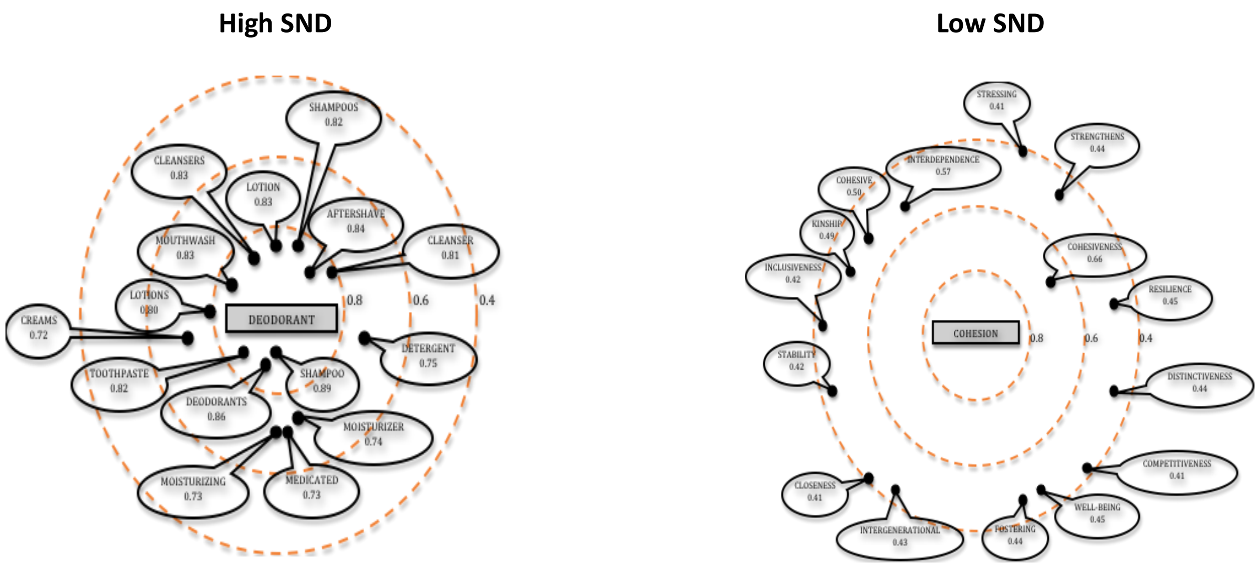
Despite these embodied accounts that aim to ground abstract concepts, language is thought to play a crucial function in abstract concept representation (Dove, 2009; Binder et al., 2016). Similar to Pavio’s (1971) and Vigliocco et al. (2009) theories, multiple representation theories propose two types of semantic representations: a perceptual and motor mediated representation and a language-based one (Borghi et al., 2017). Abstract concepts are thought to be asymmetrically ingrained within the latter linguistic representation, and as such, acquire meaning through their associations with other linguistic representations (Dove, 2011). This notion is instantiated in Dove’s (2011, 2014) model of Representational Pluralism, which proposes that both forms of representation are simulation based and integrative, rather than separated into embodied and disembodied systems as in his earlier proposal (Dove, 2009). Thus, Dove (2011, 2014) postulated that both concrete and abstract concepts engage sensorimotor simulation, but the latter concepts are “dis-embodied” in the sense that these simulations do not facilitate access to semantic meaning. Dove (2011, 2014) drew on the existing behavioral, neuroimaging, and clinical data that have already been alluded to for support of his theory, such as meta-analyses that demonstrate greater recruitment of language neural regions during abstract concept processing (e.g., Binder et al., 2009; Wang et al., 2010).

Similarly, the Words as Social Tools (WAT; Borghi et al., 2013; Borghi & Binkofski, 2014) asserted that language functions as a scaffold for the acquisition of abstract concepts, and it additionally emphasized the importance of social experiences in abstract knowledge construction. In support of these models, Gleitman et al. (2005) suggested that the acquisition of abstract concepts in children relies on mastering words and linguistic knowledge. Furthermore, Recchia and Jones (2012) used a feature generation task in one group of participants that would serve to guide identification of the concept for another group of participants. They found a dissociation of the type of information generated based on the concreteness of the concept; abstract concepts were characterised by many semantic neighbours in fruitful linguistic contexts whereas concrete concepts benefitted from many features in strong physical contexts (Recchia & Jones, 2012).

In line with the notion that linguistic information is central to abstract concept representation, researchers have examined how concreteness and language-based variables interact. For instance, Danguecan and Buchanan (2016) manipulated concreteness and semantic neighbourhood density in a series of single word processing experiments with increasing explicit engagement. Semantic neighbourhood density (SND), a measure derived from the WINDSORS lexical co-occurrence model (Durda & Buchanan, 2008), captures the variability in how semantic neighbours are dispersed within a target word’s semantic neighbourhood and provides an estimate of semantic richness. To expand, a semantic neighbourhood may be described as dense (i.e., high SND) if the target word has close associations with semantic neighbours, or it may be characterised as sparse (i.e., low SND) with a distribution of loose associations (see Figure 1; Danguecan & Buchanan, 2016).

Figure 1

Dense and Sparse Semantic Neighbourhood Densities



Note. A simplified illustration of a target word with closer relations to semantic neighbours on average (i.e., high SND) and a target word with weaker relations to semantic neighbours on average (i.e., low SND; Danguecan & Buchanan, 2016).

In their experiments, Danguecan and Buchanan (2016) found an inhibitory effect of SND for abstract concepts across tasks, in which denser SNDs were associated with slower word recognition times, although this was not consistently found for concrete concepts. Similarly, in offline and online metaphor comprehension tasks, Al-Azary and Buchanan (2017) found an inhibitory effect of SND for metaphors with concrete topics (e.g., *A Pen is a Sword*), which were rated as less sensible and processed slower than abstract metaphors with high SND topics (e.g., *Language is a Bridge*), although no differences were found for low SND metaphors based on topic concreteness. This inhibitory effect of SND and concreteness was also largely replicated in a participant with deep dyslexia, in which high SND metaphors were rated as nonsensible regardless of topic concreteness, and abstract low SND topics (e.g., *Daydream is a Trip*) were rated as most comprehensible (Al-Azary et al., 2019). Taken together, interactions between language-derived variables and concreteness lend additional support to differences in the representation and processing of concrete and abstract concepts.

Challenges to Accounting for Conceptual Combination

The theories described above outlined models of semantic representation and processing that pertain to a simple concept (e.g., BALL, DREAM), whereas linguistic productivity allows for the composition of complex concepts (e.g., also referred to as noun phrases, compounds, conjunctions, or combinations, e.g., SNOWBALL, EMPTY DREAM) from its constituent components. For simplicity, this dissertation will label “compounds” as those composed of two nouns without an inter-space between constituents (e.g., SNOWBALL) and “combinations” as those either assembled as two spaced noun constituents (e.g., MARSHMELLOW DREAM) or an adjective and noun constituent (e.g., EMPTY DREAM). Collectively, these will be referred to as “noun phrases”. Conceptual combination has primarily been examined in the context of the latter combination types, although recent extensions to encompass compounds has been made (e.g., Gagné & Spalding, 2009). Hampton (1997) stated that the process of conceptual combination is at the center of knowledge representation, as it demands understanding of how the underlying meaning of complex noun phrases are constructed from its constituent parts. Therefore, conceptual combination is crucial to forming a complete theory of concepts. In their current form, some of the theories of concepts present challenges in extending to incorporate conceptual combination.

The early classification theories, for instance, rely on prototypical examples or stored instances, whereas conceptual combination allows for the construction of novel and unfamiliar complex concepts. Further, object-based models are challenged by non-compositionality effects, in which some attributes are salient to the conjunction but are not found within either constituent concept (e.g., *live in cages* and *talk* are unique to a PET BIRD but not a PET or a BIRD). In these instances, emergent features are thought to be due to extensional feedback (i.e., experiences and stored knowledge about the concept; Hampton, 1987) or to promote a coherent and sensible concept through theory-driven relations (Murphy & Medin, 1985). Distributional semantic models, in contrast, have the capacity to account for emergent features, as these models are derived from an associative network based on language use, rather than typical features of objects. Similarly, embodied theories incorporate a large network of information from a variety of sources. These latter models can also capture differences between concrete and abstract concept representations, whereas classification and object-based models are more applicable to concrete concepts. Notably, language-based models are thought to be central to abstract concept representation (Meteyard et al., 2012). Nevertheless, many of the described theories have provided a foundation for theories of conceptual combination.

In addition, object-based theories are thought to be applicable to the comprehension of familiar noun compounds (e.g., SKATEBOARD) in which our knowledge is "lexicalised" or can be derived from prior memory instances. Traditionally, psycholinguistic theories have largely focused on noun compounding (i.e., the processing of familiar noun-noun compound words) whereas cognitive psychological theories have studied the process of conceptual combination, which arguably involves novel combinations that engage combinatorial processing for meaning construction (Günther et al., 2020). These two approaches have functioned relatively independently, although attempts to reconcile these disparate literature bases exist (e.g., Gagné & Spalding, 2009; Günther et al., 2020; Ji et al., 2011). To further unify these pursuits, a brief overview of compound processing will be delineated to set the stage for theories of conceptual combination.

Representation and Processing of Compound Words

Similar to simple concepts, compound words are denoted by a single linguistic unit (e.g., CHEESECAKE), although importantly, two simple noun concepts are embedded within a compound and can be decomposed into the modifier constituent (e.g., CHEESE) and the head constituent (e.g., CAKE), in which the latter typically reflects the syntactic or semantic category of the compound in the English language (Gagné & Spalding, 2009; Spencer, 1991). Thus, compounds words are considered complex concepts. Noun compounding has primarily been studied from a psycholinguistic lens that differentiates between lexical and semantic representations. Lexical representations refer to the storage of word-forms in the mental lexicon whereas semantic, or conceptual, representations contain information about word meanings (Gagné & Spalding, 2009).

Historically, there have been two dominant approaches to compound word comprehension (Arcara et al., 2014; Libben & Jarema, 2006). The direct access, or whole word representation, models propose that compound words are stored and processed holistically (e.g., Butterworth, 1983; Bybee, 1995) whereas morphological decomposition, or full-parsing, models generally hypothesize that modifier and head lexical and conceptual representations are extracted from compounds during processing (Libben, 1998; Schreuder & Baayen, 1997; Taft, 2004). In the latter approach, constituent lexical and conceptual representation are hypothesized to be conjointly activated through facilitatory links via a spreading activation mechanism, which in turn indirectly activates the compound's lexical and conceptual representation. Thus, meaning is evoked through the indirect activation of a pre-stored conceptual representation of the compounds' concept via co-activation of the constituent components.

Morphological decomposition models gained popularity in examinations of the processing of semantically transparent compound words (e.g., Libben, 1998), wherein "semantic transparency" refers to the extent that a constituent contains semantic information that is relevant to the compound word's meaning (Libben & Jarema, 2006). Compounds may be described as "fully transparent" meaning that each constituent provides semantic content that is relevant to the interpretation of the compound (e.g., BLUEBERRY is a berry that is blue). Thus, morphological decomposition in these compound types would aid in accessing the compound's meaning. In contrast, the meaning of "fully opaque" compounds have no relation to the semantic content of its' constituents (e.g., MOONSHINE is a type of alcohol).

Compounds can also take an intermediate form, with opaque-transparent (e.g., STRAWBERRY is a berry but not made of straw) as well as transparent-opaque (e.g., JAILBIRD is not a bird but is rather a person in jail) types (Libben & Jarema, 2006). Opaque constituents were thought to present a challenge to morphological decomposition models, given that constituent conceptual representations would not aid in the retrieval of the compound words meaning (Libben, 1998). However, emerging research has suggested that morphological decomposition is insinuated early during processing for all compound types, regardless of the semantic transparency of constituents (Brooks & Cid de Garcia, 2015).

On the basis of conflicting evidence, dual-route models that include both whole word representation and morphological decomposition have been proposed (Baayen et al., 1997; Isel et al., 2003; Koester et al., 2009; Zwitserlood, 1994). Such models are supported by research investigating compound word frequency. For instance, high frequency compounds are thought to be accessed by their whole word representations and lead to faster response times whereas low frequency compounds are more likely to be accessed by the decomposition of their constituents (Arcara et al., 2014; Kuperman et al., 2009). Further neuroimaging evidence suggests that parallel dual route processing occurs in semantically transparent compounds, whereas opaque compounds are primarily processed via the whole-form access route (MacGregor & Shtyrov, 2013).

Nevertheless, research manipulating the lexical and semantic properties of the constituents of compound words provides compelling evidence for decomposition models. For example, constituent frequency has also been shown to affect lexical processing (e.g., Andrews et al., 2004). Additionally, constituent "family size", or the number of unique compounds a given constituent is involved with (e.g., SNOWBALL and SNOWMAN belong to the family with the shared modifier constituent "snow"), has been found to facilitate the lexical processing of compounds in English and Dutch (De Jong et al., 2002). Further, the above effects are not independent; Kuperman et al. (2009) identified multiple interactions among compound frequency, constituent frequency, and constituent family size.

Novel compounds (e.g., SUNBRICK), however, present a challenge to the described models, in which a pre-stored compound representation could not exist. According to Schreuder and Baayen's (1995, 1997) meta-model, when a novel compound is encountered, a concept node (i.e., memory trace) is created through activation of the constituent components to compute meaning. Over time, this novel concept node will succumb to decay if not frequently re-activated. Alternatively, rather than assuming a compound words' meaning is computed passively as a by-product through activation of its' constituents, Gagné and Spalding (2009) suggested that constituent integration is an active process that constructs meaning through semantic composition. In its' essence, *conceptual combination* is proposed to be involved in compound processing, and Gagné and Spalding (2006) advocate for such an active meaning construction process when interpreting both familiar and novel compounds (also see Günther & Marelli, 2016; Libben, 2014). Moreover, the meaning of a compound is not a simple intersection of the constituent concepts (Gagné & Spalding, 2009), as has been demonstrated in the conceptual combination literature (e.g., Hampton, 1987; Medin & Shoben, 1988) and will be discussed further within the conceptual combination section.

Furthermore, an active meaning construction process in compound word processing has a theoretical and empirical basis. Although less efficient than accessing a whole-word representation, Libben (2014) argued that from a communication standpoint, initiating a compositional process upon encountering a compound word would enhance our understanding of the conveyed meaning thus capturing the overarching intent of language use. Consistent with this view, a meaning construction process appears to occur rapidly upon presentation of a compound in numerous psycholinguistic tasks, without recruiting top-down processing (Coutanche et al., 2019; Günther et al., 2020; Ji et al., 2011; Marelli & Luzzatti, 2012). Further, compound processing is affected by the ease of constituent integration (Günther et al., 2020). For example, in their study of German compound words and eye fixations, Inhoff et al. (2000) found that stimulus presentation revealed a decomposition and integration process based on the finding of longer fixation times for inter-spaced compounds. When compounds were presented as spaced (e.g., BLUE BERRY), they proposed that access to constituent lexical representations was facilitated. In contrast, when compounds were presented in their typical form (i.e., without a space as in BLUEBERRY), the constituent integration process was hindered. Juhasz et al. (2005) replicated these findings in English compounds. Additionally, Libben et al. (2003) found that compound words with semantically transparent heads (e.g., BLUEBERRY and STRAWBERRY) were processed faster than compounds with semantically opaque head constituents (e.g., MOONSHINE and JAILBIRD), suggesting that constituent integration was facilitated in the former type.

As aforementioned, compound processing has been viewed as the “linguistic counterpart” of conceptual combination, as both types of noun phrases have largely been studied as independent phenomena (or at least, from distinct theoretical backgrounds), although there have been attempts to align these (e.g., Gagné & Spalding, 2009; Günther et al., 2020). For instance, Günther et al. (2020) stated that compounds are inherently compositional expressions that evolved to simplify complex stimuli, similar to Wisniewski’s (1997) explanation of conceptual combination in communicative contexts. Additionally, an active combinatorial process underlying the processing of all noun phrases (e.g., SNOWBALL, CREDIT CARD, SMART CLOCK) would be parsimonious, and research in behavioral and neuroimaging studies support such a notion. For instance, similar patterns of response times were observed for both familiar and novel noun phrases in priming experiments with lexical decision and sense/nonsense judgment tasks (Gagné, 2001; Gagné & Spalding, 2004; Estes & Jones, 2008). Additionally, brain imaging studies (e.g., Coutanche et al., 2019) and cognitive neuroscience methodology (such as ERPs, e.g., El Yagoubi et al., 2008) converge on structural and functional processing of familiar and novel noun phrases. For example, the left anterior temporal lobe has been identified as a neural substrate relevant to the processing of compound nouns (Brooks & Cid de Garcia, 2015; Flick et al., 2018), noun combinations (Zhang & Pylkkänen, 2015), and adjective-noun combinations (Baron & Osherson, 2011; Bemis & Pylkkänen, 2013; Flick et al., 2018; Parrish & Pylkkänen, 2022), highlighting its importance in the composition of complex concepts across noun phrase types.

In line with this notion, Gagné and Spalding (2009) have proposed a unitary mechanism underlying the semantic composition of noun phrases. Their theory, which originally emerged in the conceptual combination literature (Gagné & Shoben, 1997), proposed that constituents are integrated through the formation of a relation that links the constituents. Additionally, multiple relational interpretations may be viable, and these various interpretations will compete during the meaning construction process (Gagné & Shoben, 1997). In order to determine the role of relational structures in the processing of familiar compound words, Gagné and Spalding (2009) found evidence in support of relational priming in a series of sense/nonsense judgment tasks, in which the response to a target word (e.g., SNOWBALL; a ball *made of* snow) was facilitated by a preceding prime that implemented the same relational structure (e.g., INKBLOT; a blot *made of* ink). In addition to the influence of conceptual relational structures, Gagné and Spalding (2009) found that psycholinguistic knowledge, including prior experience with a compound as well as prior experience with the constituent in a designated role (i.e., as a modifier or head), facilitated compound processing.

In a subsequent study, Spalding and Gagné (2014) extended their findings to semantically opaque compound words, where a meaning construction process would not necessarily yield an interpretation for the compound. That is, they conducted priming lexical decision tasks and found that the availability of distinct relational interpretations created competition among potential candidates, which in turn, hindered processing of the target compound (also see Schmidtke et al., 2016). Further, Ji et al. (2011) demonstrated a processing cost for the semantic composition of opaque compounds (also refer to El-Bialy et al., 2013). Taken together, Gagné and Spalding (2009) highlight the recruitment of conceptual combination during the processing of all noun phrases, whether familiar or novel, and hypothesize the implementation of prior knowledge of how concepts combine with other concepts within this process.

Notably, distributed semantic models have attempted to model language compositionality as well in noun phrase forms (e.g., Baroni & Zamparelli, 2010; Guevara, 2011; Günther & Marelli, 2019; Mitchell & Lapata, 2010; Vecchi et al., 2017). These models have been especially successful at modeling human language acquisition and behaviours, such as similarity judgments (Vecchi et al., 2017) and plausibility judgments (Günther & Marelli, 2016; Vecchi et al., 2011). Aside from the recruitment of language-based co-occurrences, embodied information, such as vision-based representations, has also been implemented in compositional models for concrete compound words (Günther et al., 2020). Thus, emerging research acknowledges the similarity in semantic representation and processing of compounds and combinations. Importantly, Mitchell and Lapata (2010) emphasize the importance of studying novel combinations, in addition to familiar compounds, to appreciate a full account of compositional processing. This is challenging for distributional semantic models, which rely on existing phrases in corpora. Thus, psychological theories of conceptual combination, which primarily aim to account for novel noun-noun combinations (e.g., CAT BOMB) and adjective-noun combinations (e.g., WISPY MEMORY) will be examined in the next section.

Theories of Conceptual Combination

Conceptual combination is a form of language compositionality, as it is an active meaning construction process used to simplify complex concepts in the environment. Like theories of conceptual representation, many psychological theories of conceptual combination assume that concepts have an underlying schematic representation that exhibit a causal role in the cognitive mechanisms involved in producing complex concepts (e.g., constructing a composite combination, slot filling approaches; Ran & Duimering, 2009). The majority of theories also converge on the notion that the two constituents asymmetrically contribute to generate the meaning of the complex concept, in which the first constituent is typically considered the linguistic modifier and the second constituent is labelled the head noun, or linguistic head (Ran & Duimering, 2009), similar to noun compounds (Spencer, 1991). Theories of conceptual combination elaborate on the role of each constituent in the processing mechanism and propose how the constituents interact to produce meaning that is consistent with an individual's prior experiences and context. The modifier constituent can be composed of different syntactic classes (e.g., noun or adjective) whereas the head constituent is typically a noun. Theories that describe first noun-noun combinations and then adjective-noun combinations will be reviewed, based on how the models were classified by Ran & Duimering (2010), although refer to Table 1 for a chronological summary of all theories.

Table 1

Overview of Psychological Theories of Conceptual Combination

Theory	Nature of representations	Constituent type	Interpretation strategy	Mechanism of interpretation	Familiar vs. novel	Concrete vs. abstract	Empirical Support
Fuzzy sets (Zadeh, 1965)	schematic representation	both	conjunction of constituents	overlap of extensional sets denoted by constituents with logical gradedness; applies to intersective combinations	familiar	concrete	Mathematical model in Zadeh (1965)
Amalgam Theory (Thagard, 1984)	schematic representation	both	slot-filling approach	six procedural rules that reconcile and modify the slots of the constituent components to be consistent with prior experience	both	not specified	None
Concept specialization model (Cohen & Murphy, 1984)	schematic representation	both	slot-filling approach	noun roles are activated by context and attribute values are filled based on prior background knowledge	both	concrete, although some stimuli were abstract	Medin & Shoben (1988); Murphy (1988, 1990); Springer & Murphy (1992)
Composite prototype model (Hampton, 1987)	schematic representation	noun-noun	conjunction of constituents	attribute inheritance from constituents to form a unifying composite prototype	familiar	concrete	Hampton (1985, 1987, 1988)
Selective modification model (SMM; Smith et al., 1988)	schematic representation	adjective-noun	slot-filling approach	serial process in which the adjective selects an appropriate slot in attribute of the noun, votes are shifted to increase salience of filled slot, and the diagnosticity of the modified attribute is increased	both	concrete	Smith and Osherson (1984), Smith et al., (1988)
Coherence Theory (Thagard, 1997)	connectionist network of associated concepts with schemas and relations	both	construct constraint network and apply connectionist algorithms	reconciling positive and negative coherence between concepts to maximize satisfaction of constraints	both	not specified	None
Competition among relations in nominals (CARIN; Gagné and Shoben, 1997)	statistical distributional knowledge of relational frequency in modifier constituent	noun-noun, including compounds	relational linking	16 thematic relations used based on modifier relational frequency, and competition is resolved by availability of relation	both	concrete, although some examples used were abstract	Gagné (2000, 2001); Gagné and Shoben (1997); Gagné and Spalding (2004); Estes et al., (2008)
Dual-process theory (Wisniewski & Gentner, 1991; Wisniewski, 1997)	schematic representation	noun-noun	relational linking, property mapping, and hybridization	scenario integration for relational linking and comparison and construction for property-mapping/hybridization; parallel processes with one selected based on plausibility, diagnositicity, and communicative context	both	concrete	Estes (2003); Wilkenfield & Ward (2001); Wisniewski 1996, 1997; Wisniewski & Love (1998), Wisniewski & Markman (1993);

							Wisniewski & Murphy (2005)
Constraint model (Costello & Keane, 2000)	not implied, but described as schematic representations with extensions to domain knowledge	noun-noun	conjunction, relational linking, property mapping, and hybridization	propose three constraints to account for polysemy and guide interpretation including diagnosticity, informativeness, and plausibility in a three-step serial interpretation process	both	concrete	Costello & Keane (1997, 2000, 2001)
Interactive property attribution model (Estes & Glucksberg, 2000)	schematic representation	noun-noun	slot filling approach applied to property mapping interpretations	head constituent supplies relevant dimensions, modifier constituent selects compatible attributes	novel	concrete	Bock & Clifton (2000); Estes & Glucksberg, (2000); Raffray et al. (2007)
Relational interpretation competitive evaluation (RICE; Gagné & Spalding, 2013; Spalding et al., 2010)	statistical distributional knowledge of relational frequency is accessed but not necessarily only tied to constituent conceptual representations	noun-noun, including compounds	relational linking	relational availability is accessed within the modifier constituent and compete for selection; potential candidates are evaluated in context of semantic and relational information in both constituents; the final interpretation is elaborated on	both	Concrete, although some examples used were abstract (e.g., student accusation)	Gagné (2000); Gagné & Spalding, (2009); Spalding & Gagné (2008); Spalding et al. (2010)
Interactional hypothesis (Maguire et al., 2010)	statistical distributional knowledge of relational frequency in modifier and head constituents	noun-noun	relational linking	relational frequency as a function of the interaction between modifier and head noun, based on the semantic categories of each constituent	both	concrete	Maguire et al., (2007); Maguire et al., (2010)
Embodied conceptual combination (ECCo; Lynott & Connell, 2010)	statistical linguistic distributional information as well as situated simulations	noun-noun	destructive and non-destructive processing	destructive processing involves deconstructing a representation of a concept whereas non-destructive processing leaves both concept representations intact; one process is preferentially selected based on context and plausibility	both	concrete and abstract*	Connell & Lynott (2011a, 2013)

*Only two noun-noun combinations were abstract of their 27 combinations (Connell & Lynott, 2011a; 2013)

Noun-noun Combinations

Early classification theories sparked research on the process of conceptual combination given its strong theoretical importance. Many theories focused on the *intensional* sets of concepts, or the attributes that are commonly shared by members of the same class of objects. Hampton (1987) proposed the formation of a composite prototype in semantic memory based on the interaction of the intensional sets of the constituents, which can be easily illustrated by the concept PET FISH. PET FISH contain some attributes of PET (e.g., *live in a home)* that are not generally true of FISH as well as some attributes of FISH (e.g., *have scales, breathe underwater*) that are not generally true of PET. Thus, all attributes defined within the intensional sets of PET and FISH are not maintained in the conjunction PET FISH. In Hampton's (1987) Composite Prototype Model, constituent concepts form a combined concept through a union of a set of attributes that are subject to necessity and consistency constraints. Attributes that are necessary to the constituents are kept in the composite prototype (e.g., *breathes underwater* is a necessary attribute for FISH, so this attribute remains in the concept PET FISH even though most pets breathe air) and conflicting necessary attributes between two constituents deem an implausible conjunction (i.e., an inheritance failure). Hampton's (1987, 1988) model is broadly defined within a class of models that propose that the conjunction of constituents, in which the product inherits a subset of constituent properties, is involved in the meaning construction process. This is similar to fuzzy set theory proposed by Zadeh (1965) and described in the adjective-noun section below, although emergent features present a challenge to these models. Remaining theories of noun-noun conceptual combination can be further classified by their proposed interpretation mechanism, including relational, property mapping, and dimension-based strategies.

Relational theories propose that identifying the relation evoked between constituent components derives the meaning of a novel noun phrase (Wisniewski & Markman, 1993). For example, to interpret the phrase SKUNK BOX, a relational theorist may propose that the <u>has</u> relation is evoked leading to the interpretation “a box that has skunks” The competition among relations in nominals (CARIN; Gagné & Shoben, 1997) is one such psychological model based on linguistic models that propose a taxonomy of relations between familiar noun-noun and adjective-noun combinations (e.g., Downing, 1977; Levi, 1978). CARIN proposes that prior statistical distributional knowledge (i.e., relational information) is stored with lexical entries and retrieved when interpreting complex concepts. Further, CARIN places emphasis on the modifier, rather than the head noun, in selecting the most appropriate relation during the process of conceptual combination. Gagné and Shoben (1997) proposed a taxonomy of 16 thematic relations with the intent of capturing most underlying relations between concepts, with some relations occurring more frequently for certain modifiers (e.g., the made of relation occurs most frequently for CHOCOLATE) resulting in competition amongst thematic types that is resolved by availability of a relation. The ease of interpretation is dependent on how noun constituents are used in other combinatory pairs, as some nouns have certain relational preference. Relation frequency of the modifier constituent was successful at accounting for reaction times in a sense/nonsense judgment task, supporting their model (Gagné & Shoben, 1997; although see Wisniewski & Murphy, 2005 for an alternative explanation of their findings). However, Gagné and Shoben’s (1997) model has been criticized for being too abstract (i.e., lacks precision) to capture the extent of diverse meaningful interpretations (Ran & Duimering, 2009, Wisniewski, 1997). For example, BIRTHDAY CAKE and BRAVERY MEDAL both rely on the for relation but overlook important differences between interpretations (cake used for a birthday; medal rewarded because of bravery).

Gagné (2000) extended CARIN to include a role for the head noun, in which it functions to assess the plausibility of the relational candidates within the modifier noun. Additionally, Spalding et al. (2010) proposed a formal extension and refinement of the CARIN model called the Relational Interpretation Competitive Evaluation (RICE) theory (Gagné & Spalding 2013; Spalding et al., 2010) which also encompasses noun-noun compounds, as alluded to in the prior section. With RICE, noun phrase interpretation proceeds via a “suggest evaluate elaborate” process occurring in parallel. Similar to CARIN, relational availability of the modifier constituent yields initial relational interpretations that compete for selection. To resolve competition, relational interpretations are evaluated in the context of the semantic and relational availability contained within both constituents. The final interpretation is then elaborated upon by recruiting extralinguistic knowledge, such as prior background information (Gagné & Spalding 2013; Spalding et al. 2010). Similarly, Maguire and colleagues (2010) postulated the Interactional Hypothesis, which makes two refinements to the CARIN theory. First, Maguire et al. (2010) suggest that relational preferences vary by semantic category and should be modeled within this framework, rather than at the level of individual concepts. Second, they suggest that relational availability and preference is a function of an interaction between modifier and head constituents, similar to RICE (Maguire et al., 2010).

Another strategy to interpreting novel noun phrases is called property mapping (also called attributive combinations; Wisniewski & Markman, 1993), in which a salient property of the linguistic modifier is asserted, or mapped, onto the linguistic head. As such, using a property mapping strategy, SKUNK BOX may produce an interpretation such as “a smelly box”. Wisniewski and colleagues (Wisniewski & Markman, 1993; Wisniewski 1996, 1997) found that this interpretation was employed frequently for combinations that had constituents composed from the same ontological category (e.g., CAR MOTORCYCLE interpreted as a motorcycle with four wheels). Wisniewski and Markman (1993) proposed an alignment mechanism to explain this process, in which commonalities and alignable differences are considered in the interpretation. For instance, both cars and motorcycles have wheels (i.e., a commonality) but cars have four wheels and motorcycles have two (i.e., an alignable difference). Hybridization, in which a new concept is formed through the amalgamation of constituents (e.g., a DONKEY HORSE as a new animal containing properties of both donkeys and horses), was considered an extreme form of property mapping, although it appears similar to the notion of the conjunction of constituents that was aforementioned (Hampton, 1987).

Wisniewski (1997) proposed a dual-process approach to account for the multiple interpretations derived from noun-noun combinations. After examining participants' explicit interpretations to numerous familiar and novel noun-noun combinations, he proposed three primary approaches to interpretation including relational linking, property mapping, and hybridization, with two distinct cognitive mechanisms. In relational-based interpretations, Wisniewski (1997) proposed a process of scenario integration, in which a plausible scenario is created to link the constituents. The scenarios coincide with verbs that contain roles to describe actions, events, or states. For instance, OIL may possess a *cooking* scenario with *recipient*, *agent*, and *instrument* roles to coincide with what is cooked, who cooked it, and what was used to assist in cooking. In contrast, a property-based or hybridization interpretation would utilize a process of comparison and construction, which can also be used to interpret nominal metaphors (Wisniewski, 1997). The comparison process is similar to that proposed by Wisniewski and Markman's (1993) structural alignment method described above, whereas the second construction phase instantiates a new property in the combined concept by inducing conceptual change. Conceptual change is constrained by ensuring to preserve the modifier that the property originated from while also maintaining the integrity of the head noun concept. For example, FORK SPOON would be interpreted as "a spoon with shortened fork prongs on the end of the little bowl" to ensure the integrity of the property being mapped from the modifier (i.e., prongs on a fork) as well as the functionality of the head noun (i.e., the little bowl to scoop things; Wisniewski, 1997). Wisniewski (1997) emphasized how plausibility, diagnosticity of the modifiers' features, and the communicative context are relevant in determining which interpretation strategy is utilized.

However, Gagné (2000) ran a series of online processing tasks (e.g., sense/nonsense judgment tasks) and concluded that property mapping strategies were not commonly used (i.e., were a last resort if a relational interpretation was available), and combinations that evoked property mapping interpretations in Wisniewski's (1997) offline processing task were deemed nonsensical in her task. In addition, Gagné (2000) proposed that both relational linking and property mapping could be captured in a single relation-based framework by including a "resemblance" relation or "object-like-subject" relation for property mapping interpretations. Wisniewski and Love (1998) tested this "last resort hypothesis" by using combinations that had plausible relations between the two constituents and found evidence against a serial processing model. Their results suggested instead that use of a property mapping technique was mediated by similarity of the constituent representations, could be flexibly applied when primed, and was commonly used in multiple text sources, although less so for artifact noun referents (Wisniewski & Love, 1998).

Acknowledging the multiple creative interpretations (i.e., polysemy) that conceptual combination can yield, Costello and Keane (1997) proposed a functional pragmatic account that imposed three constraints on the outlined strategies. These constraints include a diagnosticity constraint (i.e., able to be distinguished from related concepts and incorporates diagnostic attributes of both constituents), a plausibility constraint (i.e., consistent with prior experiences), and an informative constraint (i.e., refers to something novel that is not already captured by existing information), which are implemented serially in the meaning construction process. In their Constraint Model, the communicative context facilitates the appropriate interpretation. In support of their model, participants' comprehension and production of property-mapping complex concepts was predicted by diagnosticity, whether the attributes were alignable or not (Costello and Keane, 2001). For example, BUMBLEBEE MOTH yielded interpretations like "a moth that stings", which is a non-alignable (i.e., not related to commonalities between BUMBLEBEE and MOTH) but diagnostic property mapping approach.

Dimension-based models, also called slot-filling approaches, primarily dominate theories of adjectival modification described below, although some extend to noun-noun combinations (e.g., concept specialization model; Cohen & Murphy, 1984). These models have been described as a subset of both relational-based and property-based models within the literature. Slot-filling approaches embody the notion that the head concept has a frame or schematic representation (Minsky, 1975; Rumelhart, 1980) composed of attributes (i.e., slots with fillers) that become altered by the preceding modifier. Thus, interpretation of noun phrases proceeds by the modifier filling an attribute slot of the head. Returning to the SKUNK BOX example, the modifier SKUNK could occupy the <u>contents</u> attribute of the head noun BOX leading to the interpretation "a box containing skunks". Alternatively, it could occupy the <u>odour</u> attribute to yield an interpretation of "smelly box". Notice that dimension-based models derive interpretations that are identical to relation and property-mapping techniques, further supporting a single cognitive mechanism that parsimoniously underlies these different strategies.

Estes and Glucksberg (2000) applied a slot-filling model to guide property mapping interpretations called the Interactive Property Attribution Model. In contrast to the exhaustive similarity alignment process proposed by Wisniewski and Markman (1993), their model proposed a precise feature interaction process that yields property mapping interpretations by placing equal emphasis on both constituents in the mechanism. In their model, the head noun contributes pertinent dimensions whereas the modifier selects possible attributable candidates that are compatible with the head noun dimensions. Estes and Glucksberg (2000) also place emphasis on the linguistic context, as the salience of an attribute relies on the constituent it conceptually combines with. For example, FEATHER LUGGAGE can yield an interpretation of "light luggage" considering that the modifier has a characteristic property of being light and the head noun has a relevant weight dimension that can be altered. In contrast, FEATHER STORAGE is unlikely to yield a similar interpretation given that the head noun does not possess weight as a relevant dimension. Similarly, for the head noun BOX, odour may not typically be a relevant dimension until combined with the modifier SKUNK, which activates and alters this dimension to yield "smelly box". Although only property mapping interpretations were examined, they hypothesized the existence of a unitary schema model that contains storage of both featural and relational information in constituents to yield all interpretations (although see Wisniewski, 2000 for a re-analysis of their data that is consistent with the dual-process model).

Aside from schema-based models, Lynott and Connell (2010) proposed a more recent embodiment account of conceptual combination, called Embodied Conceptual Combination (ECCo). Similar to relational theories, conceptual representations contain linguistic distributional information about how concepts interact in language. However, perceptual simulation systems are also included in conceptual representations. For example, the concept SKUNK will activate associative linguistic information such as *animal, furry, stinks, black and white, striped*, as well as corresponding representations in the simulation system in their respective modalities. Similar to CARIN (Gagné and Shoben, 1997), relational information, or "affordances" (Gibson, 1977), characterize how concepts can mesh together in an interactive and complementary manner, although the ECCo model argues for the role of both constituents in this process. Based on how intact the concepts remain in the final coherent combination yields two distinct processing types, destructive or non-destructive, in which one processing type is committed to early on based on the meshed affordances and plausibility drawn from prior knowledge and experience. The former destructive type is akin to property mapping, or other construal interpretation strategies, whereas the latter non-destructive type is similar to relational linking (although Lynott & Connell, 2010 argue that relation linking is reflective of both types of processes). Their model is the first to extend to concepts where at least one noun constituent is abstract (e.g., ELEPHANT COMPLAINT, STRESS SEASON). Similar to mixed embodied accounts, they proposed the recruitment of the simulation system only for tasks that engage deeper processing tasks (e.g., offline interpretation tasks) but not for shallowing processing tasks (e.g., lexical decision, sense/nonsense judgments), and deemed this phenomenon the "linguistic short-cut hypothesis".

In support of their model, Connell and Lynott (2011a) used a meaningfulness task in which participants were instructed to respond “yes” if they could produce an interpretation and “no” if they could not. If a “yes” response was provided, participants were prompted to specify their explicit interpretation. They operationalized linguistic distributional information as a 5-gram frequency measure (i.e., the co-occurrence of words such as OCTOPUS and APARTMENT were calculated as the summed occurrence in a large corpus with zero, one, two, and three intervening words), which captures linguistic co-occurrence in local contexts, whereas evidence of recruitment for the perceptual system appeared to be reflected in the processing type (i.e., processing should be faster for non-destructive), although the mechanism of recruitment for perceptual inputs was unclear. They found that for yes responses, non-destructive interpretations yielded faster reaction times whereas for no responses, rejections times were negatively related to linguistic frequency (i.e., participants were faster to reject low frequency combinations), as predicted. In further support of their linguistic short-cut hypothesis, Connell & Lynott (2013) used a sense/nonsense judgment task and found that linguistic frequency was positively related to sense judgment times and negatively related to nonsense judgment times. Embodied theories in conceptual combination are additionally supported by research showing attentional switching costs when different modalities are used for adjective-nouns pairs (Connell & Lynott, 2011b) and with property generation tasks in which generated properties were modulated by effects of occlusion (Wu & Barasalou, 2009) and equivalent when neutral versus mental imagery instructions were used (Barsalou, Soloman, & Wu, 1999).

To summarize, conceptual combination in the context of noun-noun combinations has received much empirical attention. Many of the theories assert a schematic representation of each constituent concept, although some have incorporated statistical distributional information (e.g., CARIN, Gagné & Shoben, 1997; ECCo, Lynott & Connell, 2010) as well as perceptual simulations (i.e., ECCo, Lynott & Connell, 2010) into the nature of representations. The primary interpretation strategies include relational linking and property mapping, although theories diverge on the specific mechanisms of interpretation and can be subsumed under a dimension-based approach. Most theories aim to address both familiar and novel combinations. Like simple concepts, many theories of noun-noun combinations are targeted towards concrete complex concepts, although some may extend applicability to abstract concepts (e.g., CARIN, Gagné & Shoben, 1997; ECCo, Lynott & Connell, 2010). Yet, abstract complex concepts were underrepresented in stimulus sets and concreteness was not examined as a distinct variable, despite the mounting evidence suggesting distinct conceptual representations for concrete and abstract concepts (e.g., Huang and Federmeier, 2015; Huang et al., 2010; Pexman et al., 2007). In the next section, conceptual combination theories that extend to adjective-noun combinations are explored.

Adjective-noun Combinations

In noun-noun combinations, Wisniewski and Love (1998) state that the modifier noun can function as an adjective, particularly for construal interpretations, as the noun serves to identify a salient property in these instances. Nouns are thought to be more conceptually complex than adjectives (Murphy, 1990), and capture more descriptive information and knowledge in the modifier position (Wisniewski & Love, 1998). For instance, Wisniewski and Love (1998) state that in construal, compared to adjectives modifiers, nouns can denote the essence of the representation better (e.g., ZEBRA MOLLUSK versus STRIPED MOLLUSK), form a distinct category by referring to a specific referent rather than a general property (e.g., ELEPHANT SEAL and ELEPHANT GARLIC), convey a complex set of adjectives at once (e.g., APARTMENT DOG can refer to dogs that are both small and quiet), and ground experiences (e.g., OVEN DESERT can evoke perceptions of extremely hot and dry). Adjective-nouns are typically processed faster than noun-noun combinations, although this could in part be due to modifier nouns violating syntactic structure (Murphy, 1990).

Clearly, there are inherent differences between noun-noun and adjective-noun combinations, although the latter have been largely regarded as being conceptually simple. Adjectives do serve to express a feature of a noun concept (Ran & Duimering, 2009), but in contrast to the belief that adjectives can be captured by a value along a single dimension, some adjectives can reflect changes in multiple attributes (e.g., RIPE may modify colour, texture, smell, and taste; Medin & Shoben, 1988). In children, adjectives serve to organize lower-order categories (e.g., subordinate concepts) and emphasize differences between similar objects (Waxman & Markow, 1998). Additionally, adjectives do not refer to the same feature across noun contexts, and instead, can modify the connotation of the expression based on the noun (e.g., SLIMY MUD versus SLIMY PERSON) and create a united concept that is more complex than its' constituent components (e.g., *warm* and *outgoing* are unlikely to be salient characteristics of FRIENDLY COMPUTER; Medin & Shoben, 1988). Even simple adjectives that appear to be ubiquitous across contexts (e.g., colour and size adjectives), can have differences relative to one another (e.g., the *red* in RED CHERRY is qualitatively different than the *red* in RED LEAF) and have the capacity to capture distinct essences (e.g., GRAY is understood as indication of the aging process in the context of HAIR but as a storm condition in the context of CLOUDS; Medin & Shoben, 1988).

Adjectives can also be described as predicating, in which the adjective is sensical regardless of the position (e.g., slimy mud; the mud is slimy) or non-predicating if the adjectives are nonsensical when following the noun (e.g., corporate building; the building is corporate). Non-predicating adjectives have been shown to yield slower processing times compared to predicating adjectives (Murphy, 1990). Additionally, Kamp and Partee (1995) distinguished between subsective adjectives, in which adjectival modification is dependent on the noun (e.g., SKILLFUL) and intersective adjectives, which modify nouns ubiquitously (e.g., CARNIVOROUS). Drašković et al. (2013) conducted a speeded semantic classification task and found that participants were quicker to judge intersective combinations as meaningful compared to subsective combinations, which they attributed to the latter requiring more elaborate processing of the noun-related features. Many of the theories in this section account for both noun-

noun combinations and adjective-noun combinations. Only one model applies exclusively to adjective-noun combinations (i.e., Selective Modification Model; Smith et al, 1988).

Zadeh (1965) presents the first attempt to mathematically describe conceptual combination in his *fuzzy sets* theory. Briefly, this theory was applied to intersective combinations, which weigh both constituents equally in the interpretation from a logical standpoint (i.e., symmetric or true conjunctives). It is based on referential semantics which posits that words denote an extensional set of things that are captured in a concept's representation. In his model, comprehension of true conjunctives proceeds by considering the overlap in extensional sets denoted by the two constituents (e.g., RED APPLES are the overlap of categories of red things and apples; PET FISH are pets that are also fish). Fuzzy logic (Zadeh, 1965) proposed a method of modelling the logic of gradedness that constitutes the basis of categorization in prototype theory (i.e., similarity to prototype; Rosch & Mervis, 1975). The truth value of combinatory concepts can take a continuous value between 1 (true) and 0 (false). For example, the truth of the set "tall man" is dependent on the comparison to the height of a prototypical man. Zadeh (1965) proposed a minimum rule, in which the truth value of the combined concepts was the minimum of the two constituent truth values.

However, Osherson and Smith (1981) highlighted indisputable criticisms of the logic underlying *fuzzy sets* and demonstrated contradictory evidence of the minimum rule with examples of highly related concepts (e.g., X is a mammal animal) as well as examples in which the conjunction is a better prototypical example (e.g., X is a striped apple) than either constituent independently (i.e., X is striped, X is an apple). Smith and Osherson (1984) demonstrated empirical support for this latter criticism using participant typicality ratings and deemed the finding a "conjunction effect". Fuzzy logic does have some success in accounting for intuitions about the conjunction of two unrelated statements (Oden, 1977) and is able to model degrees of category membership in line with psychological studies of concepts (Cohen & Murphy, 1984). However, the extensional view of concepts applies to words that have confined extensional sets (e.g., descriptive notions like triangular and red), but it's difficult to extend to intricate concepts (e.g., belief), and does not apply to non-intersective combinations (e.g., non-predicting adjectives paired with nouns; Murphy 1988).

To address the shortcoming of *fuzzy sets* and provide empirical support for the claims made by Osherson and Smith (1981), Smith et al. (1988) proposed the Selective Modification Model (SMM), the only model developed to solely describe adjective-noun combinations. The SMM is similar to the Composite Prototype Model (Hampton, 1988) described above, in that conceptual representations include intensional sets and constituent concepts conjoin to form a composite prototype, with adjectives acting to modify the noun concept. Noun schematic representations consist of properties based on one's knowledge, which are decomposed to reflect attribute-value pairs (i.e., also called slot-filler pairs), with "slot fillers", in this case adjectives, occupying the value entity (Smith et al., 1988). For example, an APPLE has attribute-value pairs such as colour-red, shape-round, and taste-sweet. Conceptual representations also include an associated weight (i.e., also called salience or votes) for each value, with some values having more salience relative to others based on prototypicality, or similarity to the prototype (e.g., *red* is a more salient value for APPLE than *round*), and some values having more relevance relative to other concepts (e.g., *red* of an apple is more salient than *red* of a brick). The salience of the value is thought to be a function of subjective frequency and perceptibility (Smith et al., 1988). Lastly, conceptual representations contain information about the diagnosticity of attributes for noun concepts, or how valuable the attribute is at differentiating the noun concept from contrasting concepts (e.g., the *shape* attribute distinguishes APPLES and BANANAS, which are both members of the same category, fruit). Thus, the SMM (Smith et al., 1988) embodies the contrast rule (Tversky, 1977) as well as the notion of similarity to prototype (Rosch & Mervis, 1975) and applies it to adjectival modification of nouns.

To interpret a complex concept in SMM (Smith et al., 1988), a filler (i.e., adjective) occupies the slot of a head concept and acts as a modifying concept. They proposed that the selection and modification processes occur in three serial steps. First, the adjective fills the relevant attribute of the noun concept. Once the attribute is modified, “vote shifting” occurs, in which salience of the value occupied by the adjective increases. Finally, the diagnosticity of the modified attribute will be enhanced. For example, in the modified concept RED APPLE, the adjective RED selects the colour attribute of the noun concept APPLE and shifts votes towards red and away from other potential slot fillers, such as green and brown. As a result of this process, the colour attribute will have increased diagnositicity for the modified concept RED APPLE. In support of their model, Smith et al. (1988) demonstrate how it can predict typicality ratings for property generation tasks for adjective-noun conjunctive concepts and draw on their earlier demonstration of the conjunction effect based on such ratings (Smith & Osherson, 1984).

However, the SMM has received much criticism since its inception. For example, Murphy (1988) noted that the model is over-specified, as the second and third steps could be subsumed under one process. Additionally, the model cannot extend to more complex concepts, such as non-predicating adjectives, or noun concepts which may not have a distinct dimension for the modifier to act on (e.g., abstract concepts; Murphy, 1990). Further, the model propagates the notion that adjectives have a simple conceptual structure (Murphy, 1988), and assumes a direct mapping between value and attribute, when in fact a single value (e.g., BROWN) may influence multiple correlated attributes (e.g., colour, shape, smell, and taste; Medin & Shoben, 1988; Smith et al., 1988).

For instance, Springer and Murphy (1992) found that true phrase statements (e.g., *peeled apples are white*) were verified more quickly and accurately compared to true noun statements (e.g., *peeled apples are round*), with no difference found between false statement types (e.g., *peeled apples are red*; *peeled apples are square*) providing evidence against a serial model that predicts activation of noun features prior to constructing features of the combination. Similarly, Medin and Shoben's (1988) findings demonstrated that participant typicality ratings reflected sensitivity to correlated attributes (e.g., WOODEN SPOON affected the expected size of the spoon), noun context (e.g., BRASS and GOLD RAILING rated as most similar due to colour whereas SILVER AND GOLD COIN rated as most similar due to value), and property centrality (e.g., shape attribute may be more central to BOOMERANG than BANANA due to dependence on functionality). They conclude that prototype- or feature-based models of concepts that assume stored rigid instances cannot explain the increasing complexity and context dependence of adjective-noun combinations, in which recruitment of prior exemplars and background knowledge is needed (Medin & Shoben, 1988). Nonetheless, Smith et al. (1988) qualify that their model is better suited at determining typicality of simple adjective-noun concepts rather than a model capable of categorizing complex categories.

Thus, empirical evidence supports a contextual view of conceptual combination, wherein the noun supplies the context to aid in accurate retrieval of emergent attributes and avoids activation of context inappropriate features (Maguire et al., 2007). Adjectives can serve to resolve ambiguities in homonyms almost instantaneously. For example, WOODEN TABLE would direct the interpretation for a piece of furniture as the referent, rather than a data-organizing tool. Likewise, Potter and Falcouner (1979) found that adjectives and nouns interact automatically when a sentence is presented, suggesting holistic retrieval of the noun phrase facilitated through spreading activation, rather than independent activation of the constituent components. However, Potter and Falcouner (1979) noted that this latter slower mechanism may be recruited for novel combinations (e.g., FURRY UMBRELLA).

In contrast to the SMM, the Concept Specialization Model (Cohen & Murphy, 1984) incorporates prior background knowledge in directing the selection of the ideal slot to modify in head nouns, although it adopts a similar notion of noun modification via slot-filling. In their model, concepts are schematic representations within a hierarchical and flexible category network, and attribute values are weighted by typicality derived from knowledge of family resemblance within concept categories (Cohen & Murphy, 1984). Further, interpretation of complex concepts is constrained by the context, and prior knowledge and experience contribute to an elaborated interpretation that is not captured by the individual constituents (i.e., consistent with the Theory View; Murphy & Medin, 1985). For example, APARTMENT DOG would be interpreted by selecting the slot that apartment would best modify based on prior experience (i.e., habitat slot) and further elaboration would identify additional features such as cleaner, quieter, and smaller than other dogs.

In support of this model, Murphy (1988) used predicating adjective-noun pairs in a typicality rating task and an explicit interpretation task. Concept elaboration was supported by emergent features (e.g., *lose money* for EMPTY STORE) and context dependence of the noun. For example, OPEN YEAR was interpreted as flexible whereas OPEN EYE was interpreted as alert. Thus, Murphy (1988) concluded that meaning construction was derived through an interaction between the constituent components. In a series of meaningfulness judgment tasks, Murphy (1990) found that novel noun-noun combinations (e.g., PARK OLIVE) yielded slower meaningful decision times than adjective-nouns where the adjective modified a slot atypically (e.g., SWEET OLIVE). These combinations, in turn, were processed more slowly than adjective nouns that modified a slot in a typical manner (e.g., SALTY OLIVE). When exploring adjective-nouns further, non-predicating adjective-noun pairs did not reliably differ in response times from atypical adjective-noun pairs, although both were processed slower than typical adjective-noun pairs. In a post-hoc explicit interpretation task, participants produced more meaningful interpretations for noun-noun and non-predicating adjective-noun phrases compared to predicating adjective-noun phrases. Murphy (1990) concluded that noun-nouns and non-predicating adjective-nouns were conceptually richer and proposed that a comprehensive model needs to address the structural differences between combination types. Murphy (1988, 1990) acknowledged that the Concept Specialization Model incorporates the richness of conceptual representations that was lacking in the

SMM, but he noted that the model still fails to specify how information is accessed and is limited in extending to combinations that do not modify an intuitive slot in the head noun (e.g., noun-nouns, non-predicating adjective-nouns).

Lastly, philosopher Paul Thagard (1984) proposed the Amalgam model which adopts a schematic representation similar to the above models, with concepts defined as frames consisting of data structures as slots (Minsky, 1975). Concepts were conceptualized as akin to stereotypes, with default values based on expectation and experience. His theory proposed six procedural rules to reconcile conflict between features and prior instances of constituent components in order to guide interpretation for the new combined concept. Within these rules, he acknowledged resolution arising from both context-independent (i.e., pure) elements existent within the constituent concepts, from those derived from context (e.g., data-driven or goal directed). As an example of a goal-directed concept, Thagard (1984) attributed the development of scientific concepts (e.g., RADIO WAVE, NATURAL SELECTION) to conceptual combination, as these novel concepts aim to capture their explanatory scope. However, Thagard (1984) did not specify the cognitive mechanisms that underlie his rules, nor did he elaborate on how novel combinations can be produced in the absence of prior instances.

In an extension of his model, called the Coherence Theory, Thagard (1997) proposed that concept attributes have the ability to either cohere (i.e., fit together) or incohere, and reconciling the coherence problem produces a combined concept that maximizes the overall satisfaction of the constraints. As opposed to assuming that conceptual representations exist within a schematic framework, Thagard (1997) proposed a connectionist network which implements an algorithm to identify positive and negative constraints in the concepts to achieve coherence. Positive constraints are exemplified by concepts with statistical or causal relations between them whereas negative constraints arise from contradictions or weak negative correlations. Balancing these associations to maximize the potential satisfaction of the constraints achieves coherence. Within this model, Thagard (1997) propagates the assumption of harmony or goodness of fit borrowed from Gestalt psychology, although he acknowledged that his extended model fails to account for non-predicating combinations and some novel combinations (e.g., WEB POTATO) which may only be understood through failure to find coherence between the concepts. In these instances, Thagard (1997) proposed the implementation of broader thematic relations with recruitment of more elaborate semantic, factual, and contextual knowledge, or reliance on higher-order cognitive strategies such as analogy (i.e., essentially proposing an untestable model).

In summary, slot-filling approaches are thought to be involved in the meaning construction of adjective-noun combinations. Therefore, most models that include adjective-noun combinations instantiate a schematic prototype of the head noun, with the adjective acting to modify the head noun concept. Schema-based models, however, primarily assume that interpretation is directed by the content of the constituent concepts, rather than prior knowledge about their history of use in combinations (Spalding et al., 2010). Additionally, they model conceptual change in the head noun schematic concept only, whereas statistical distributional representations exploit knowledge about relational frequency for modifier and head constituents (e.g., RICE; Gagné & Spalding, 2013) as well as the co-occurrence between the constituents (ECCo; Lynott & Connell, 2010). Schematic conceptual combination models incorporate a role for context and the integration of prior background knowledge to inform interpretations, although the mechanism is unclear (Maguire et al., 2007). Moreover, these models are generally limited to adjectives paired with concrete concepts. Collectively, conceptual combination appears to involve the recruitment of several cognitive processes involved in identifying combinable features, selecting features for transfer, integrating features into a unitary representation, and assessing the plausibility of the interpretation (Coutanche et al., 2019).

Overview of the Present Study

Familiar and novel adjective-noun combinations warrant further investigation and in the current study that investigation goes beyond “simple” adjective constituents (e.g., colour, shape; Smith et al., 1988). Rather than assuming a schema-based model, incorporation of distributional semantic information into concept representations appears to be a promising avenue in the field of noun compounding and noun-noun conceptual combination. These statistic-based models have yielded accurate predictions on several psycholinguistic tasks of other complex noun phrases including priming experiments (e.g., relational priming; Gagné, 2001; Gagné & Shoben, 2002; Gagné & Spalding, 2009), lexical decision tasks (Gagné & Spalding, 2004; Estes & Jones, 2008), and sense/nonsense judgment tasks (Connell & Lynott, 2011a). Further, these empirical pursuits have revealed that the processing of noun phrases is sensitive not only to variables linking the constituents, such as 5-gram frequency co-occurrence (Connell & Lynott, 2011a) and typicality (Murphy, 1988; 1990), but also its’ constituent composition, including modifier and head frequency (Andrews et al., 2004; Kuperman et al., 2009), relational availability (Gagné & Shoben, 1997; Maguire et al., 2010), and semantic transparency (Libben, 2003). Given these findings in other noun phrases, it is also likely that the processing of adjective-noun combinations is sensitive to both its constituent and phrase composition.

Relevant to this dissertation is the concept of semantic richness, which refers to the amount of variability contained within a concept’s meaning (Pexman et al., 2008). Concepts that are associated with large amounts of semantic information are deemed semantically rich and are often associated with quicker response times compared to semantically sparse concepts (e.g., Pexman et al., 2008). Semantic neighbourhood density (SND; Durda & Buchanan, 2008), is one measure of semantic richness derived from language co-occurrence models. As described earlier (refer to Figure 1; Danguecan & Buchanan; 2016), SND captures the variability of semantic neighbours within a concept’s semantic neighbourhood, with dense SNDs occurring in concepts with many close semantic neighbours on average and sparse SNDs demonstrated in concepts with loose associations with their semantic neighbours. As stated, SND has been shown to interact with concreteness, causing an inhibitory effect for abstract simple concepts on online processing tasks (Danguecan & Buchanan; 2016) as well as an inhibitory effect in concrete metaphors on an offline processing task (Al-Azary & Buchanan, 2017). Thus, investigating how the SND constituent composition of adjective-noun combinations affects constituent integration is another aim of the current study.

Furthermore, few theories of conceptual combination incorporate concreteness into their models despite the growing literature base for unique conceptual representations for concrete and abstract concepts (Borghi et al., 2017). As abstract concepts are increasingly complex and lack an identifiable referent (Borghi and Binkofski, 2014; Borghi et al., 2017; Schwanenflugel, 2013), language is thought to play a critical role in abstract concept representation (Dove, 2009; Binder, 2016). Most theories of conceptual combination have primarily focused on concrete concepts. In noun-noun phrases, Lucas et al. (2017) studied a variable related to concreteness called imageability (i.e., how easily a concept can evoke a mental image). Their study investigated how imageability of the modifier noun affected the ease of interpreting noun-noun phrases with concrete head nouns and recorded brain activity with an electroencephalogram (EEG) while participants rated the ease of interpretation. They found that participants rated noun-noun phrases with a higher imageable modifier noun as easier to interpret than those with a lower imageable modifier noun. The latter noun types were also associated with a larger N700 potential, which they interpreted as suggestive of the recruitment of an imagery-based strategy when interpreting higher imageable noun-noun phrases (Lucas et al., 2017). Thus, examining the concreteness (and abstractness) in the conceptual combination of adjective-noun phrases, along with its interactions with SND, is an additional aim of this study.

Lastly, the stimuli and methodology employed in the current study are reflective of both psycholinguistic and psychological endeavours in order to provide an interface for these two literature bases. The meaningfulness of adjective-noun combinations, defined as the ease of meaning construction, is used as a proxy for the degree of familiarity and plausibility of the phrase. Adjective-nouns were stratified into low, intermediate, and high meaningfulness to represent a range of novelty with the stimuli. Further, the methodology spans online processing tasks with traditional psycholinguistic tasks (i.e., lexical decision tasks) and those that recruit deeper semantic engagement (i.e., sense/nonsense judgment task), as well as an offline processing task (i.e., explicit judgment task). Such tasks have produced disparate results in the literature (e.g., Günther et al., 2020), potentially due to distinct processing mechanisms. A complete theory of conceptual combination requires a comprehensive understanding of processing across differential task demands and this dissertation will move us closer to that understanding.

This dissertation serves as an extension of McAuley (2018) which also examined semantic effects in adjective-noun pairs. In McAuley (2018), adjectives paired with concrete head nouns were processed faster on lexical decision and sense/nonsense judgment tasks than adjectives paired with abstract head nouns. Importantly, this effect was modulated by the plausibility of the adjective-noun pair, with high and intermediate plausible adjectives paired with concrete head nouns yielding faster processing times than adjectives paired with abstract head nouns, although no differences were found for implausible pairs based on head noun concreteness. Additionally, no differences were observed in processing times between intermediate and implausible adjective-noun pairs with abstract head nouns. Moreover, an adjective paired with a noun that was a close semantic neighbour facilitated processing times on a sense/nonsense judgement task, although this was not found in double lexical decision tasks. In contrast, distant and unrelated adjective-noun pairs did not yield a processing advantage across tasks. This dissertation expands on these findings by modifying the operational definition of concreteness and plausibility (i.e., labelled meaningfulness) through collection of a large database of subjective participant ratings using a novel quantitative measure. Further, semantic neighbourhood density of constituent concepts was examined, rather than the semantic distance between adjective and nouns within a pair.

Lastly, an explicit judgment task was included to span both online and offline processing tasks. A more in-depth discussion of the present study is described in the following section.

CHAPTER 2

DESIGN AND METHODOLOGY

Research Objectives and Hypotheses

The research objective of this study is to examine the semantic processing of adjective-noun combinations as an extension of McAuley (2018). Within this scope, semantic variables related to the constituent and phrase composition were manipulated to measure the impact on a meaning construction process (i.e., conceptual combination). The semantic neighbourhood density (SND) of adjective and noun constituents was derived from a semantic distributional co-occurrence model (Windsor Improved Norms of Distance and Similarity of Representations of Semantics: WINDSORS; Durda & Buchanan, 2008). The novelty and thus the familiarity of the adjective-noun phrases was stratified into low (e.g., CLUMSY REALM), intermediate (e.g., SPACIOUS REALM), and high (e.g., MAGICAL REALM) meaningfulness. These adjective-noun phrases varied in meaningfulness and concreteness with ratings for both obtained from participants using a novel quantitative measure in Experiment 1: Stimulus Development. The processing of adjective-noun combinations was evaluated in a series of tasks ranging from shallow (i.e., Experiments 2 and 3: Lexical Decision Tasks) to deep (i.e., Experiment 4: Sense/nonsense Judgment Task) online processing tasks to yield response times and error rates. In addition, participants were asked for their explicit interpretation of novel adjective-noun combinations, or those of low meaningfulness, on an offline comprehension task (refer to Experiment 5: Explicit Judgment Task) to identify underlying strategies gathered from interpretations. Taken together, the following hypotheses are tested:

H1: High meaningful adjective-noun combinations will yield faster response times than intermediate meaningful combinations, which will yield faster response times than low meaningful combinations (Experiments 2-4).

Based on a re-analysis of empirical findings by Wisniewski and Murphy (2005), familiar and plausible noun-nouns were processed faster than unfamiliar and implausible noun-noun combinations in a sense/nonsense judgment task. Additionally, using the same task, Murphy (1990) found that nouns paired with an adjective reflecting a typical attribute were processed faster than nouns paired with an adjective indicating an atypical attribute. In the current study, familiarity with the adjective-noun phrases and plausibility of the adjective-noun phrase are thought to be two components captured in meaningfulness (i.e., the ease of meaning construction) ratings. Thus, a similar pattern of response times was expected. Further, when a crude measure of plausibility was used in McAuley (2018), no differences were found in processing times for intermediate and low plausible pairs, although high plausible pairs were processed faster than the former two. The current study uses a more precise measure of meaningfulness to determine the sensitivity of response times as a function of the meaningfulness of the combination. This finding was expected to be most robust for sense judgments in Experiment 4, which requires the deepest level of semantic processing.

H2: Concrete adjective-noun combinations will be processed faster than abstract adjective-noun pairs (Experiment 2-4).

A concreteness effect, in which concrete concepts are processed more quickly than abstract concepts, is a robust finding in the psycholinguistic literature (Kroll & Merves, 1986; Paivio, 1971) and has been found in shallow and deep processing tasks. McAuley (2018) found that adjectives paired with concrete head nouns yielded faster response times and fewer errors than adjectives paired with abstract head nouns across all tasks. In the present study, rather than examining concreteness for only head noun constituents, concreteness was measured at the phrase level. Intuitively, as adjectives serve to label a specific referent and distinguish among related concepts (Medin & Shoben, 1988; Ran & Duimering, 2009; Waxman & Markow, 1998), this may promote concretizing of the noun by identifying specific features. For instance, the concept PILLOW is concrete, given there is a physical referent associated with sensory features, but FLUFFY PILLOW provides an even more elaborate description of the referent. Similarly, IDEA is an abstract concept, but CLOUDY IDEA evokes a more precise referent.

H3: The above effects may interact, with concrete combinations yielding faster processing times than abstract combinations of high and intermediate meaningfulness, with no differences found for low meaningful combinations based on concreteness.

Hypothesis 3 is entirely motivated by McAuley (2018), who found this effect using similar measures.

H4: Modifier and head noun constituents with denser semantic neighbourhoods will yield faster response times than modifier and head noun constituents with sparser semantic neighbourhoods (Experiment 4).

Semantic richness, defined as the extent of the variability in the meaning of a concept, has been shown to facilitate processing times in simple concepts (Pexman et al., 2008). Borrowing from a related field, Kintsch (2000) proposed a predication algorithm to compute the meaning of metaphors, such as *A Pen is a Sword*, in which PEN is the topic and SWORD is the vehicle. Specifically, Kintsch's (2000) model predicted that the vehicle's semantic neighbourhood is activated through a spreading activation mechanism, which selects relevant semantic neighbours and inhibits irrelevant semantic neighbours in the context of the topic. Thus, a denser semantic neighbourhood was predicted to facilitate this process (Kintsch, 2000) and may similarly do so with adjective-noun pairs. Semantic neighbourhood density (SND) was used as a measure of semantic richness in the present study, derived from WINDSORS (Durda & Buchanan, 2008) using the semantic neighbourhood app (Lutfallah & Buchanan, 2018). SND effects are thought to be particularly prominent in the head noun constituent, as the head noun identifies the semantic category of the phrase (Gagné & Spalding, 2009; Spencer, 1991), although modifier constituent effects have been found in noun-noun combinations (Estes & Glucksberg, 2000; Spalding et al., 2010). Further the present study aimed to delineate the effect of the modifier SND from the head noun SND within each meaningful group, as each head noun will be paired with three modifying adjectives to constitute the high, intermediate, and low meaningful combination types. SND effects were expected in Experiment 4, the deepest processing task.

H5: SND of the constituents will interact with concreteness (Experiment 4).

Language is thought to be critical to abstract concept representation (Binder, 2016; Dove, 2009) whereas concrete concepts may incorporate information from a variety of sources (e.g., sensory-motor features; Wiemer-Hastings & Xu, 2005). Therefore, abstract adjective-noun combinations may be more sensitive to language-based measures. Two studies have demonstrated the sensitivity of abstract concepts to SND. First, online processing tasks with simple concepts has demonstrated an inhibitory effect for high SND abstract nouns compared to low SND abstract nouns, with no differences found in concrete concepts based on SND (Danguecan & Buchanan, 2016). On the other hand, Al-Azary and Buchanan (2017) found a facilitatory effect for high SND metaphors with abstract topics, which were rated as more comprehensible and processed faster than concrete high SND metaphors, with no differences found for low SND metaphors varying on topic concreteness. As such, examining the interactions between concreteness and SND of constituents was of interest. An interaction between concreteness and SND was expected to be most robust in Experiment 4, though interactions have been found in single lexical decision tasks as mentioned (Danguecan & Buchanan, 2016). Also of interest is whether any of the hypothesized effects outlined above are modulated by task demands as has been shown for noun-noun combinations (e.g., Gagné & Shoben, 1997).

Regarding the offline comprehension task, which only uses low meaningful (i.e., novel) adjective-noun combinations, the general hypothesis was as follows:

H6: More unique interpretations will be produced for adjective-noun combinations with abstract head nouns relative to adjective-noun combinations with concrete head nouns. Constituents with high SNDs will also yield more unique interpretations.

Novel concepts have been shown to yield more creative interpretations than familiar concepts (Murphy, 1990) and superordinate referents (e.g., ANIMAL) yielded more alternative interpretations than basic-level concepts (e.g., DOG; Costello & Keane, 2000) in noun-noun combinations. Relative to concrete concepts, abstract concepts are inherently less stable, lack identifiable referents, and are shaped by personal experiences (Borghi & Binkofski, 2014; Schwanenflugel, 2013) whereas concrete concepts are defined by their intrinsic properties (Wiemer-Hastings & Xu, 2005). Thus, combinations with an abstract head noun are likely to have multiple alternative interpretations. Similarly, a high SND constituent composition provides a greater network of information from which to draw knowledge to implement into the final product.

Aside from the quantity of interpretations, the quality of the interpretations was also examined in an exploratory fashion. Similar to noun-noun combinations, it is likely that adjective-noun combinations benefit from various interpretive approaches. However, the exact nature of these interpretations, and whether they are akin to those employed when interpreting noun-noun combinations, has not yet been studied. Property-mapping for instance, can likely apply to adjective-noun combinations, because some (but not all) features of the adjectives can be attributed to the noun. For instance, the features characteristic of FRIENDLY map on differently when the referent is a person (e.g., *warm, outgoing*) versus a computer (e.g., *user-accessible, simple formatting*). Thus, underlying themes of interpretation strategies were explored in the present study.

In summary, a series of online experimental processing tasks were conducted to examine the constituent composition (i.e., SND) and phrase composition (i.e., concreteness and meaningfulness) to evaluate their independent and interactive effects in the processing of adjective-noun phrases. Further, an offline comprehension task was used to identify strategies used to interpret novel adjective-noun phrases containing different semantic properties. Collectively, the findings are considered in light of current theories of conceptual combination.

Operational Definitions

Meaningfulness

Meaningfulness refers to the ease of meaning construction when encountering an adjective-noun combination and is thought to capture both familiarity and plausibility with the phrase. Meaningfulness ratings were obtained through participant ratings using a quantitative measure ranging from -2 to +2 (see Experiment 1a: Stimulus Development). Meaningfulness ratings were stratified into high, intermediate, and low meaningful adjective-noun pairs to create a categorical variable. Categorical variables for familiarity and plausibility are common in the field (e.g., Murphy, 1990; Wisniewski, 1997; Wisniewski & Murphy, 2005), and in the current case, it allows for inclusion of an intermediate group rather than solely comparing extreme ends. High meaningful pairs were operationalized as those with mean values above .75, low meaningful as pairs as those below -.75, and intermediate as those between -.75 and .75. For example, FIZZY SODA represented a high meaningful pair (*M*=1.73, *SD*=.81), TENSE SODA reflected a low meaningful pair (*M*=-1.70, *SD*=.81), and DILUTE SODA was rated as intermediate (*M*=-.24, *SD*=.94). Each head noun was paired with three adjectives that varied in low, intermediate, and high meaningfulness based on participant ratings.

Concreteness

Concreteness reflects a continuum based on whether a concept is more experience-based, or a physical entity associated with sensory experiences (i.e., concrete), or more language-based, in which the concept cannot be experienced directly and instead the meaning is defined through its association with other words (i.e., abstract; Brysbaert et al., 2014). Concreteness ratings were also obtained through participant ratings using a quantitative measure ranging from -2 (more abstract) to +2 (more concrete; see Experiment 1b: Stimulus Development). For example, RAW CHAOS was rated as more abstract (*M*=-1.32, *SD*=0.90) and CRUNCHY PASTRY was rated as more concrete (*M*=1.60, *SD*=.74). Despite this continuum, the ratings yielded a bimodal distribution and supported the utility of categorizing a concrete/abstract variable. A mean split was performed (*M*=0.00), and those below the mean were categorized as abstract words and those above the mean were categorized as concrete words for Experiments 2 to 4.

For the explicit judgment task (Experiment 5), concreteness was examined at the constituent level rather than at the phrase level. Concreteness of adjectives and head nouns were obtained from ratings collected by Brysbaert et al. (2014). Their concreteness ratings range from 1 to 5, with higher ratings reflecting more concrete concepts. To yield a categorical variable, concepts with mean ratings equal to or below 3.00 were considered abstract and those above 3.00 were considered concrete. For instance, RISKY LUGGAGE has an abstract adjective (*M*=1.31) and concrete head noun (*M*=4.83), whereas CLOUDY DILEMMA has a concrete adjective (*M*=4.00) and an abstract head noun (*M*=2.00). In total, there were 17 concrete-concrete pairs, 18 abstract-concrete pairs, 13 concrete-abstract pairs, and 15 abstract-abstract pairs.

Semantic Neighbourhood Density

Semantic neighbourhood density (SND) is defined as the variability in the distribution of semantic neighbours around a target word's semantic neighbourhood and acts as a proxy for semantic richness. SND values were obtained for each adjective and noun constituent (for its closest 25 neighbours) from a global co-occurrence model (WINDSORS; Durda & Buchahan, 2008) using the semantic neighbourhood app created by Lutfallah and Buchanan (2018). SND was used as a continuous variable in Experiments 2 to 4, with values ranging from .31 (sparser SND) to .82 (denser SND). For example, BLANK CANVAS has a sparser SND for the adjective (.38) relative to the head noun (.53) whereas the opposite pattern is seen for SHINY PEBBLE (.61 and .42, respectively). Refer back to Figure 1 for examples of sparse and dense semantic neighbourhoods.

For Experiment 5, constituent SNDs were categorized into "sparse SND" and "dense SND" based on a median split of the adjective and noun SND (both *Med's*=.53). Constituents equal to and below the median were categorized as "sparse SND" and those above .53 were categorized as "dense SND". For example, HILLY BANNER had a sparse adjective SND (.38) and noun SND (.42) whereas SLEEPY VACCINE had dense SND constituents (.59 and .78, respectively). In total, there were 18 dense-dense SND pairs, 12 sparse-dense SND pairs, 13 dense-sparse SND pairs, and 20 sparse-sparse SND pairs.

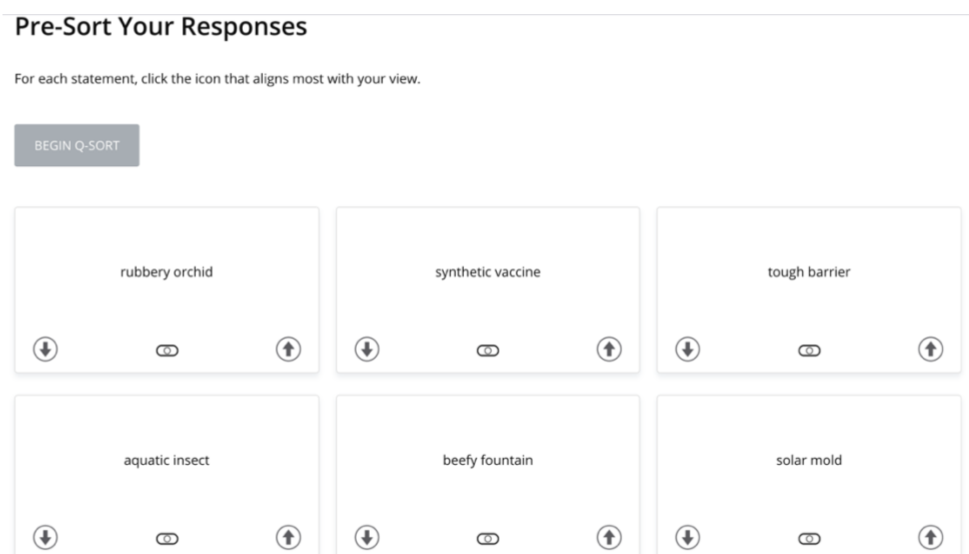
Methodology Overview

Experiment 1: Stimulus Development

The stimulus set was developed via Q-sort methodology to quantify participants' subjective ratings using a computerised web-based application (QMethod Software, Lutfallah & Buchanan, 2019). Once participants initiate the app they are presented with the instructions of the task and then enter a "Pre-Sort" stage, which allows them to view the stimuli on cards and make judgments about the stimuli. For example, for the meaningfulness rating task, participants were asked to judge whether the adjective-noun pair was high meaningful by clicking an "up arrow" icon, intermediate meaningful by clicking the "middle" icon, or low meaningful by clicking the "down arrow" icon. See Figure 2 for an example of the pre-sort stage.

Figure 2

The Pre-Sort on QMethod Software

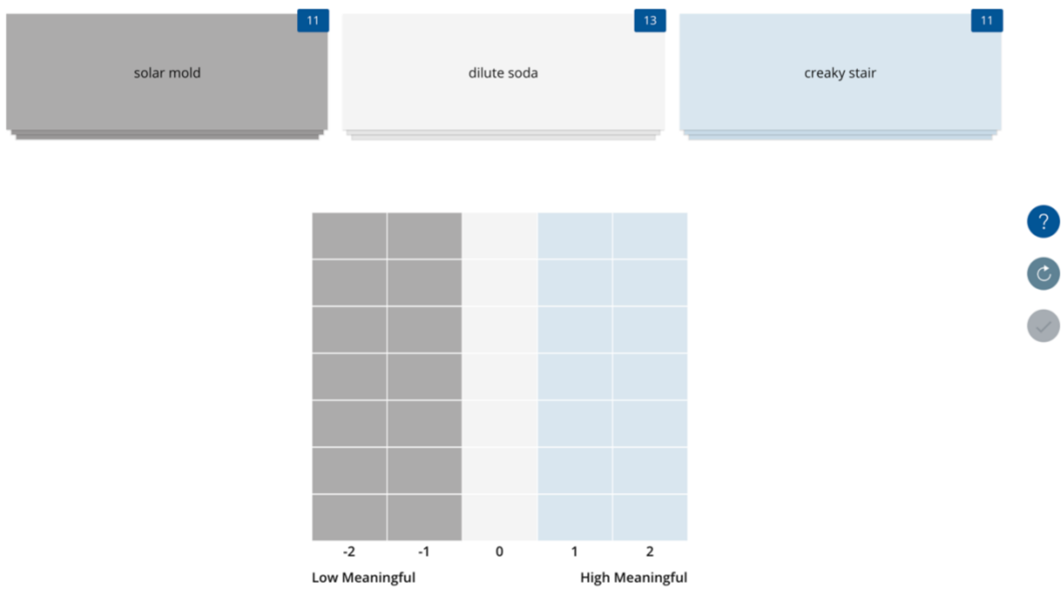


Note: An example of the pre-sort stage in QMethod Software from the meaningfulness rating task.

Once completed, participants were automatically directed to the "final-sort" page, where they were instructed to drag and drop their adjective-noun pre-sorted cards along a flat 5-column distribution ranging from -2 (low meaningful) to +2 (high meaningful). In this way, participants could rank the meaningfulness of an adjective-noun pair against all other adjective-noun pairs and adjust their ratings accordingly. See Figure 3 for an example of the final-sort stage.

Figure 3

The Final Sort on QMethod Software



Note: An example of the final-sort step in QMethod Software from the meaningfulness rating task.

Once all adjective-noun pairs were sorted along the distribution, participants submitted their responses. The collective ratings submitted by all participants was exported to produce an output file that assigns a numeric value, ranging from -2 to +2, to reflect the meaningfulness of each adjective-noun pair rated by each individual participant. Q-sort methodology was also used to collect concreteness ratings, with -2 reflecting more abstract entities and +2 reflecting more concrete entities.

Experiments 2 and 3: Double Lexical Decision Tasks

In a standard lexical decision task, participants are presented with a single string of letters, and they decide whether the letter string is a real word or not. In a double lexical decision task (DLDT), participants are presented with two letter strings and are asked to determine whether both letter strings are real words (e.g., FLUFFY PILLOW). Experiments 2 and 3 were double lexical decision tasks. In Experiment 2, distractor letter strings were orthographically illegal, non-pronounceable non-words (e.g., FLZFFY PXLLZW) created by modifying a key vowel in real words thus recruiting the shallowest level of processing of the online tasks. In Experiment 3, distractor letter strings were orthographically legal, pronounceable non-words (e.g., FLEFFY PALLOW) produced by a pseudoword generator called Wuggy (Keuleers & Brysbaert, 2010) thus requiring deeper levels of processing than the non-pronounceable non-word version of the task (James, 1975; Schulman & Davidson, 1977).

Experiment 4: Sense/nonsense Judgment Task

In the sense/nonsense judgment task, participants were presented with adjective-noun pairs and asked to determine whether the phrase made sense. This task represents the deepest level of semantic processing of the online processing task, as it requires participants to construct meaning of adjective-noun phrases (or fail to) via conceptual combination (e.g., Gagné & Spalding, 2013). Response times for both sense and nonsense judgments were analyzed separately.

Experiment 5: Explicit Judgment Task

In an explicit judgment task, participants were presented with novel (i.e., low meaningful) adjective-noun pairs and asked to provide an interpretation of the phrase. This is an offline deep level processing task, and participant responses were coded into themes of interpretation approaches.

CHAPTER 3

EXPERIMENT 1a and 1b: STIMULUS DEVELOPMENT

The purpose of Experiment 1 was to construct a stimulus set that varied in adjective-noun meaningfulness and concreteness at the phrase level. A novel quantitative measure was used to gather meaningfulness and concreteness ratings unique to the stimulus set.

Preliminary Development

A total of 86 head nouns were generated by the researcher, with half determined to denote concrete concepts and half abstract concepts. Notably, a more precise measure of concreteness was gathered through participant ratings in Experiment 1b. Each noun was paired with three adjectives and the resulting three adjectives paired with a single noun will be collectively referred to as "adjective-noun sets". Two of the adjectives paired with the noun were deemed to make sense by the researcher and ranked in terms of meaningfulness, whereas the third adjective was deemed to yield a nonsensical pair. In total, the original stimuli comprised 258 adjective-nouns pairs (or 86 adjective-noun sets). All adjective and noun constituents selected had an orthographic frequency (OF) of less than 30 per million words of text based on WINDSORS found on Wordmine2 (Durda & Buchanan, 2008), given that frequency effects are a robust finding in psycholinguistic literature with a tendency to conceal semantic effects (Durda & Buchanan, 2008).

This large initial set was further reduced by eliciting interrater reliability ratings from twelve lab members blind to the study's objective. Lab members were asked to rank order meaningfulness from 1 (lowest meaningful pair in set) to 3 (highest meaningful pair in set). Adjective-noun sets were removed if one adjective-noun pair in the set had less than 69% agreement (9/13) on rankings. This resulted in the removal of 18 adjective-noun sets (54 adjective-noun pairs). The resulting stimulus base (68 adjective-noun sets; half abstract) was then used to identify adjective-noun sets that could be stratified into high, intermediate, and low meaningfulness pairs through subjective rating measures (see below).

Experiment 1a: Meaningfulness Ratings

The intent of the meaningfulness rating study was to create adjective-noun sets that could be stratified into high, intermediate, and low meaningful pairs with the same head noun. These adjective-noun sets served as the stimuli in the online experimental tasks (Experiment 2-4). A second aim was to identify adjective-noun pairs that were rated as low in meaningfulness, as these served as the novel stimuli for the explicit judgment task (Experiment 5).

Method

Participants

Three hundred and seventy-three participants were recruited through Prolific (prolific.co). Inclusion criteria were that participants had to be native speakers of English and were only able to participate in one of our posted studies. Participants were paid £1.88 for their participation on abstract adjective-noun lists and £1.50 for their participation on concrete adjective-noun lists, as the latter lists took less time on a pilot study.

Materials

The 68 adjective-noun sets were divided into 34 adjective-noun sets with concrete head nouns and 34 adjective-noun sets with abstract head nouns. This was to ensure that concreteness did not create a confound in participants' meaningfulness ratings. Each adjective-noun set was further divided into its three adjective-noun pairs and randomly assigned to a list in order to create lists that were composed of unique nouns (i.e., no list contained the same noun) that varied in meaningfulness. The lists were double-checked to ensure that one-third of the stimuli from each list were made up of adjective-noun pairs that were ranked low in meaningfulness. In total, six lists of 34 items were created. The task was administered using Q-sort methodology on QMethod Software (Lutfallah & Buchanan, 2019), which was described in detail above (refer to Figures 2 and 3).

Procedure

If inclusion criteria were met, participants on Prolific were able to view the study once published. Upon seeing the study advertisement, participants were directed to click a link (QMethod Software) if interested in conducting the study. Then, participants were directed to the consent form. If they consented by clicking "I agree", they were asked to provide their Prolific ID and given instructions for the task. Participants were instructed to determine whether an adjective-noun combination makes sense as a pair, and they were provided with an example of a high (e.g., DEADLY POISON), intermediate (e.g., SWEET POISON), and low (e.g., MERRY POISON) meaningful combination. In the pre-sort stage, participants were presented with all adjective-noun combinations and asked to judge whether the pair was high meaningful (indicated with an up arrow), intermediate meaningful (indicated with a middle icon), or low meaningful (indicated with down arrow). Once completed, participants were automatically advanced to the final-sort stage. In this stage, participants pre-sort judgements divided the stimuli into the respective three sorting options to facilitate the final sort. Participants were asked to "drag and drop" the pairs along a distribution ranging from -2 (low meaningful) to +2 (high meaningful) until all items were sorted along the 5-column flat distribution. A validity check was included, in which participants were directed to drop a card in a specified column. Once completed, participants submitted their responses for review and received between £1.50 to £1.88 for their participation if they passed the validity check. The entire task was completed in approximately 12-15 minutes.

Results

All data were exported from each list and these data provided a range of 56 to 66 participant ratings per item. Ninety-six participants were removed for failing a validity check, leaving 277 participants for further analysis. Data cleaning was conducted in Excel. For each item, descriptive statistics were calculated including the mean, standard deviation, mode, median, minimum value, and maximum value.

The following criteria was used to evaluate an adjective-noun set for inclusion in the final stimulus set: the high pair needed to have a mean meaningfulness value above .75, the intermediate pair needed to have a mean value between -.75 and .75, and the low pair needed to have a mean value of less than -.75. This resulted in the removal of 29 adjective-noun sets. The final stimulus set contained 39 adjective-noun sets made up of 117 adjective-noun pairs, in which three adjective-noun pairs contained the same head noun but varied in meaningfulness (see Appendix A). Of the final 39 adjective-noun sets, 21 had concrete head nouns and 18 had abstract head nouns.

Data analyses were conducted in IBM SPSS Statistics Software (Version 22). Data were determined to meet the assumptions of a univariate ANOVA, including equal groups sizes and normal distributions. The assumption of homogeneity of variance was met based on Levene's Test (p=.23). Using an ANOVA with alpha set at .05, a difference was found in meaningfulness values between high, intermediate, and low meaningful groups, $F(2,114)=781.25$, $p<.001$, $\eta^2=.9$. A Bonferroni post-hoc analysis showed that all groups differed from one another (p's<.001). Refer to Table 2 for meaningfulness means and standard deviations of each meaningful group.

Other lexical variables were examined to ensure equality among groups. These variables included combined letter length of adjectives and nouns in each pair, as well as the mean orthographic frequency (OF) of the adjective and noun in each pair. Semantic neighbourhood density of the adjective was also examined[1]. A multivariate ANOVA (MANOVA) was conducted to compare letter length, mean OF, and the SND of adjectives between groups. All data were independent and generally normally distributed. Equality of variances and covariances met Box's Test (p=.98) and equality of variances met Levene's Test (p's>.05). The MANOVA revealed no differences between meaningful groups across all variables ($F(6, 224)=.41$, p=.78). See Table 2 for a summary of all results.

Table 2

[1] SND of the noun was equivalent across groups because identical nouns were used in each group.

Summary of Variables for Meaningful Groups

Meaningful Group	*N*	Meaningfulness	Combined Letter Length	Mean OF	Adjective SND	Noun SND
High	39	1.23 (.26)	11.79 (2.83)	8.41 (5.09)	.52 (.10)	.59 (.11)
Intermediate	39	.26 (.32)	11.74 (2.70)	7.94 (5.18)	.54 (.09)	.59 (.11)
Low	39	-1.30 (.27)	11.10 (2.89)	7.39 (5.66)	.54 (.09)	.59 (.11)

For novel items for Experiment 5, the 277 items were examined and all items with mean meaningfulness values above -0.75 were removed (211 items). An additional three items were removed because they contained a duplicate noun (the less meaningful item was retained). In sum, 63 adjective-noun pairs were selected to be used as novel stimuli for Experiment 5 (see Appendix B).

Discussion

The QMethod Software was used to generate average participant meaningfulness ratings of adjective-noun combinations. These ratings were used to reduce the stimulus set to 39 adjective-noun sets that could be stratified into a high, intermediate, and low meaningful adjective-noun pairs with identical head nouns, which were used in the concreteness rating task and online processing tasks (Experiments 2-4). These meaningfulness groups were matched on other relevant variables, including combined word length and mean orthographic frequency of the constituents, though these variables were still included in future analyses as covariates for robustness. Additionally, the meaningfulness rating task produced 63 novel (i.e., low meaningful) adjective-noun combinations for the explicit judgment task (Experiment 5).

Experiment 1b: Concreteness Ratings

The purpose of the concreteness rating study was to quantify the concreteness of adjective-noun phrases to serve as a categorical variable in the online processing tasks (Experiment 2-4).

Method

Participants

Two hundred and thirty-two participants were recruited through Prolific (prolific.co). All participants were native speakers of English and only able to participate in one of our posted studies. Participants were paid £1.88 for their participation on meaningful lists (List 2 and 3) and £2.50 for their participation on the novel list (List 1), as the latter was shown to require additional time on a pilot study.

Materials

The thirty-nine adjective-noun sets derived from the meaningfulness rating task were used as the stimuli (Appendix A). Three lists were created to separate the low, intermediate, and high meaningful adjective-noun pairs (i.e., List 1, 2, and 3, respectively), so that meaningfulness would not confound concreteness ratings. Each list contained 39 adjective-noun pairs from one meaningful group. Like the meaningfulness rating task, QMethod Software (Lutfallah & Buchanan, 2019) was used to record participant ratings using a 5-column flat distribution ranging from -2 (abstract) to +2 (concrete).

Procedure

If inclusion criteria were met, participants on Prolific were able to view the study once published. Upon seeing the study advertisement, participants were directed to click a link (QMethod Software) if interested in conducting the study. Then, participants were directed to the consent form to purview. If they consented by clicking “I agree”, they were asked to provide their Prolific ID and given instructions for the task. Prior to beginning the task, participants were asked to read the instructions carefully. Instructions were adapted from Brysbaert et al. (2014) who collected concreteness ratings for simple concepts. Participants were told the essence of a concrete word (i.e., experience-based) and the essence of an abstract word (i.e., language-based), and how some words fall in-between these two extremes. Examples of concrete adjective-noun pairs (i.e., SWEET CANDY, JUMPY FROG, and COMFY COUCH) and abstract adjective-noun pairs (i.e., FAIR JUSTICE, STICKY LIE, CURIOUS QUESTION) were provided. An additional instruction was added to explain that some word pairs may not make sense (e.g., STEEP RIDDLE, SPICY YARN), but participants were encouraged to imagine if the word pair could be directly experienced through senses (concrete) or if its meaning depended on language (abstract). Participants were instructed to determine whether an adjective-noun combination was concrete, abstract, or in-between. In the pre-sort stage, participants were presented with all adjective-noun combinations and asked to judge whether the pair was concrete (indicated with a filled-in circle), in-between (indicated with a half- filled circle), or abstract (indicated with an empty circle). Once completed, participants were automatically advanced to the final-sort stage. In this stage, participants pre-sort judgements divided the stimuli into the respective three sorting options to facilitate the final sort. Participants were asked to “drag and drop” the pairs along a flat distribution ranging from -2 (abstract) to +2 (concrete) until all items were sorted along the 5-column flat distribution. A validity check was included, in which participants were directed to drop a card in a specified column. Once completed, participants submitted their responses for review and received between £1.88 to £2.50 for their participation as long as they passed the validity check. The entire task was completed in approximately 15-20 minutes.

Results

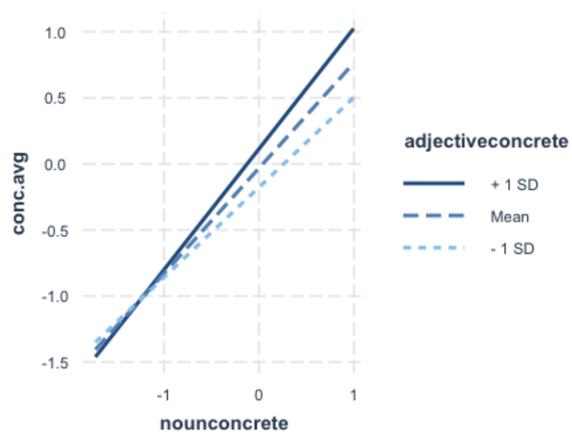
Data were exported from QMethod Software and compiled, with a total of 77-78 ratings per item. Descriptive statistics were computed in Excel. Thirteen participants were removed for failing a single validity check, leaving 219 participant ratings for further analysis. For each item, descriptive statistics were calculated including the mean, standard deviation, range, minimum value, and maximum value. Means and standard deviations of concreteness ratings for each adjective-noun pair can be found in Appendix C. Descriptive information regarding other lexical and semantic variables for each adjective-noun pair can also be found in Appendix C.

Data analyses were conducted in IBM SPSS Statistics Software (Version 22). Assumptions were determined to be met for a Pearson correlation. A Pearson correlation matrix was computed with all 117 adjective-noun pairs to determine the relations between variables, including concreteness, SND of the adjective, letter length combined, and mean orthographic frequency (OF). Concreteness was weakly negatively correlated with combined letter length ($r(116)=.30$, $p=.001$) and mean OF ($r(116)=-.23$, $p=.012$) , meaning that more abstract pairs were associated with longer combined letter length and higher mean OF. SND of the adjective was also weakly negatively correlated with mean OF ($r(116)=-.26$, $p=.005$), such that denser adjective SND was associated with a lower mean OF. No other correlations were significant. For each separate list of the 39 adjective-noun pairs, a Pearson correlation was examined between concreteness and SND of the noun. For high and low meaningful adjective-noun pairs, a weak positive correlation was observed, with higher concreteness associated with a denser noun SND, $r(38)=.38$, $p=.019$ and $r(38)=.37$, $p=.020$, respectively. Intermediate meaningful adjective-noun pairs also yielded a moderate positive correlation, with higher concreteness associated with a denser noun SND, $r(38)=.53$, $p=.001$. A correlation table of meaningfulness as a continuous variable and its associations with the other independent variables and covariates is included in Appendix D. Recall that meaningfulness was converted into a categorical variable for the value of the middle intermediate group and comparison across a gradient of meaningfulness for adjective-noun sets.

An additional analysis was conducted in R (R Core Team, 2017) version 3.4.3 with the tidyverse (Wickham et al., 2019) and broom (Robinson et al., 2022) packages to examine the relationship between the adjective-noun phrase concreteness ratings obtained in the present study compared to adjective and noun constituent concreteness ratings in Brysbaert et al. (2014). The assumption of normality was violated so a Kendall's tau-b correlation was used. A moderate positive correlation was found between phrase concreteness and noun concreteness ratings ($\tau_b=.64$, $p<.001$), and a weak correlation was found between phrase concreteness and adjective concreteness ratings ($\tau_b=.30$, $p<.001$). Notably, there was no relation found between phrase concreteness and meaningfulness ratings ($\tau_b=.00$, $p=1.00$). A multiple regression was carried out to investigate whether an interaction between constituent concreteness values could significantly predict phrase concreteness ratings. Assumptions of normality, independence, multicollinearity, homogeneity of variance, and linearity were met. Predictor variables were centered and entered into the model. The model was significant, $F(3, 113)=230.80$, $p<.001$, $R^2=.86$. There were significant effects for noun concreteness ratings ($b=.78$, $p<.001$), adjective concreteness ratings ($b=.15$, $p<.001$), and an interaction of adjective and noun concreteness ratings ($b=.12$, $p=.001$). For the interaction (refer to Figure 4), adjective concreteness did not affect overall phrase concreteness ratings when the noun was abstract but did modify phrase ratings when the noun was concrete. Higher phrase concreteness ratings were observed when a concrete noun was paired with a concrete adjective and lower ratings were observed when a concrete noun was paired with an abstract adjective.

Figure 4

Interaction Between Constituent Noun and Adjective Concreteness on Phrase Concreteness Ratings



Note: Higher positive values are associated with higher concreteness. The y-axis (conc.avg) is adjective-noun phrase concreteness.

When examining the distribution of adjective-noun phrase concreteness ratings, there was a clear bimodal distribution reflecting abstract entities and concrete entities. As such, rather than keeping concreteness as a continuous measure, a categorical measure was derived by a mean split ($M=0.00$). Abstract adjective-noun pairs were labelled as those equal to and below the mean ($N=54$), whereas concrete adjective-noun pairs were categorized as those above the mean ($N=63$).

Discussion

In summary, QMethod Software was used to gather mean concreteness ratings of 117 adjective-noun pairs to be used as an independent categorical variable in the online processing tasks (Experiment 2-4). Generally, weak associations were observed between concreteness and other lexical variables, which again suggests that the stimulus set is well controlled for extraneous variables such as combined letter length and mean OF, though these were included as covariates in future analyses. Additionally, independent variables that were used in subsequent analyses (i.e., SND of constituents and concreteness) do not appear to have an issue of multicollinearity, although this was measured more rigorously in future data analyses as well. Lastly, concreteness of the adjective affected overall phrase concreteness ratings when paired with a concrete noun but not with an abstract noun, which was contrary to what was predicted. That is, concrete nouns paired with concrete adjectives were rated as more concrete whereas concrete nouns paired with abstract adjectives were rated as less concrete, but phrases with abstract nouns were rated as less concrete regardless of adjective concreteness.

CHAPTER 4

EXPERIMENTS 2-4: ONLINE PROCESSING TASKS

The purpose of the experimental, online processing tasks was to measure the independent and interactive effects of semantic variables related to the constituent (i.e., SND) and phrase (i.e., meaningfulness and concreteness) composition of adjective-noun combinations to gain a better understanding of a conceptual combinatory process within these combinations. Three online processing tasks were selected that differentially engage shallow to deep levels of semantic processing to compare these effects across different task demands, as different patterns of response times have been found in noun-noun combinations (e.g., Gagné & Shoben, 1997). These three online processing tasks include a double lexical decision task with non-pronounceable non-words, a double lexical decision task with pronounceable non-words, and a sense/nonsense judgment task.

Method

Participants

Four hundred and fifty-seven participants (*N*=145 to 157 per experiment) were recruited through Prolific (prolific.co). Participants varied in sex (276 female, 181 male), age (range=18-75; *M*=32 years, *SD*=12.68), and education level (range=less than high school to doctorate; *M*=14 years, *SD*=2.36). All participants were native speakers of English and able to participate in only one of our posted studies. Participants were paid £1.25 for their participation in Experiments 2 and 4 and £1.88 for their participation in Experiment 3, as the latter was deemed to require additional time on a pilot study.

Material

The 117 adjective-noun pairs (Appendix A) were quasi-randomly distributed to create three lists with 39 adjective-noun pairs varying in meaningfulness and concreteness, with no list containing the same head noun. Therefore, each experiment had three versions comprising the different lists. For the double lexical decision tasks, 39 non-words were also generated, either by manipulating the vowels to make non-pronounceable non-words (e.g., CJRNY FRKUD; Experiment 2) or using the pseudoword generator Wuggy (e.g., COFFY FRAUN; Keuleers & Brysbaert, 2010; Experiment 3). Thus, for the double lexical decision tasks, participants viewed 78 stimuli in total, half real word pairs and half non-word pairs. For the sense/nonsense judgment tasks, an additional 13 novel adjective-nouns pairs (i.e., low meaningful) were randomly selected to be included in each list (see Appendix B). This was in an attempt to equalize the potential ratio of sense and nonsense judgements made so as not to bias a key response. In total, participants viewed 52 stimuli on this task.

The experiments were created in an online open-source software for designing online behavioral experiments to collect chronometric data called Psychopy3 (psychopy.org; Pierce et al., 2019). Psychopy3 has been shown to present stimuli and record reaction times with a precise level of temporal resolution (Gallant & Libben, 2019). The experiments were conducted virtually using Pavlovia (pavlovia.org) as it is the interface that allows experiments created in Psychopy3 to run online. Demographic information was also collected through a Qualtrics (qualtrics.com) survey. The studies were advertised and accessed through Prolific.

Procedure

Eligible participants were able to view the study on Prolific (prolific.co) once published. Upon seeing the study advertisement, participants were directed to click a link (Qualtrics survey) if interested in conducting the study. Then, participants were directed to the consent form and if they consented by clicking "I agree", they were asked to provide demographic information, including their age, sex, and education, as well as their Prolific ID. Participants were then re-directed to conduct the study through Pavlovia.org, which provides an online platform to conduct experiments created in Psychopy3. Participants were again asked to provide their Prolific ID, and then provided the instructions for the task with *please read carefully* noted at the top of the instruction window. For the DLDT tasks, participants were instructed that they would be presented with two letter strings and asked to determine if they were real words (by pressing the A key) or not (by pressing the L key), with an example of each provided. Participants were encouraged to respond as quickly as possible and provided six practice trials which required the correct response to successively advance through the trials. For the sense/nonsense judgment task, participants were instructed that they would be presented with two words and asked to determine whether the pair makes sense (by pressing the A key) or not (by pressing the L key). An example of a sensical (i.e., DEADLY POISON) and nonsensical (i.e., MERRY POISON) pair were provided for reference. Participants were encouraged to make their decisions as quickly as possible. Six practice trials were provided to acclimate to the task, though corrective feedback was not provided given sensical or nonsensical judgments are unique to the individual. In each task, two validity checks that instructed participants to “press the A key” or “press the L key” were included. Once completed, participants were thanked for their participation and their responses were submitted for review. They received between £1.25 to £1.88 for their participation as long as they passed both validity checks. Each task was completed in approximately 10-15 minutes.

Data Analysis Assumptions and Overview

All data from the three lists were compiled for each experiment, and separate analyses were conducted for each experiment. For outlier analyses on the double lexical decision tasks (DLDTs), a minimum accuracy rate of 70% was used as a cut-off for participants and for words. Additionally, only correct responses were analyzed, and those that were faster than 300ms or slower than 3000ms were removed. For the sense/nonsense judgment task, a less stringent cutoff was applied given the nature of the task (i.e., there were no wrong answers) and because individuals appeared to trend towards longer response latencies. For this experiment, any responses that were faster than 300ms or slower than 5000ms were removed. Response times for sense and nonsense judgments were analyzed separately. All data were analyzed using R (R Core Team, 2017) version 3.4.3. The lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017) packages were used. Probability values (*p* values) were obtained using the lmerTest package with Satterthwaite approximations to degrees of freedom (Kuznetsova et al., 2017). The ggplot2 package (Wickham, 2016) was used to generate plots of findings.

To meet the assumptions of analyses, reaction times (RTs) were log transformed to ensure normal distribution of the dependent variable (Keene, 1995). The sample sizes were large and well over ten observations per independent variable, varying between 4,633 and 5,096 observations. Data were visualized to ensure a linear relationship between the predictor variables and the dependent variable. There was an equal number of observations for each variable and no missing data. The variance inflation factor (VIF), produced with the car package (Fox & Weisberg, 2019), was used to examine multicollinearity among predictors and deemed acceptable if less than five; VIFs of predictor variables and covariates were less than or equal to three across analyses. An assumption of homogeneity of variances, homoscedasticity, or sphericity are not required for linear mixed effects models (Quené & Van der Bergh, 2008). Continuous predictors and covariates were centered and standardized to provide uniformity as well as assist models with converging. Lastly, normality of the residuals is an assumption that was assessed through examination of q-q plots and addressed through trimming the residuals (see below).

For reaction time (RT) analyses, data were analyzed in two linear mixed effects models. To compare effects across meaningful groups, fixed effects included meaningful group, concreteness, adjective SND, meaningful X concreteness, and concreteness X adjective SND. To evaluate response times within meaningful groups, fixed effects included concreteness, adjective SND, noun SND, and their three-way and two-way interactions. Covariates for both analyses included combined letter length and mean orthographic frequency of the adjective-noun pair. Subjects and items were included in all models as random effects, and the models were fitted with random slopes by subject and by item: subjects with random slopes for meaningful group and concreteness, and items with random slopes for concreteness were considered across meaningful groups, whereas subjects and items with random slopes for concreteness were considered within each meaningful group. Models were built using a data-driven, bottom-up approach advocated by Nagle (2019), by backward-testing fixed effects followed by forward-testing random effects. The optimizer bobyqa was used to address convergence problems, and if unsuccessful the simpler model was retained. Models were specified to use maximum likelihood estimation to allow for comparison of models with different fixed effects while backward-fitting, and restricted maximum likelihood estimation to allow for model comparison with different random effect structures when forward-fitting (Nagle, 2019). Models were compared using an ANOVA, and a significant chi-square value, combined with lower AIC and BIC values, was supportive of model improvement.

Models were evaluated by examination of plots of fitted against residual values, and model trimming was used to remove outliers with standardized residuals that were 2.5 standard deviations away from 0. Model trimming was only conducted if <5% of the data were removed, and this was also to primarily to address the assumption of normality of the residuals. Skewness and kurtosis values are reported for the residuals using the moments package (Lukasz & Novomsetsky, 2015). R^2 values, including the marginal R^2 (i.e., the variance explained by fixed effects) and the conditional R^2 (i.e., the variance explained by both fixed and random effects), are also reported for each model with the MuMIn package (Bartoń, 2020), and are based on how Nakagawa et al. (2017) suggest that R^2 values should be calculated for linear mixed models. Additionally, Tukey adjusted pairwise comparisons were examined using the emmeans packages (Lenth, 2021) and effect sizes were estimated for each independent variable using *t* values and degrees of freedom of the error estimate to calculate partial eta squared through the effectsize package (Ben-Shachar et al., 2020). Models were reported in tables based on a best practice approach advocated by Meteyard and Davies (2020).

For error analyses, the binomial dependent variable (i.e., correct or incorrect) was analyzed using a mixed logit model (generalized linear mixed model; Jaeger, 2008). The same fixed and random effects as the RT analyses were used to compare effects across meaningful groups. The same procedure for model fitting and evaluation was also used, though no model trimming was conducted. Error analyses within meaningful groups were not conducted as errors made up <6% of the total data for both DLDTs. An error analyses was not conducted for the sense/nonsense judgment task.

Results

Experiment 2: Double Lexical Decision Task with Non-pronounceable Non-words

Data Cleaning

Non-pronounceable non-word trials were removed, as only real adjective-noun word trials were analyzed further. Eighteen participants were removed from the analyses; fourteen were removed for failing a validity check, three were removed for having less than 70% accuracy, and one was removed for having an unusually slow response trend (i.e., all reaction times >2000ms). Forty-five responses (.84% of data) were removed for being less than 300ms or greater than 3000ms. Item accuracy was above 70% for all word pairs. In total, 8.7% of the data were removed. Final reaction time analysis included only correct responses and 137 participants.

Reaction Time (RT) Analyses

Mean RTs, standard deviations, and error rates per meaningful group and concrete group for the final data set are displayed in Table 3. To compare effects across meaningful groups, backward and forward fitting was conducted to remove variables from the model (as described in detail above), and

Table 3

Mean RTs (with SDs) and Error Rates Per Group in Experiment 2

Group		*N*	Mean Log RT (ms)	Mean Raw RT (ms)	Mean error rate (%)
Low	Abstract	733	6.63 (.33)	805.42 (329.84)	4.16
Meaningful	Concrete	952	6.64 (.33)	809.25 (324.52)	5.07
Intermediate	Abstract	783	6.62 (.31)	791.50 (307.40)	4.37
Meaningful	Concrete	917	6.61 (.33)	793.88 (336.09)	4.06
	Abstract	833	6.63 (.31)	799.02 (297.13)	3.24

High Meaningful	Concrete	877	6.58 (.32)	765.99 (298.60)	3.40

the final model included fixed effects for combined letter length and mean orthographic frequency (OF; i.e., the covariates), and random intercepts for subjects and items. After the model was fitted, data were trimmed by removing outliers with a standardized residual at a distance greater than 2.5 standard deviations from 0. This resulted in the removal of 138 observations (2.71% of data). Skewness of the residuals was .42 and kurtosis was 3.11, and this is considered to be normally distributed according to Hair et al. (2010) and Bryne (2010). There was a main effect of combined letter length [b = -.027, $t(109.76) = 4.13$, $p < .001$, $\eta_p^2 = .13$], with shorter combined letter length yielding shorter RTs than adjective-noun pairs with longer combined letter length. There was a main effect of mean OF [b = -.018, $t(110.72) = -2.77$, $p = .007$, $\eta_p^2 = .06$], with adjective-noun pairs with higher mean frequency yielding shorter RTs than adjective-noun pairs with lower mean frequency. The marginal R^2 and the conditional R^2 for the overall model were .01 and .50, respectively.

Within each meaningful group, the same procedure was followed as outline above. For adjective-noun pairs of low meaningfulness, all fixed effects were removed during model fitting, and a model with covariates was no better than the null model (i.e., one without any fixed effects or covariates). As such, no model is reported.

For intermediate meaningful adjective-noun pairs, the final model included fixed effects for mean OF and random intercepts for subjects and items. Data were trimmed resulting in the removal of 45 observations (2.65% of data). For the residuals, skewness was .44 and kurtosis was 3.26. There was a main effect of mean OF [b = -.02, $t(36.48) = -2.33$, $p = .03$, $\eta_p^2 = .13$], with adjective-noun pairs of higher mean frequency yielding shorter RTs than those with lower mean OF. Marginal R^2 was .01 and conditional R^2 was .48 for the final model.

For high meaningful adjective-noun pairs, the final model included fixed effects for mean OF and combined letter length (i.e., the covariates), and random intercepts for subjects and items. Data were trimmed resulting in the removal of 40 observations (2.34% of data). Skewness was .46 and kurtosis was 3.08 for the residuals. There was a main effect of combined letter length [b = .040, t(36.53) = 3.47, p =.001, η_p^2=.25], with shorter combined letter length yielding shorter RTs than adjective-noun pairs with longer combined letter length. The effect for mean OF was non-significant (p=.06) after trimming the residuals. Marginal R^2 was .03 and conditional R^2 was .53 for the final model.

Error Analyses

Errors made up 4.01% of the data. To compare effects across meaningful groups, backward and forward fitting was conducted to remove variables from the model. All fixed effects were removed during model fitting, and a model with covariates was no better than the null model (i.e., one without any fixed effects or covariates). As such, no model is reported.

Experiment 3: Double Lexical Decision Task with Pronounceable Non-words

Data Cleaning

Pronounceable non-word trials were removed, as only real adjective-noun word trials were analyzed further. Sixteen participants were removed from the analyses; three were removed for failing a validity check and thirteen were removed for having less than 70% accuracy. Two items were removed from all participant responses for having accuracy below 70%. Forty-eight responses were removed for being less than 300ms or greater than 3000ms. In total, 8.62% of the data were removed. Final reaction time analysis included only correct responses and 129 participants.

Reaction Time (RT) Analyses

Mean RTs, SDs, and error rates per meaningful group and concrete group for the final data set are displayed in Table 4. Across meaningful groups, forward and backward fitting was used to remove

Table 4

Mean RTs (with SDs) and Error Rates Per Group in Experiment 3

Group		*N*	Mean Log RT (ms)	Mean Raw RT (ms)	Mean error rate (%)
Low	Abstract	649	6.89 (.33)	1039.09 (383.32)	9.23
Meaningful	Concrete	890	6.83 (.33)	975.66 (374.89)	5.62
Intermediate	Abstract	775	6.83 (.32)	977.06 (345.99)	4.67
Meaningful	Concrete	763	6.82 (.33)	965.66 (366.28)	4.03
High	Abstract	782	6.81 (.28)	948.05 (298.27)	4.05
Meaningful	Concrete	774	6.75 (.31)	900.04 (326.31)	3.85

variables, and the final model included fixed effects for meaningfulness, concreteness, and the covariates, and random intercepts for items and participants, and participants with random slopes by concreteness. After the model was fitted, data were trimmed resulting in the removal of 116 observations (2.50% of data). Skewness was .52 and kurtosis was 3.29. The final model is shown in Table 5. For the main effect of meaningfulness, reaction times to high meaningful adjective-noun pairs were

Table 5

Summary of the Final Model Across Meaningful Groups in Experiment 3

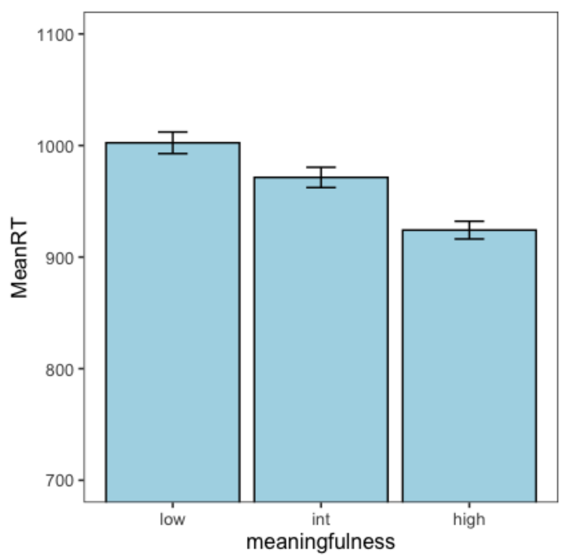
Fixed Effects					
	Estimate/ Beta	SE	95% CI	t	p
Intercept	6.87	.02	6.82 to 6.91	313.12	<.001
mean OF	-.03	.01	-.04 to -.01	-3.77	<.001
meaningfulness – int	-.03	.02	-.06 to .01	-1.63	.11
meaningfulness – high	-.06	.02	-.10 to -.03	-3.73	<.001
concreteness- concrete	-.06	.01	-.09 to -.03	-3.99	<.001
Random Effects					
			Variance	S.D.	Correlation
Participant (Intercept)			.0361	.19	
Items (Intercept)			.0043	.07	
Concreteness | Participant (slope)			.0013	.04	0.11
Model fit					
R^2			Marginal		Conditional
			.02		.51

Key: p-values for fixed effects calculated using Satterthwaite's approximations. Confidence Intervals have been calculated using the Wald method.
Model equation: logRT ~ mean OF + meaningfulness + concreteness + (1 + concreteness| Participant) + (1 | Item)

significantly faster than reaction times to low meaningful adjective-noun pairs [$t(109)=-3.73$, $p<.001$, $\eta_p^2=.11$]. Response times to intermediate meaningful adjective-noun pairs did not significantly differ from high or low adjective-noun pairs (*p*'s= .10 and .24, respectively; see Figure 5). For the main effect

Figure 5

Main Effect of Meaningfulness in Experiment 3

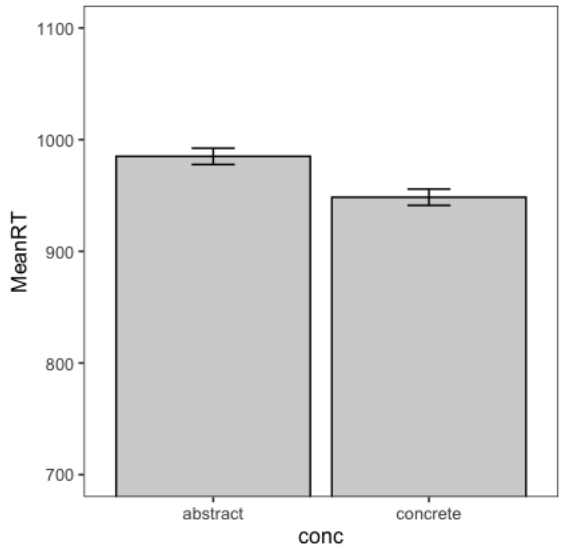


Note. A significant difference was observed between high and low meaningful groups. Reaction times are displayed in milliseconds.

of concreteness, response times were faster for concrete items compared to abstract ones [$t(115)=-3.25$, $p=.002$, $\eta_p^2=.12$; see Figure 6]. There was also a main effect of mean OF [$b = -.03$, $t(108.89) = -3.77$,

Figure 6

Main Effect of Concreteness Across Meaningful Groups in Experiment 3



Note. Reaction times are displayed in milliseconds.

p =<.001, η_p^2=.12], with a higher mean OF yielding shorter reaction times compared to lower mean OF pairs.

Within each meaningful group, the same model fitting procedure was used. For adjective-noun pairs of low meaningfulness, the final model included a fixed effect for concreteness and random intercepts for subjects and items. Data were trimmed resulting in the removal of 42 observations (2.73% of data). For the residuals, skewness was .53 and kurtosis was 3.18. There was a main effect of concreteness, [*b*= -.07, *t*(33.95) = -2.72, *p* =.01, η_p^2=.18], with concrete adjective-noun pairs yielding faster reaction times than abstract adjective-noun pairs. The marginal R^2 was .01 and the conditional R^2 was .48 for the final model.

For adjective-noun pairs of intermediate meaningfulness, the final model included a fixed effect for mean OF and random intercepts for subjects and items. Data were trimmed resulting in the removal of 40 observations (2.60% of data). Skewness was .48 and kurtosis was 3.30. There was a main effect of mean OF [*b*= -.04, t(35.94) = -2.71, *p* =.01, η_p^2=.17], with higher frequency pairs yielding faster reaction times compared to lower frequency adjective-noun pairs. The marginal R^2 was .01 and the conditional R^2 was .52 for the final model.

For high meaningful adjective-noun pairs, the final model included a fixed effect for combined letter length and random intercepts for subjects and items. Data were trimmed resulting in the removal of 39 observations (2.51% of data). Skewness was .45 and kurtosis was 3.22. There was a main effect of combined letter length [*b*= .04, t(34.59) = 3.49, *p* = .001, η_p^2=.26] with longer combined letter length yielding slower reaction times compared to shorter combined letter length. For the final model, marginal R^2 was .03 and the conditional R^2 was .53.

Error Analyses

Errors made up 5.18% of the data. To compare effects across meaningful groups, backward and forward fitting was conducted to remove variables from the model. The final model included fixed effects for meaningfulness and random intercepts for participants and items. There was a main effect of meaningfulness [*b*= .63] though follow-up comparisons with Tukey adjustment revealed no significant difference in errors between low meaningful pairs and high meaningful pairs (*p*=.05) or intermediate meaningful pairs (*p*=.08), and high meaningful pairs did not significantly differ from intermediate meaningful pairs (*p*=.98).

Experiment 4: Sense/nonsense Judgement Task

Data Cleaning

Additional low meaningful stimuli (13 items) that were not a part of the meaningful groups were removed. Four participants were removed; three were removed for failing validity checks and one for having nearly all responses <300ms. Sixty-one responses were removed for being less than 300ms or greater than 5000ms. In total, 9.58% of the data were removed. Final reaction time analysis included 153 participants with 3,583 sense judgments and 2,323 nonsense judgments. Refer to Table 6 for a summary of mean RTs and SDs of sense and nonsense judgments per meaningfulness and concreteness group. Given differences in group sizes, low meaningful stimuli were dropped from sense judgment analyses, whereas high meaningful stimuli were dropped from nonsense judgment analyses.

Table 6

Mean RTs (with SDs) For Sense and Nonsense Judgments in Experiment 4

		Sense Judgments			Nonsense Judgments		
Group		*N*	Mean Log RT	Mean Raw RT (ms)	*N*	Mean Log RT	Mean Raw RT (ms)
Low Meaningful	Abstract	151	7.38 (.46)	1780.67 (834.71)	705	7.27 (.39)	1567.73 (718.39)
	Concrete	141	7.29 (.41)	1585.93 (665.13)	969	7.18 (.35)	1398.77 (587.05)
Intermediate Meaningful	Abstract	643	7.30 (.39)	1608.32 (699.41)	264	7.27 (.45)	1609.99 (838.66)
	Concrete	803	7.23 (.38)	1493.56 (650.50)	258	7.26 (.42)	1558.39 (736.39)
High Meaningful	Abstract	889	7.13 (.37)	1344.43 (587.72)	66	7.28 (.48)	1657.09 (829.13)
	Concrete	956	7.04 (.34)	1218.20 (521.89)	61	7.11 (.40)	1327.13 (572.19)

Reaction Time (RT) Analyses – Sense Judgments

To compare sense judgments across meaningfulness groups, forward and backward fitting was used to remove variables, and the final model included fixed effects for meaningfulness, concreteness, and mean OF, and random intercept for items, participants, and participants with random slopes by meaningfulness. After the model was fitted, data were trimmed resulting in the removal of 78 observations (2.37% of data). Skewness was .51 and kurtosis was 3.38 for the residuals. The final model is shown in Table 6. For the main effect of meaningfulness, reaction times were significantly faster for

Table 6

Summary of the Final Model for Sense Judgments in Experiment 4

Fixed Effects					
	Estimate/ Beta	SE	95% CI	t	p
Intercept	7.31	.03	7.25 to 7.36	269.02	<.001
mean OF	-.04	.01	-.06 to -.02	-3.88	<.001
meaningfulness – high	-.19	.02	-.23 to -.15	-8.76	<.001
concreteness-concrete	-.11	.02	-.15 to -.06	-5.02	<.001
Random Effects					
			Variance	S.D.	Correlation
Participant (Intercept)			.0603	.25	

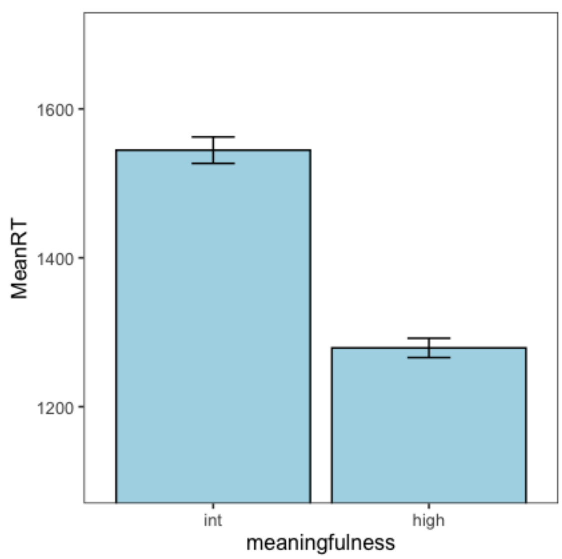
Items (Intercept)	.0067	.08	
Meaningfulness | Participant (slope)	.0078	.09	-0.54
Model fit			
R^2	Marginal		Conditional
	.10		.55

Key: p-values for fixed effects calculated using Satterthwaite's approximations. Confidence Intervals have been calculated using the Wald method.
Model equation: logRT ~ mean OF + meaningfulness + concreteness + (1 + meaningfulness| Participant) + (1 | Item)

high meaningful adjective-noun pairs compared to intermediate meaningful pairs [*t*(91.70)=-8.76, *p*=<.001, η_p^2=.46; see Figure 7]. For the main effect of concreteness, response times were faster for

Figure 7

Main Effect of Meaningfulness in Experiment 4

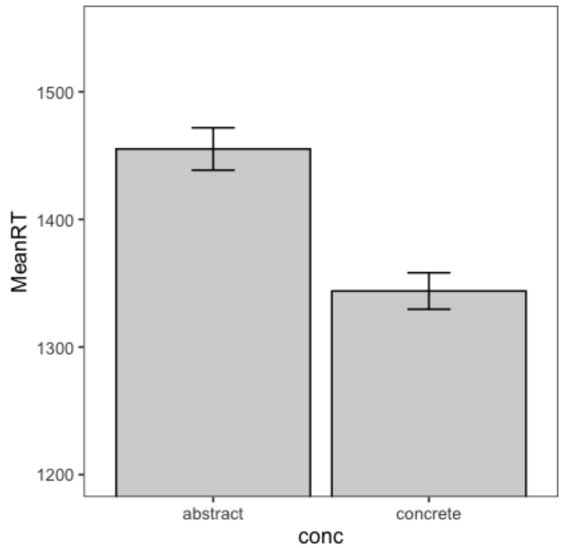


Note. Reaction times are for sense judgments and are displayed in milliseconds.

concrete items compared to abstract ones [*t*(75.70)=-5.01, *p*=<.001, η_p^2=.25; see Figure 8]. There was also a main effect of mean OF [*b* = -.04, t(75.08) = -3.85, *p* =<.001, η_p^2=.16], with a higher mean OF yielding shorter reaction times compared to adjective-noun pairs with lower mean OF.

Figure 8

Main Effect of Concreteness Across Meaningful Groups in Experiment 4



Note. Reaction times are for sense judgments and are displayed in milliseconds.

Within each meaningful group, the same model fitting procedure was used. For sense judgments for adjective-noun pairs of intermediate meaningfulness, the final model included fixed effects for concreteness, noun SND, and an interaction between adjective SND and concreteness. Random effects included random intercept for items, participants, and participants with random slopes by concreteness. After the model was fitted, data were trimmed resulting in the removal of 34 observations (2.35% of data). Skewness was .39 and kurtosis was 2.99 for the residuals. The final model is shown in Table 7.

Table 7

Summary of the Final Model for Sense Judgments of Intermediate Meaningful Items

Fixed Effects					
	Estimate/ Beta	SE	95% CI	t	p
Intercept	7.32	.03	7.26 to 7.38	235.79	<.001
adj SND	.03	.02	-.02 to .07	1.05	.29
concreteness-conc	-.11	.03	-.18 to -.05	-3.48	.001
noun SND	.04	.02	.00 to .08	2.47	.02
adjSND*Concrete	-.05	.03	-.11 to .02	-1.52	.14

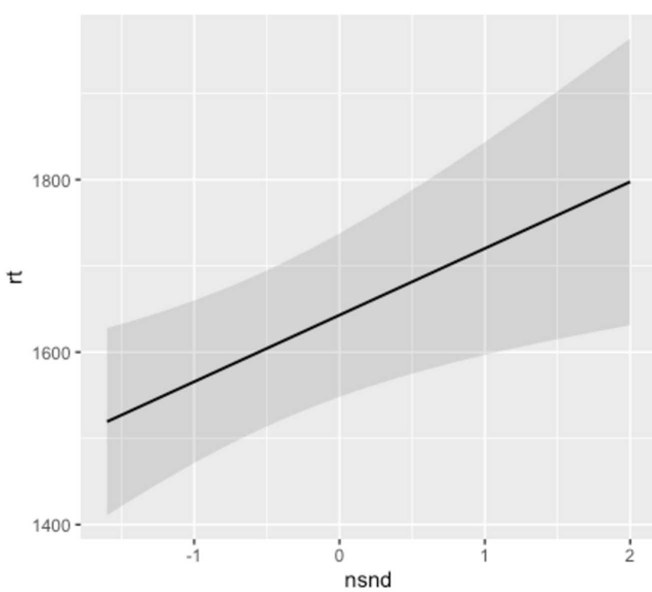
Random Effects			
	Variance	S.D.	
Participant (Intercept)	.0710	.27	
Items (Intercept)	.0052	.07	
Concreteness | Participant (slope)	.0051	.07	-0.57
Model fit			
R^2	Marginal		Conditional
	.02		.51

Key: p-values for fixed effects calculated using Satterthwaite's approximations. Confidence Intervals have been calculated using the Wald method.
Model equation: logRT ~ nounsnd + adjsnd*concreteness + (1 + conc | Participant) + (1 | Item)

A main effect of noun SND was observed, with faster sense judgments observed for adjective-noun pairs with a sparser noun SND compared to adjective-noun pairs with a denser noun SND [*b*=.04, *t*(36.31)=2.47, *p*=.02, η_p^2=.14; see Figure 9].

Figure 9

Main Effect of Noun SND for Sense Judgments for Intermediate Meaningful Adjective-Noun Pairs



Note: Mean reaction times are displayed in milliseconds.

There was also a typical main effect of concreteness, with concrete adjective-noun pairs yielding faster RTs compared to abstract adjective-noun pairs [$t(38.07)=3.92$, $p<.001$, $\eta_p^2=.25$]. The nonsignificant interaction between concreteness and adjective SND improved the model fit overall.

For sense judgements of high meaningful adjective-noun pairs, the final model included fixed effects for concreteness, adjective SND, and the covariates. The random effects included random intercepts for stimuli and participants, as well as participants with random slopes by concreteness. After the model was fitted, data were trimmed resulting in the removal of 57 observations (3.09% of data). Skewness was .55 and kurtosis was 3.51 for the residuals. The final model is shown in Table 8.

Table 8

Summary of the Final Model for Sense Judgments of High Meaningful Items

Fixed Effects					
	Estimate/ Beta	SE	95% CI	t	p
Intercept	7.12	.03	7.06 to 7.16	265.33	<.001
combined letter	.05	.01	.02 to .07	3.66	<.001
mean OF	-.06	.01	-.09 to -.04	-4.51	<.001
concreteness-concrete	-.09	.03	-.15 to -.04	-3.43	.002
adj SND	-.03	.01	-.05 to -.01	-2.31	.03

Random Effects			
	Variance	S.D.	Correlation
Participant (Intercept)	.0543	.23	
Items (Intercept)	.0046	.07	
Concreteness | Participant (slope)	.0025	.05	-0.82

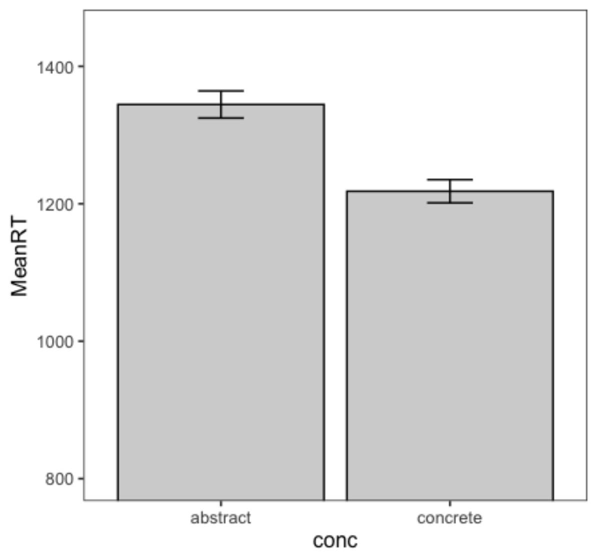
Model fit		
R^2	Marginal	Conditional
	.08	.56

Key: p-values for fixed effects calculated using Satterthwaite's approximations. Confidence Intervals have been calculated using the Wald method.
Model equation: logRT ~ combined letter + mean OF + concreteness + adjsnd+ (1 + concreteness| Participant) + (1 | Item)

There was a typical concreteness effect, in which sense judgments had faster RTs for concrete adjective-noun pairs compared to abstract adjective-noun pairs [t(34.1)=3.43, p=.002 η_p^2=.26; see Figure 11].

Figure 11

Main Effect of Concreteness for Sense Judgments for High Meaningful Adjective-Noun Pairs

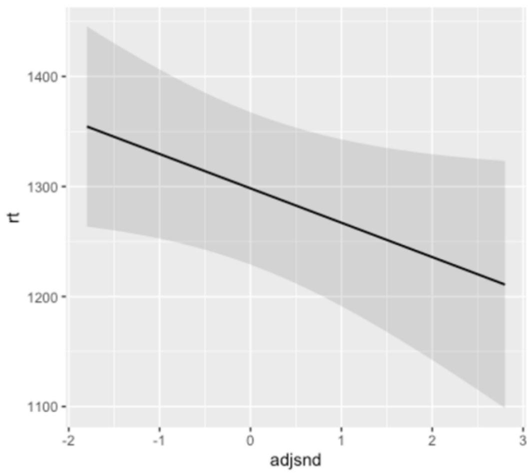


Note. Reaction times are displayed in milliseconds.

There was also a main effect for adjective SND, with faster sense judgments made for adjective-noun pairs with a denser adjective SND compared to a sparser adjective SND [b=-.03, t(32.67)=-2.31, p=.03

Figure 12

Main Effect of Adjective SND for Sense Judgments for High Meaningful Adjective-Noun Pairs



Note. Mean reaction times are displayed in milliseconds.

η_p^2=.14; see Figure 12]. Both covariates were also significant in the model. Higher frequency of adjective-noun pairs yielded faster RTs compared to lower mean OF of pairs. Shorter combined letter length also yielded faster RTs compared to adjective-noun pairs with longer combined letter length. Low meaningful stimuli were not analyzed due to a low sample size (i.e., sense judgments were not commonly made for low meaningful stimuli).

Reaction Time (RT) Analyses – Nonsense Judgments

To compare nonsense judgments across meaningfulness groups, forward and backward fitting of models was conducted resulting in variables being removed. The final model included fixed effects for meaningfulness and concreteness, and random intercepts for items and participants, and participants with random slopes by meaningfulness. After the model was fitted, data were trimmed resulting in the removal of 64 observations (2.91% of data). Skewness of residuals was .48 and kurtosis was 3.27. The final model is shown in Table 9.

Table 9

Summary of the Final Model for Nonsense Judgments in Experiment 4

Fixed Effects

	Estimate/ Beta	SE	95% CI	t	p
Intercept	7.24	.03	7.19 to 7.29	282.98	<.001
meaningfulness – intermediate	.09	.02	.05 to .13	4.14	<.001
concreteness-concrete	-.06	.02	-.10 to -.02	-3.27	.002
			Random Effects		
			Variance	S.D.	Correlation
Participant (Intercept)			.06	.24	
Items (Intercept)			.004	.07	
Meaningfulness | Participant (Slope)			.006	.08	.70
			Model fit		
R^2			Marginal		Conditional
			.02		.56

Key: p-values for fixed effects calculated using Satterthwaite's approximations. Confidence Intervals have been calculated using the Wald method.
Model equation: logRT ~ meaningfulness + concreteness + (1 + meaningfulness| Participant) + (1 | Item)

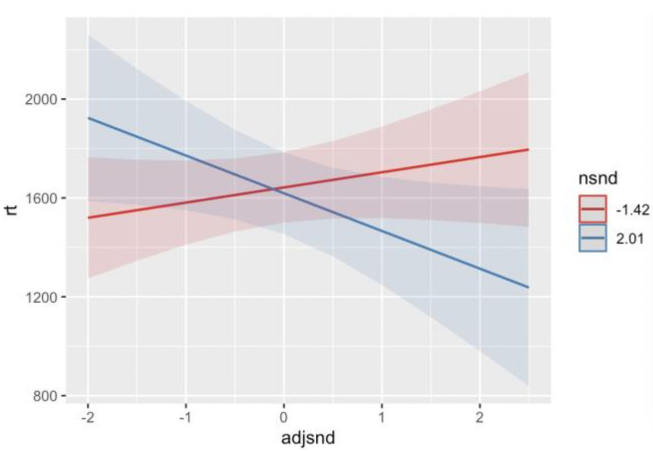
There was a main effect of meaningfulness, with faster nonsense RTs observed for low meaningful adjective-noun pairs compared to intermediate adjective-noun pairs [$t(91.40)=-4.14$, $p=<.001$, $\eta_p^2=.16$]. A typical concreteness effect was also found, with faster response times for concrete stimuli relative to abstract stimuli [$t(70.30)=3.27$, $p=.002$, $\eta_p^2=.13$].

Within each meaningful group, the same modelling procedure was followed. For nonsense judgments of low meaningful adjective-noun pairs, the final model included fixed effects for concreteness, mean OF, and combined letter length, and random intercepts for subjects and items. Data were trimmed resulting in the removal of 49 observations (2.93% of data). Skewness was .57 and kurtosis was 3.40. There was a main effect of concreteness, [$b = -.09$, $t(37.72) = -3.32$, $p = .002$, $\eta_p^2=.22$], with concrete adjective-noun pairs yielding faster reaction times than abstract adjective-noun pairs. Combined letter length and mean OF were not significant in the model ($p=.31$ and $p=.43$ respectively) but improved the model fit overall. The marginal R^2 was .02 and the conditional R^2 was .55 for the final model.

For adjective-noun pairs of intermediate meaningfulness, the final model included fixed effects for adjective SND, noun SND, and their two-way interaction and random intercepts for subjects and items. Data were trimmed resulting in the removal of 11 observations (2.11% of data). Skewness was .27 and kurtosis was 2.85. There was a two-way interaction between adjective and noun SND [b = -.04, t(54.12) = -2.18, *p* =.04, η_p^2=.09; see Figure 14], in which nonsense judgments were facilitated if both constituents of the pair had dense SNDs but response times were slower if the adjective had a sparse SND paired with a dense SND noun. Main effects for adjective SND (*p*=.42) and noun SND (*p*=.67) were not significant. The marginal R^2 was .01 and the conditional R^2 was .55 for the final model. High meaningful stimuli were not analyzed due to a low sample size (i.e., nonsense judgments were not commonly made for high meaningful stimuli).

Figure 13

Interaction Between Adjective and Noun SND for Nonsense Judgments of Intermediate Meaningfulness



Note. Mean reaction times are displayed in milliseconds.

Discussion

This series of experiments examined adjective-noun phrase processing in online tasks (Experiments 2-4) using a carefully constructed and normed stimulus set (Experiment 1). In Experiments

2-4 adjective-noun phrase processing was examined using three online processing tasks that ranged from shallow to deep levels of processing. Across these tasks, main effects of meaningfulness and concreteness with shorter RTs to highly meaningful and concrete adjective-noun phrases were expected. Whether concreteness and meaningfulness would interact was also of interest, as McAuley (2018) found no difference in low plausible stimuli based on head noun concreteness.

In the shallowest lexical decision task where non-words were non-pronounceable (Experiment 2), orthographic properties of the adjective-noun pairs were predictive of reaction times. When comparing across meaningful groups, adjective-noun combined letter length and mean orthographic frequency were predictive of response times, with faster reaction times observed for adjective-noun pairs with shorter combined letter length and higher mean orthographic frequency. Differences were observed within each meaningful group. That is, no model was predictive of reaction times for low meaningful stimuli, whereas higher mean orthographic frequency yielded faster reactions times for intermediate meaningful adjective-noun phrases, and shorter letter length yielded faster reaction times for high meaningful adjective-noun phrases. Collectively, and in contrast to predictions, in the shallowest online processing task there were neither main effects of, nor an interaction between, meaningfulness and concreteness. Rather, orthographic properties (i.e., letter length and frequency) were predictive of reaction times, and this differed depending on the meaningfulness of adjective-noun phrases. For adjective-noun pairs of low meaningfulness, orthographic properties were not predictive of reaction times. With greater meaning between the adjective-noun phrases, a frequency effect was observed. For the most familiar stimuli (i.e., those of high meaningfulness), the most surface-level orthographic property, combined letter length, contributed to differences in response times.

Experiment 3 was designed to elicit a deeper level of processing than Experiment 2 and this depth was obtained by having participants complete a double lexical decision task with pronounceable non-words as opposed to the non-pronounceable nonwords in Experiment 2. In Experiment 3, in contrast to Experiment 2, main effects of meaningfulness and concreteness were obtained. When comparing across meaningful groups, high meaningful adjective-noun pairs yielded faster responses times than low meaningful adjective-noun pairs, although no significant differences were observed for intermediate meaningful pairs. Additionally, a concreteness effect was observed, in which concrete adjective-noun pairs were processed faster than abstract adjective-noun pairs, and a typical frequency effect was also found. There were differences within each meaningful group. In contrast to Experiment 2, in Experiment 3, a typical concreteness effect was observed for the least meaningful adjective-noun group. In Experiment 2, only orthographic covariates were predictive of reaction times for intermediate and high meaningful groups. The same pattern was observed in Experiment 3, where adjective-noun pairs with higher mean orthographic frequency yielded faster reaction times than those with lower frequency for intermediate meaningful pairs, and adjective-noun pairs with shorter combined letter length yielded faster reaction times for high meaningful adjective-noun pairs. In Experiment 3 then, the hypotheses were partially supported, with differences in reaction times observed between the high and low meaningful groups (but not intermediate), and a concreteness effect observed for the least meaningful stimuli.

As the deepest level of semantic processing task, Experiment 4 was anticipated to have greater semantic effects than those found in Experiments 2 and 3. For Experiment 4, a dense SND composition of adjective and noun constituents was anticipated to yield faster response times. Based on prior research, an interaction between phrase concreteness and constituent SND was predicted (Al-Azary & Buchanan, 2017; Danguecan & Buchanan, 2016), though the exact nature of this interaction was uncertain. Prior research has found an inhibitory effect for simple abstract concepts with dense SNDs (Danguecan & Buchanan, 2016) whereas an inhibitory effect was observed for concrete metaphors with dense SNDs in a metaphor comprehension task (Al-Azary & Buchanan, 2017). In Experiment 4, participants were explicitly asked to make sense and nonsense judgments, and these responses times were analyzed separately. Compared to Experiments 2 and 3, Experiment 4 had the most semantic effects. For sense judgments, participants were faster to judge high meaningful adjective-nouns phrases compared to intermediate meaningful adjective-noun phrases. A typical concreteness effect was also observed within intermediate and high meaningful groups, in which participants were faster to respond to concrete adjective-noun pairs relative to abstract adjective-noun pairs. Additionally, constituent semantic neighbourhood (SND) effects were found. For intermediate meaningful pairs, a dense noun SND yielded an inhibitory effect, meaning slower sense judgments were observed compared to a sparser noun SND. In contrast, for adjective-noun phrases of high meaningfulness, adjective SND demonstrated a facilitative effect, in which a dense adjective SND yielded faster reaction times compared to a sparse adjective SND. Orthographic properties, including letter length and orthographic frequency, were also significant predictors of sense reaction times for high meaningful stimuli. Collectively, main effects of meaningfulness and concreteness were observed, as predicted. Further aligned with the hypotheses, a dense SND adjective constituent facilitated sense judgments of high meaningful stimuli, though noun SND effects were not observed in this group. In contrast, a dense noun SND slowed sense judgments for intermediate meaningful stimuli, contrary to what was predicted.

For nonsense judgments in Experiment 4, only low and intermediate meaningful groups were included in analysis. Nonsense judgments were quicker for low meaningful stimuli compared to intermediate meaningful stimuli, as predicted. A typical concreteness effect was observed within low meaningful adjective-noun pairs, in which nonsense judgments were faster for concrete pairs compared to abstract ones. For intermediate meaningful adjective-noun pairs, there was an interaction between constituents SNDs, in which nonsense judgments were made quicker if both adjective and noun SNDs were dense. Taken together, hypotheses were partially supported, as a main effect of meaningfulness was observed, though a concreteness effect was only observed in low meaningful adjective-noun phrases. Further aligned with hypotheses, dense SND constituents of intermediate meaningful pairs facilitated nonsense judgments.

Overall, semantic effects were sensitive to task demands as well as the amount of meaning contained within the adjective-noun phrase. As tasks required a deeper level of processing and greater semantic engagement, semantic variables such as meaningfulness, concreteness, and constituent SND played a larger role in predicting reaction times.

CHAPTER 5

EXPERIMENT 5: EXPLICIT JUDGMENT TASK

The purpose of the explicit judgment task was to identify underlying interpretation strategies of novel adjective-noun combinations that vary on semantic properties, including concreteness and semantic neighbourhood density of the adjective and noun constituents.

Method

Participants

Sixty-five participants were recruited through Prolific (prolific.co). Participants varied in sex (47 female, 18 male), age (range=18-61; *M*=35 years, *SD*=12.44), and education level (range=high school to doctorate; *M*=14 years, *SD*=2.28). All participants were native speakers of English. Participants were paid £3.75 for their participation in the 30-minute task.

Material

The 63 low meaningful adjective-noun pairs (Appendix B) were used. Constituent adjective and head noun concreteness ratings were obtained from Brysbaert et al. (2014). Constituent semantic neighbourhood densities (SNDs) were derived from WINDSORS (Durda & Buchanan, 2008) using the semantic neighbourhood app (Lutfallah & Buchanan, 2018). The experiment was designed entirely in Qualtrics (qualtrics.com) as a questionnaire with open-ended responses for participants to record their interpretations.

Procedure

Eligible participants were able to view the study on Prolific once published. Upon seeing the study advertisement, participants were directed to click a link (Qualtrics survey) if interested in conducting the study. Then, participants were directed to the consent form to read. If they consented by clicking “I agree”, they were asked to provide demographic information, including their age, sex, and education, as well as their Prolific ID. Participants were then presented with the instructions of the task, which were adapted from Murphy (1988). Participants were asked to imagine that they had heard a novel adjective-noun phrase in conversation and to come up with a meaning of the phrase that seemed most natural. They were asked to generate a meaning of adjective-noun phrases as a whole and to elaborate on the meaning rather than just simply re-state the meaning in a circular manner. Examples of acceptable (e.g., STONY MOUSTACHE is a “moustache covered in rocks”) and circular (e.g., STONY MOUSTACHE is a “moustache that is stony”) interpretations were provided. They were instructed to indicate adjective-noun phrases that they did not know the definitions for by typing “unknown” as the response. Participants were then provided the adjective-noun phrases in a random order, with an open-ended response field to enter their interpretations for each one. Once completed, participants were thanked for their participation and submitted their responses. The task took approximately 30 minutes to complete, and participants were awarded £3.75 for their participation.

Results

Data Cleaning

Five participants were removed in total; two were removed for putting “unknown” for all adjective-noun phrases and three were removed for writing nonsensical single word responses and/or leaving the majority of responses blank. One adjective-noun phrase was also removed from the analysis (i.e., mobile chaos) as upon inspection of interpretations, many participants interpreted it as a noun-noun combination (e.g., “a mobile phone in which the apps are completely disorganized”, “a vehicle that has gone crazy and caused an accident”). The remaining 60 participants and 62 adjective-noun phrases were analyzed further.

Thematic Analysis

An exploratory thematic analysis was conducted to determine the underlying strategies used to interpret adjective-noun phrases. Initially, interpretation types from noun-noun combinations, such as property-mapping and relation-linking, were identified. For example, one interpretation for SPONGY

RUMOUR was “a rumour with a lot of holes in it”, which can be conceived as property-mapping, where a salient property of spongy (i.e., porous) is mapped onto the head noun. As an example of relation-linking, one interpretation for DUSTY HEADACHE was “a headache caused by dust”, which uses the noun-modifier “causes” relation. However, the majority of interpretations did not code neatly into a property-mapping or relation-linking approach. Instead, for all adjective-noun phrases, three to four consistent themes emerged as interpretation strategies. Descriptions and examples of these interpretation types are provided in detail in Table 10.

One strategy, labelled slot-filling based on the adjective-noun conceptual combination literature (Smith et al., 1988), was a common strategy used by participants in which the adjective would modify an attribute or multiple attributes of a noun schema. Both property-mapping and relation-linking interpretation types were generally subsumed under slot-filling. Another strategy was named noun elaboration, and it emerged as a consistent theme in which participants would modify an attribute or multiple attributes of a noun schema by adding in elaborate knowledge, such as drawing on general information, personal relevance, and background knowledge. In both interpretation types the noun schema remained in the final product, but the interpretation was elaborated on in the latter strategy. For example, for ROBUST CHILL, “a very cold sensation” was coded as slot-filling whereas “cold weather that is so intense you feel it in your bones and can’t seem to warm up once inside” was coded as noun elaboration. In both instances, the underlying schema remains as CHILL, but it is clear the latter interpretation incorporated greater background knowledge. Notably, participants may have interpreted a given noun differently, particularly for abstract nouns; for example, “solo” may have been interpreted as a dance solo or musical solo. Similarly, some participants gave a superordinate (e.g., flower for orchid), subordinate (e.g., pudding for dessert), or synonym (e.g., snowstorm for blizzard) response in their interpretation. In these instances, the overall approach distinguished the interpretation types, even when the final noun referents differed among participants. For instance, “loud snowstorm” and "loud blizzard" were both coded as slot-filling and both coded as modifying the same slot (i.e., the sound of a blizzard).

A third strategy, labelled abstraction, is when the underlying noun schema was entirely changed in the final product (and was more than simply a superordinate, subordinate, or synonym). For example, "Refrigerator" was an interpretation given for ROBUST CHILL and was coded as an abstraction. As another example, one interpretation for ITCHY SKELETON was "a deep feeling of unease", which changes the noun referent entirely. The fourth and less frequent strategy identified was called adjective-reversal, which occurred when the adjective became the referent or schema, such as "green sun" for SOLAR MOLD or "controversial lyrics" for LYRICAL CONSPIRACY. Other interpretations that did not match the themes identified were labelled as "miscellaneous", either because they were circular interpretations, they did not make sense, or words were misspelled. For example, one miscellaneous response for CLUMSY REALM was "clumsy" and one for RIPE HIKE was "ready preamble". Additionally, some responses were "unknown", as participants were told to record this when they did not know the definition of the adjective or noun.

Table 10

Examples of Four Primary Interpretation Strategies Identified

		dense toad	**clingy umbrella**	**wormy solo**	**illegal boredom**
Constituent Properties		concrete/concrete low/low SND	abstract/concrete high/low SND	concrete/abstract high/low SND	abstract/abstract low/high SND
Slot-filling	adjective modifies attribute(s) of noun	"a toad with little to no intelligence", "a plump, heavy toad"	"an old, damp umbrella", "sticky umbrella"	"a solo that goes on a long time", "a dance solo with worm-like movements"	"being very bored", "when one is bored doing illegal things"

Noun Elaboration	adjective modifies attribute(s) of noun by adding in background information, general knowledge, and/or personal relevance	“a hibernating toad, curled into a small ball to outlast winter”	“a umbrella with a purposefully sticky handle, meant to be easier to hold onto even in high winds”	“Piece of music performed on a piano which is riddled with wood worm, affecting sound”	“a level of boredom so high it should be criminal”, “bored when surrounded by many Christmas gifts”
Abstraction	noun schema/referent is modified entirely	“a person who is perceived as unattractive and unintelligent”	“The desire to protect someone from the rain at all times”, “a coat with a hood, is a clingy umbrella”	“when someone gets ahead by doing something underhanded”	“an idle mind”, “how jurors feels when listening to a case”
Adjective-reversal	adjective is the final schema/referent and may be modified into a noun	--------	--------	“a worm that is alone”, “nematode alone”	“something illegal that has lost its novelty”

Coding Reliability

To ensure reliability of coding, four undergraduate research volunteers were each given ten interpretations to code (eight unique and two the same between coders). In total, 34 interpretations were randomly selected from different adjective-noun phrases. Coders were provided a description of the interpretation types and examples of each. For the two interpretations that were common to all coders, there was 100% agreement between coders and the researcher’s coding. Among the remaining 32 interpretations, twenty-one interpretation codes were aligned with the original coding and eleven were discrepant from the researcher’s coding (65% agreement). One of the coders, who was generally far more familiar with the research background and had 100% agreement with the researcher’s coding, was asked to code the 11 discrepant interpretations for a third opinion. In total, 9 were coded in agreement with the researcher and 2 were discrepant. When considering the third opinion on discrepant items, there was 94% agreement overall with the researcher’s coding.

Data Analysis

All preliminary analyses were conducted in Excel. Once coding was reliably checked, the interpretations were grouped into their respective four interpretation codes. The number of responses within each interpretation approach was tallied, as were interpretations coded as “miscellaneous” and “unknown”. The total number of interpretations for each adjective-noun phrase was calculated as the combined tallies for the four interpretation types, whereas miscellaneous and unknown responses were removed and not analyzed further. Proportions were then calculated for each interpretation type by dividing the number of interpretations within an interpretation type by the total number of interpretations for each adjective-noun phrase. For example, HILLY BANNER had 33 interpretations that were identified as slot-filling and had a total of 46 responses between the four interpretation types. As such, the proportion of slot-filling interpretations for HILLY BANNER was .72.

Unique interpretations were also summed for each adjective-noun phrase. For slot-filling interpretations, responses were only calculated as unique if the interpretation modified a different slot or modified the same slot in a different way. For example, for TENSE SODA, some responses were “a drink which tastes bad”, “bad tasting drink”, and “soda that doesn’t taste very nice”. Each of these responses modify the “taste” slot of SODA in the same way, so they are counted as a single unique interpretation. Another interpretation was “strong tasting drink”. This also modified the “taste” slot, but it did so differently and was counted as a unique interpretation. Another slot-filling interpretation was “warm soda”, which modifies the “temperature” slot and is a unique interpretation. Responses that were similar for noun elaboration, abstraction, and adjective-reversal were also grouped and only coded as a single unique interpretation. For instance, for DUSTY HEADACHE, two abstraction responses were similar, “hangover” and “like a hangover”, so they were grouped as one unique interpretation. In contrast, the response “old person” was a separate unique abstraction interpretation. Once sorted, the proportion of unique interpretations was calculated for each adjective-noun phrase by tallying the total number of unique interpretations and dividing it by the total number of interpretations. For example, CHATTY CIDER had 15 unique interpretations out of 54 total responses, which is equivalent to a proportion of .28. See Appendix E for a summary of proportions of interpretation strategies and unique interpretations for each adjective-noun phrase.

Mean and standard deviations of unique proportions and interpretation strategy proportions were computed in R and categorized based on semantic properties of the constituents. As mentioned earlier, abstract constituents were categorized as those with concreteness ratings equal to or below 3 according to Brysbaert's et al. (2014) dataset, whereas concrete constituents were categorized as those above 3. Refer to Table 11 for a summary of the proportions of interpretations based on constituent concreteness. An ANOVA was conducted in R to determine if there were any differences in the proportion of unique interpretations based on constituent concreteness. The "car" package (Fox & Weisberg, 2019) was used, which allows for Type III sums of squares for unbalanced group sizes. Assumptions of normality and homogeneity of variances were met. A significant main effect for noun concreteness was found, in which adjective-noun phrases with abstract nouns yielded more unique interpretations compared to adjective-noun phrases with concrete nouns [$F(1, 58)=9.57$, $p=.003$, $\eta_p^2=.14$]. Adjective concreteness and the interaction variable were not significant (*p'*s=.60 and .26, respectively).

Table 11

Proportion Means (SDs) Based on Constituent Concreteness

Adjective Concreteness	Noun Concreteness	N	Unique	Slot-Filling	Noun Elaboration	Abstraction	Adjective-Reversal
concrete	concrete	17	.47 (.09)	.65 (.13)	.16 (.08)	.16 (.09)	.03 (.06)
abstract	concrete	18	.42 (.12)	.70 (.14)	.15 (.09)	.14 (.12)	.02 (.05)
concrete	abstract	13	.54 (.15)	.48 (.15)	.21 (.12)	.28 (.17)	.03 (.05)
abstract	abstract	14	.56 (.14)	.46 (.17)	.31 (.12)	.22 (.12)	.01 (.02)

Proportions of interpretation strategies were analyzed descriptively. As can be seen in Table 11, a slot-filling approach was the most common strategy regardless of concreteness of the constituents. Relative to adjective-noun phrases with an abstract noun, adjective-noun phrases with a concrete noun had proportionally higher slot-filling interpretations. In contrast, adjective-noun phrases with an abstract noun had proportionally higher use of noun elaboration and abstraction strategies compared to adjective-noun phrases with a concrete noun. Notably, abstract nouns paired with concrete adjectives favoured abstraction interpretations relative to noun elaboration, whereas the opposite pattern is seen for abstract nouns paired with an abstract adjective. Adjective-reversal was the least common strategy regardless of concreteness.

Proportions were also examined based on constituent semantic neighbourhood density (SND), with sparse SND categorized as those equal to and below .53 and dense SND categorized as those above .53, based on a median split. See Table 12 for a summary of the proportions of interpretations based on constituent SND. An ANOVA was similarly conducted in R using a Type III sums of squares for an

Table 12

Proportion Means (SDs) Based on Constituent SND

Adjective SND	Noun SND	N	Unique	Slot-Filling	Noun Elaboration	Abstraction	Adjective-Reversal
dense	dense	18	.48 (.12)	.60 (.17)	.22 (.09)	.16 (.14)	.02 (.05)
sparse	dense	12	.47 (.12)	.61 (.20)	.23 (.15)	.15 (.10)	.01 (.02)
dense	sparse	12	.58 (.13)	.43 (.14)	.23 (.14)	.29 (.13)	.05 (.06)
sparse	sparse	20	.46 (.15)	.66 (.13)	.15 (.08)	.18 (.12)	.01 (.03)

unbalanced design. Assumptions of normality and homogeneity of variances were met. A significant main effect for noun SND was found, in which adjective-noun phrases with sparse SND nouns yielded more unique interpretations compared to adjective-noun phrases with dense SND nouns [$F(1, 58)=4.82$, $p=.03$, $\eta_p^2=.08$]. The interaction and main effect of adjective SND were not significant (*p's*= .09 and.88, respectively).

Proportions of interpretation strategies were analyzed descriptively. Again, regardless of constituent SND composition, a slot-filling approach was the most common strategy employed whereas adjective-reversal was the rarest strategy used. Compared to the other groups, adjective-noun phrases with a dense SND adjective and sparse SND noun had a lower proportion of slot-filling interpretations and higher proportion of abstraction interpretations. In addition, adjective-noun phrases made up of sparse constituent SNDs had a lower proportion of noun elaboration strategies compared to other groups. Lastly, adjective-noun phrases with dense noun SNDs had a higher proportion of noun elaboration interpretations compared to abstraction interpretations.

Re-Examination of Experiment 4 Analysis

Considering a wide array of meaningful interpretations were generated for the low meaningful, or novel, adjective-noun phrases, it is worthwhile to re-visit the previous experiment and examine the reaction time data for sense judgments for low meaningful adjective-noun phrases (N=292). That is, it is reasonable to postulate that making a sense judgment for a low meaningful adjective-noun phrase is due to recruiting conceptual combinatory processing and coming to a sensical interpretation of the pair. Response times were slowest for this group compared to the other two meaningful groups (refer back to Table 6). When included in the model and investigated with planned contrasts using Satterthwaite approximation for degrees of freedom (Kuznetsova et al., 2017), sense judgments for low meaningful adjective-noun phrases were significantly slower than sense judgments for intermediate meaningful phrases [$t(145.4)=3.40$, $p=.002$, $\eta_p^2=.07$] and high meaningful phrases [$t(149.3)=8.80$, $p=<.001$, $\eta_p^2=.34$]. To investigate response latencies for sense judgments of low meaningful adjective-noun phrases, the same model fitting procedure was used as outlined above. The final model included fixed effects for an interaction between adjective SND and concreteness and an interaction between adjective and noun SND. Random effects included a random intercept for participants only, as none of the error variance was related to the stimuli. After the model was fitted, data were trimmed resulting in the removal of 5 observations (1.70% of data). Skewness was .08 and kurtosis was 3.03 for the residuals. The final model is shown in Table 7. There was a significant interaction between concreteness and adjective SND,

Table 13

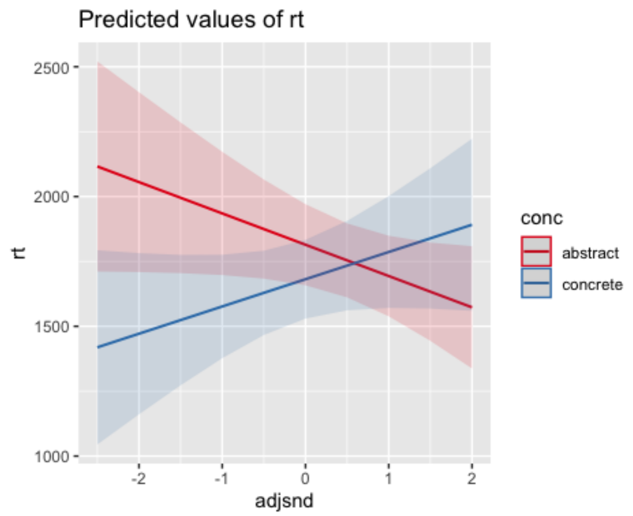
Summary of the Final Model for Sense Judgments of Low Meaningful Items

Fixed Effects					
	Estimate/ Beta	SE	95% CI	t	p
Intercept	7.41	.05	7.32 to 7.50	165.18	<.001
adj SND	-.10	.04	-.17 to -.02	-2.44	.02
concreteness-conc	-.08	.05	-.17 to .02	-1.60	.11
noun SND	-.01	.02	.06 to .03	-.59	.56
adjSND*Concrete	.15	.06	.03 to .27	2.40	.02
adjSND*nounSND	-.07	.03	-.12 to -.01	-2.36	.02
Random Effects					
			Variance	S.D.	
Participant (Intercept)			.0922	.30	
Model fit					
R^2			Marginal		Conditional
			.03		.52
Key: p-values for fixed effects calculated using Satterthwaite's approximations. Confidence Intervals have been calculated using the Wald method. Model equation: logRT ~ adjsnd*concreteness + adjsnd*nounsnd + (1 | Participant)					

displayed in Figure 17, in which there was an inhibitory effect on abstract adjective-noun pairs with a sparse adjective SND whereas a facilitative effect was observed for concrete-adjective noun pairs with a sparse adjective SND [b = .15, t(231.75) = 2.40, *p* =.02, η_p^2=.02]. No differences were observed with dense adjectives SNDs based on concreteness of the pair. There was also an interaction between adjective and noun SND, in which a dense adjective SND paired with a dense noun SND had faster

Figure 17

Interaction Between Concreteness and Adjective SND for Sense Judgments of Low Meaningful Pairs

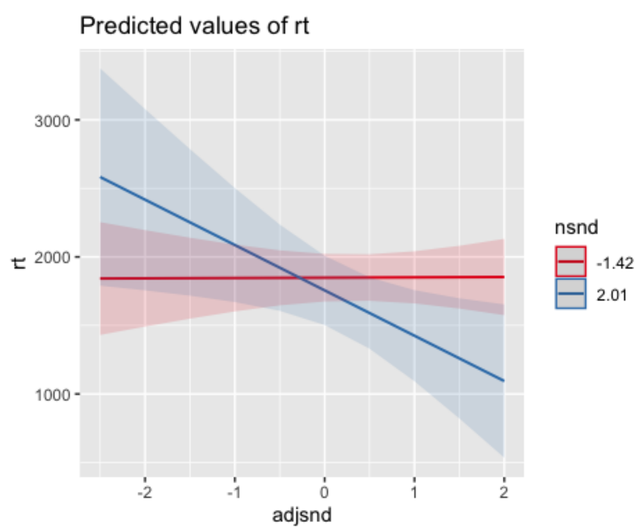


Note. Mean reaction times are displayed in milliseconds for sense judgments only.

sense reaction times relative to a sparse adjective SND paired with a dense noun SND [b = -.07, t(223.83) = -2.36, *p* =.02, η_p^2=.02]. No difference in reaction times was observed in adjective-noun phrases with sparse noun SNDs based on adjective SND. Additionally, a main effect of adjective SND was observed,

Figure 18

Interaction Between Adjective SND and Noun SND for Sense Judgments of Low Meaningful Pairs



Note: Mean reaction times are displayed in milliseconds for sense judgments only.

with faster sense judgments observed for adjective-noun pairs with a dense adjective SND relative to a sparse one [b = -.10, t(237.11) = -2.44, *p* =.02, η_p^2=.02].

Discussion

In an offline explicit judgment task, adjective-noun interpretation strategies were examined primarily in a qualitative fashion. Based on the thematic analysis, four primary interpretation approaches were identified among adjective-noun phrases with differing semantic properties. A random selection of interpretations was coded by other raters and deemed to have acceptable reliability with the researcher's coding. The four interpretation strategies include slot-filling, noun elaboration, abstraction, and adjective-reversal. The former strategy was the most common strategy used and involved modifying an attribute or characteristic of the noun schema. Noun elaboration also modified the noun schema but in an elaborate way, by incorporating background knowledge, general information, and/or personal information, whereas an abstraction approach modified the noun schema entirely. Adjective-reversal was the least common strategy used and involved a reversal where the adjective was the final schema and was often changed into its noun form, and the noun functioned as a modifier.

The overall proportion of unique interpretations and the proportion of interpretations for each of the four strategy types was compared based on constituent concreteness and constituent SND composition. Recall that adjective-noun phrases with abstract head nouns and dense constituent SNDs were hypothesized to yield the most unique interpretations, and the results partially support these hypotheses. In terms of concreteness, the results suggest that participants produce more unique interpretations for adjectives paired with an abstract noun than when paired with a concrete noun, regardless of adjective concreteness, and this is consistent with what was hypothesized. A higher proportion of slot-filling interpretations was observed for adjective-noun phrases with concrete nouns compared to abstract nouns, whereas the opposite pattern was seen for noun elaboration and abstraction strategies. Adjective concreteness also impacted the interpretation strategy used when paired with an abstract noun, in which a concrete-abstract pair was more likely interpreted with an abstraction strategy whereas an abstract-abstract noun pair was more likely interpreted with a noun elaboration strategy. With regard to SND, hypotheses were partially supported, as the most unique interpretations were observed for adjective-noun pairs with a dense adjective SND but paired with a sparse noun SND. This SND composition was shown to have a lower proportion of slot-filling interpretations but higher proportion of abstraction interpretations relative to the other SND constituent groups. In addition, a noun elaboration strategy was less commonly used if both constituents had sparse SNDs, whereas it was more common than an abstraction strategy when noun constituents had dense SNDs. Taken together, the chosen interpretation strategy was sensitive to constituent concreteness and SNDs.

Based on the qualitative findings from this study an additional analysis of data from Experiment 4 was conducted. The focus of this analysis was on sense judgments for low meaningful adjective-noun phrases, because a sense judgment suggests that participants came to a sensical interpretation via conceptual combination. Relative to high and intermediate meaningful pairs, sense judgments were slowest for low meaningful adjective-noun phrases, supporting the first hypothesis. A typical concreteness effect (Hypothesis 2) was not observed as expected. Aligned with hypotheses, a dense SND constituent composition facilitated sense judgments for low meaningful stimuli. In support of the fifth hypothesis, an interaction between SND and concreteness was observed, in which a sparse adjective SND slowed sense judgments for abstract adjective-noun phrases whereas it facilitated sense judgments for concrete adjective-noun phrases.

CHAPTER 6

GENERAL DISCUSSION

The overarching goal of this study was to shed light on semantic richness and abstractness in the conceptual combination of adjective-noun phrases, from the lens of a language-based model of concepts. To achieve this goal, this dissertation carried out an in-depth investigation of how adjective-noun phrases varying in semantic properties including meaningfulness, concreteness, and semantic neighbourhood density are processed. Unique to this study, meaningfulness and concreteness ratings of adjective-noun phrases were collected using a novel application called QMethod Software (Lutfallah & Buchanan, 2019), which utilizes Q-sort methodology to quantify participant ratings. The stimulus development (Experiment 1) resulted in an adjective-noun stimulus set stratified by meaningfulness and concreteness that was well-controlled for orthographic properties, such as combined letter length and mean orthographic frequency. Semantic neighbourhood density (SND), a measure derived from a lexical co-occurrence model (Durda & Buchanan, 2008; Lutfallah et al., 2018), served as a proxy of semantic richness for adjective and noun constituents and reflects the variability in the distribution of semantic neighbours within the semantic neighbourhood of a concept (Danguecan & Buchanan, 2016; Pexman et al., 2008). Recall that a concept can be semantically rich and surrounded by many close semantic neighbours capturing a denser SND, or semantically poor and loosely associated with semantic neighbours exhibiting a sparser SND (refer to Figure 1; Danguecan & Buchanan, 2016).

Many of the hypothesized semantic effects were observed in a graded fashion across online processing tasks based on the level of semantic processing the task demanded, in line with other studies (e.g., Danguecan & Buchanan, 2016; Siakaluk et al., 2018; Yap et al., 2011). This was similarly observed across adjective-noun phrases within tasks based on the meaningfulness between the pair. The intuitively more challenging adjective-noun phrases, such as the low meaningful, or more novel, phrases, tended to exhibit greater semantic effects. In turn, less orthographic effects were seen with increased semantic processing demands. Importantly, the presence of semantic effects suggests the activation and recruitment of conceptual knowledge in lexical or sense/nonsense decisions (James, 1975; Pexman et al., 2008), and the SND effects observed in Experiment 4 further adds to the plethora of support for language-based models of conceptual representation (e.g., see Günther et al., 2019, and Jones et al., 2015 for reviews).

Experiment 2, in which non-pronounceable non-words were used as foils in a double lexical decision task, was the shallowest level of processing task and no semantic effects were observed, despite expectations for an effect of meaningfulness (Hypothesis 1) and concreteness (Hypothesis 2). For instance, a typical concreteness effect was demonstrated with low frequency single concepts on an identical task based on a subject analysis (but not item analysis; Danguecan & Buchanan, 2016). The adjective and noun constituents used in the current study were likewise low frequency to begin with (below 30 instances per million words; Durda & Buchanan, 2008), though the current analyses opted to include orthographic frequency and letter length as control variables, and subject and item errors were analyzed within a single model. Other studies have found different effects based on whether the distractor non-words were non-pronounceable or pronounceable in lexical decision tasks. For example, James (1975) examined single concepts varying in frequency and concreteness and found a frequency effect (but not a concreteness effect) in a lexical decision task when non-words were orthographically illegal and non-pronounceable, whereas a concreteness effect was observed for low frequency (but not high frequency) nouns on a separate lexical decision task when the non-words were pronounceable. Similarly, lexical decisions for semantically related words have shown to have an advantage over semantically unrelated words, when the non-word distractor pairs were pronounceable (Meyers & Schvaneveldt, 1971), though this semantic relatedness effect is greatly reduced when the non-word pairs were non-pronounceable (Shulman & Davidson, 1977). Furthermore, findings from Gagné and Shoben (1997) demonstrated that faster lexical decisions were observed for sensical modifier-noun pairs relative to nonsensical pairs, and their lexical decision task used distractors that had a pronounceable non-word constituent in the pair.

Collectively, these findings suggest that semantic knowledge is not necessarily accessed when making lexical decisions when distractor non-words are orthographically illegal/non-pronounceable, and rather, lexical decisions in this context are driven by orthographic properties (Yap et al., 2011), such as the degree of familiarity with the word form (James, 1975). The findings of Experiment 2 support this notion, as higher mean frequency was predictive of faster reaction times for intermediate meaningful adjective-noun phrases, whereas the most surface level orthographic feature, lower combined letter length, was predictive of faster reaction times for the intuitively easiest high meaningful adjective-noun phrases. Interestingly, no model was predictive of low meaningful adjective-noun phrases regardless of the concreteness of the phrase, perhaps indicative of their higher difficulty as stimuli. Importantly, the absence of semantic effects, such as concreteness, does not necessarily mean semantic knowledge was not recruited, as differences between concrete and abstract concepts may be minimized in certain conditions (e.g., when concepts are high frequency; James, 1975), and this may similarly be the case for low meaningful adjective-noun phrases which is supported based on response times.

Further in line with these findings, hypothesized effects of meaningfulness and concreteness were observed for lexical decisions in Experiment 3, where non-words were orthographically legal and pronounceable. For this task, a main effect of meaningfulness was partially supported (Hypothesis 1), as response times were fastest for high meaningful adjective-noun phrases and slowest for low meaningful adjective-noun phrases, similar to the results of Gagné and Shoben (1997). This aligns with Potter and Falcouner's (1979) suggestion of recruitment of a slower spreading activation mechanism for novel adjective-noun combinations that requires independent activation of constituent components, relative to automatic and holistic retrieval of familiar adjective-noun phrases. Intermediate meaningful adjective-noun phrases, however, did not differ from either group. A similar pattern of response times was observed in McAuley (2018), specifically for adjective-noun pairs with an abstract head noun. A replication of this interaction (Hypothesis 3) was not observed in the current study, though there were considerable differences in the operational definitions and stimulus development of meaningfulness and concreteness between studies. Akin to Experiment 2, in Experiment 3 a frequency effect was observed in intermediate meaningful pairs and a letter length effect was seen in high meaningful pairs, whereas low meaningful adjective-noun phrases yielded a concreteness effect (Hypothesis 2; Kroll & Merves, 1986; Paivio, 1971). Taken together, it appears that semantic knowledge was recruited for lexical decisions in Experiment 3 based on the presence of semantic effects (i.e., meaningfulness and concreteness), and this was especially the case for low meaningful (i.e., less plausible and less familiar) adjective-noun phrases.

Though the lack of differentiation for intermediate meaningful adjective-noun phrases in Experiment 3 may raise skepticism of the value of this middle group and the validity of the stimulus development, the findings in Experiment 4 mitigate against these concerns. Experiment 4 required participants to make sense and nonsense judgments of adjective-noun phrases. In other words, participants needed to process the *meaning* of the phrase to make judgments, rather than simply make decisions about the *visual form* of phrases (Pexman et al., 2008; Yap et al., 2011). Full support for a meaningfulness effect (Hypothesis 1) was observed in Experiment 4, in which response times for sense judgments were graded across meaningfulness of the phrase, and nonsense judgments were faster for low meaningful adjective-nouns compared to intermediate meaningful pairs. This is consistent with findings from Wisniewski and Murphy (2005) and Murphy (1990), in which sense judgments were faster for plausible and familiar (versus implausible and unfamiliar) noun-noun phrases and for adjective-noun pairs where the adjective reflected a typical (versus atypical) attribute of the noun, respectively. Slower response times on the sense/nonsense judgment task (relative to the lexical decision tasks) are likely indicative of recruitment of deeper conceptual processing, or *conceptual combination,* as participant “sense” decisions suggest that they came to a sensical interpretation for the phrase (via conceptual combination) whereas a “nonsense” decision suggests a rejection of a sensical interpretation (Gagné & Spalding, 2013). The main effect of meaningfulness additionally supports this notion, as conceptual combinatory processing would be sensitive to the meaningfulness of the phrase, with slower sense decisions expected with less semantic meaning in the phrase. This is in line with the view that conceptual combination is recruited during the processing of both familiar and novel modifier-noun phrases based on theory (Gagné & Spalding, 2006, 2009; Günther & Marelli, 2016; Libben, 2014), empirical evidence (Gagné, 2001; Gagné & Spalding, 2004; Estes & Jones, 2008), and neuroscience studies (e.g., Coutanche et al., 2019; El Yagoubi et al., 2008).

A concreteness effect (Hypothesis 2) was also observed for nearly all meaningful groups in Experiment 4, except for sense decisions for low meaningful phrases and nonsense decisions for intermediate meaningful adjective-noun phrases. Similar to low meaningful phrases in Experiment 2, differences in concreteness and abstractness may have been minimized in these latter groups (James, 1975). Nevertheless, the current study extended the finding of a robust concreteness effect to include when concreteness is rated at the level of the adjective-noun phrase as a whole, specifically in tasks that engage semantic processing. Consistent with the second hypothesis, concreteness ratings of the head noun constituent (from Brysbaert et al., 2014) were predictive of concreteness ratings at the phrase level that were collected in Experiment 1. Contrary to predictions, adjective concreteness affected phrase concreteness ratings differently dependent on head noun concreteness. For concrete head nouns, adjective concreteness (or abstractness) yielded higher (or lower) phrase concreteness ratings, whereas adjective concreteness did not affect phrase abstractness ratings for adjectives paired with abstract head nouns (refer back to Figure 4).

This unique finding highlights the asymmetry in concrete and abstract modifier-noun relationships and in turn, their conceptual representations. As concrete concepts are believed to be defined by their intrinsic properties (Wiemer-Hastings & Xu, 2005), a concrete adjective paired with a concrete noun may further concretize the final adjective-noun representation by identifying specific concrete features that increase the specificity of the final referent (e.g., SILKY ROBE, TORN ENVELOPE), whereas an abstract adjective paired with a concrete noun is more detrimental to the concreteness of the overall phrase by adding ambiguity and instability to the overall representation (e.g., UNIQUE CANVAS, ROMANTIC PICNIC). Thus, in the former instance, concrete nouns are further constrained in their meaning by being paired with a concrete adjective, whereas meaning is less constrained when paired with an abstract adjective, consistent with the nature of concrete and abstract concepts (Crutch & Warrington, 2005). Lucas et al. (2017) similarly demonstrated differences in the processing of noun-noun phrases with concrete head nouns based on modifier imageability. In line with the proposed rationale, EEG findings showed that a higher imageability modifier was associated with a larger N700 potential (suggestive of recruitment of mental imagery) and were rated as easier to interpret relative to a lower imageability modifier-concrete head noun phrase (Lucas et al., 2017). On the other hand, abstract concepts generally have inherently less stable representations and greater interindividual variability based on linguistic and social experiences (Borghi & Binkofski, 2014; Borghi et al., 2017). When adjectives are paired with abstract head nouns, neither concrete (e.g., TOXIC CONSPIRACY, SLICK SARCASM) nor abstract (e.g., SINCERE CONFESSION, DECEPTIVE FRAUD) adjectives may resolve the ambiguity inherent in the abstract head noun concept, and in essence, adjective concreteness does not substantially contribute to the abstractness of the phrase.

In addition, SND effects from Experiment 4 demonstrate the asymmetry inherent in modifier-noun relationships (Ran & Duimering, 2009), with a novel twist, in that the nature of the asymmetric direction of the relationship differed based on the meaningfulness of adjective-noun phrases. Recall that the fourth hypothesis predicted that semantically rich constituent concepts would facilitate processing times, in line with findings in single concepts (e.g., Pexman et al., 2008). In the current study, constituent semantic richness demonstrated both facilitative and inhibitory effects on meaning construction, and this was dependent on the meaningfulness of the phrase. When adjective-noun phrases were highly meaningful, meaning construction was facilitated by a dense adjective semantic neighbourhood whereas a sparse adjective semantic neighbourhood slowed semantic processing. In contrast, for intermediate meaningful phrases, a semantically dense head noun demonstrated an inhibitory effect on sense judgments whereas a semantically sparse head noun facilitated semantic processing. Likewise, meaning construction was rejected faster for intermediate meaningful adjective-noun phrases when both constituents were semantically dense. For low meaningful adjective-noun phrases, however, meaning construction was facilitated when both adjective and noun concepts had dense semantic neighbourhoods. Furthermore, modifier SND interacted with concreteness of the phrase (Hypothesis 5), particularly for low meaningful adjective-noun phrases. For novel abstract adjective-noun phrases, a sparse adjective semantic neighbourhood slowed meaning construction, whereas for novel concrete adjective-noun phrases, a sparse adjective semantic neighbourhood facilitated meaning construction. Taken together, the cluster of SND effects and interactions observed in Experiment 4 are considered in light of findings of similar effects in related fields, as no models of conceptual combination incorporate a role for semantic richness and abstractness of the concepts.

In contrast to the dense SND inhibitory effect seen in Danguecan and Buchanan (2016) for single abstract concepts, abstract phrases with sparser adjective semantic neighbourhoods slowed semantic processing of phrases with low meaningfulness in the current study. Considering task demands, in the former case task demands emphasized word recognition whereas in the latter, task demands emphasized meaning construction. Danguecan and Buchanan (2016) interpreted their findings as reflective of the semantic complexity of abstract words, which are thought to be represented by complex linguistic associations (consistent with Crutch and Warrington, 2005; Vigliocco et al., 2009), the complexity of which can be captured by large scale co-occurrence patterns in language. If word recognition requires activation of the concept via spreading activation, this process would take longer for abstract concepts with complex linguistic associations (e.g., close neighbours are abstract and consequently have complex and unstable representations themselves), whereas concrete concepts are more largely represented by sensorimotor properties (e.g., features of nouns; Danguecan & Buchanan, 2016), perhaps circumventing the need for extensive activation. In contrast, in the current study when novel phrases were abstract with a semantically sparse adjective, meaning construction was hindered as a few complex linguistic associations may be less conducive to adjective modification relative to sensorimotor properties that define concrete noun concepts, especially considering that sensorimotor information can be linguistically denoted by adjectives and verbs (which are meant to describe features and action of nouns).

Turning to figurative language, recall that Al-Azary and Buchanan (2017) found that metaphors consisting of concrete topics and vehicles with dense semantic neighbourhoods (e.g., a Pen is a Sword) were rated as less comprehensible and processed slower on a comprehensibility task compared to semantically dense metaphors with abstract topics (e.g., Language is a Bridge). Their findings suggested that semantic richness (referring to concrete concepts with dense SNDs in this case) was not conducive to constructing the meaning of metaphors (Al-Azary & Buchanan, 2017). Interestingly, in the current study, participants were asked to make meaningful sense judgments for adjective-noun phrases, some of which were novel (i.e., low meaningful), and a different pattern of concreteness and SND interactions was seen under these task conditions. Semantically rich adjective-noun phrases, in the form of a dense semantic constituent composition for novel pairs, facilitated semantic processing in the current study. However, slower meaning construction was observed for novel concrete adjective-noun phrases with a dense modifier SND compared to those with a sparse modifier SND, which demonstrated a facilitative effect on sense judgments. This finding suggests that perhaps novel adjective-noun phrases that are “too” semantically rich (i.e., concrete phrases with dense constituent semantic neighbourhoods) are not conducive to meaning construction, which is similar to Al-Azary and Buchanan's (2017) findings in metaphors. Although their study did not find differences in metaphors with sparse semantic neighbourhoods based on topic concreteness (Al-Azary & Buchanan, 2017), the current study found that semantic processing was slower for novel adjective-noun phrases with sparse modifier semantic neighbourhoods, particularly for abstract phrases.

Al-Azary and Buchanan's (2017) findings were interpreted considering Kintsch's (2000) predication algorithm, which is a computational model of sentence meaning based on latent semantic analysis (i.e., similar to language-based co-occurrence models; Durda & Buchanan, 2008). When processing metaphors or literal sentences, Kintsch's (2000) approach involves creating a spreading activation network between the topic (akin to predicate or modifier) and vehicle (akin to subject or head noun) within high-dimensional vector space. To construct meaning, the algorithm searches the neighbourhood of the topic for associations that are related to the neighbourhood of the vehicle and activates shared neighbours while inhibiting unrelated neighbours. As a result, the meaning of the phrase is derived as a new vector composed of the merged vectors (i.e., the topic, vehicle, and their shared neighbours). Although Kintsch (2000) hypothesized that semantic richness would facilitate this process, Al-Azary and Buchanan's (2017) findings suggest that semantic richness, in the form of dense semantic neighbourhoods and concrete topics, was detrimental to metaphor comprehension in particular. They hypothesized that having many close semantic neighbours and concrete features increases the complexity in the algorithm by slowing down and disrupting the search for shared neighbours while simultaneously requiring more suppression of irrelevant properties (Al-Azary & Buchanan, 2017). Additionally, case study findings by Al-Azary et al. (2019) were consistent with this proposed extension to the predication algorithm, as a participant with deep dyslexia rated metaphors with sparse semantic neighbourhoods and abstract topics as most comprehensible and metaphors with dense semantic neighbourhoods as least comprehensible regardless of topic concreteness.

A recent empirical endeavour by Al-Azary et al. (2021) further supported this notion and is relevant to the current study. In their study, Al-Azary et al. (2021) examined modifier-noun metaphors varying in noun concreteness and constituent semantic neighbourhood density, including adjective-noun and noun-noun metaphors, and asked participants to evaluate the literalness of modifier-noun phrases (i.e., literally true versus literally false decisions). Past research has demonstrated a phenomenon called the metaphor interference effect (MIE; Glucksberg et al., 1982), in which evaluating metaphors for literalness is thought to be conflicted by the automatic computation of the metaphorical meaning produced by the predication algorithm. As such, slower literally false judgments are typically observed for metaphors compared to their scrambled metaphor counterparts (Glucksberg et al., 1982).

Al-Azary et al. (2021) demonstrated comparable MIE's in adjective-noun and noun-noun metaphors, but the size of the MIE effect was dependent on word-level semantics for adjective-noun metaphors only. They similarly found that semantic richness was detrimental to adjective-noun metaphor processing, as semantically dense adjective-noun constituent pairs as well as semantically dense adjectives paired with concrete head nouns were both rejected rapidly, resulting in a diminished MIE. In particular, a semantically rich *adjective* modifier was proposed to be asymmetrically detrimental to computing metaphoric meaning, as a sparse adjective SND paired with a concrete head noun slowed rejection of the phrase, resulting in an enhanced MIE (Al-Azary et al., 2021). This was similarly observed in the current study with novel concrete adjective-noun phrases, where *too* semantically rich was detrimental to processing and this effect was likewise asymmetrically driven by the modifier. Collectively, Al-Azary et al.'s (2021) findings extended Kinstch's (2000) predication algorithm to encompass processing of modifier-noun phrases, demonstrated differences in early activation of finer-grain semantics within phrase types (noun-noun versus adjective-noun), and proposed that semantic richness within the modifier constituent in adjective-noun metaphors was particularly unfavourable to constructing figurative meaning.

Given that Kintsch's (2000) predication algorithm has been modeled in literal and metaphoric sentence meaning construction, and recently extended to encompass meaning integration in modifier-noun phrases (Al-Azary et al., 2021), it would be a parsimonious explanation and mechanism to apply in creating meaning for other modifier-noun phrases, in essence, capturing conceptual combination. Conceptual combination is ultimately semantic composition through integration of two single concepts to create a complex concept with new meaning. Similar to the pursuits of Gagné and Spalding (2009) who propose a unitary meaning construction process when interpreting familiar or novel compound words (e.g., HUMBUG) and modifier-noun phrases (e.g., MOUSE SHOE, SILKY MOON; also see Günther & Marelli, 2016; Libben, 2014), a single base mechanism may underlie all conceptual meaning construction, at least early in processing, including modifier-noun phrases, metaphors, sentences, and essentially any processing of linguistic stimuli to generate semantic meaning (i.e., semantic processing). This notion is additionally supported by having a unitary neurobiological substrate proposed in both sentence construction and conceptual combination (i.e., the left anterior temporal lobe; Baron et al., 2010; Baron & Osherson, 2011; Pylkkänen, 2016; Vandenberghe, 2002) and clinical populations having degradation of semantic and conceptual knowledge with damage to this region (e.g., Lambon Ralph et al., 2012; Magnusdottir et al.,2013). Importantly, there may be downstream nuances in deeper processing of literal versus figurative language meaning for instance (e.g., recruitment of right hemisphere in metaphor comprehension; Schmidt & Seger, 2009), or when task demands evoke embodiment (e.g., recruitment of perceptual systems; Bergen, 2015) which is consistent with mixed embodiment accounts (Barsalou et al, 2008; Louwerse, 2007).

Experiment 4's semantic effects can largely be understood in the framework of Kintsch's (2000) predication algorithm, while expanding on the findings described above. The proposed explanation is outlined below:

- Slower processing times reflect an increased time searching for shared neighbors and inhibiting irrelevant ones to compute meaning.
    - The speed of sense decisions was dependent on phrase meaningfulness, with slower response times observed with less meaningful phrases in a graded fashion. As phrases have less meaning, there are fewer shared neighbours to begin with, effectively slowing the search.
    - The speed of nonsense decisions was dependent on phrase meaningfulness, with slower response times observed for intermediate meaningful phrases compared to low meaningful phrases. When task demands enforce a time pressure, a continued search for shared neighbours is worthwhile when the phrase has some meaning whereas the search is abandoned quicker for most low meaningful phrases as they do not have many shared neighbours to begin with, similar to scrambled metaphors (Al-Azary et al., 2021). As such, a nonsense judgment reflects a failure to find coherent meaning when attempted under time pressure.
- A semantically dense **modifier** constituent can facilitate the search for shared neighbors and consequently in computing meaning for adjective-noun phrases (i.e., resulting in faster sense judgments), as proposed by Kintsch (2000), particularly when phrases are highly meaningful. However, for novel phrases with few shared neighbours and many irrelevant neighbours, "too" semantically rich is detrimental for meaning computation, particularly in the **modifier** constituent (similar to metaphors; Al-Azary & Buchanan, 2017; Al-Azary et al., 2021).
    - For high meaningful adjective-noun phrases, the algorithm does not have to search very long for shared neighbours (like literal phrases; Al-Azary et al., 2021), so a dense modifier semantic neighbourhood further facilitates this process resulting in

faster sense decisions. Likewise, a sparser adjective SND makes it increasingly difficult to identify shared neighbors, resulting in a slower sense response latency.

    - When a search is pursued and successful in low meaningful adjective-noun phrases, dense constituent semantic neighbourhoods, as well as concrete phrases with a *sparse* adjective SND, facilitate the search to compute meaning resulting in faster sense decisions. Concrete adjective-noun phrases with a denser adjective SND are arguably “too” semantically rich, as they were slower to accept as sensical compared to a sparser adjective SND, supporting the notion that there may be a gradient of semantic richness that is beneficial for computing meaning of more novel adjective-noun phrases under time pressure.

- Relatedly, a semantically sparse **modifier** constituent slowed the search for shared neighbours and meaning construction, particularly for novel abstract phrases, again highlighting the notion that some degree of semantic richness is preferred to construct meaning under time constraints.
    - As mentioned, abstract conceptual representations are thought to be linguistically complex but less stable in their representation (e.g., vary by context, between individuals, etc.; Borghi & Binkofski, 2014; Borghi et al., 2017) compared to concrete representations, which are hypothesized to benefit from rich sensorimotor features (Wiemer-Hastings & Xu, 2005). In support of this, slower decisions were observed for abstract adjective-noun phrases across nearly all conditions, suggesting that a search for shared neighbours and inhibition of irrelevant neighbours is on average slower for abstract phrases.

    - Further strengthening this claim, meaning construction was especially slowed in novel abstract phrases when the modifier constituent was semantically sparse, making it increasingly difficult to identify shared neighbours.
- A larger effect for the **head noun** was observed for those adjective-noun phrases with intermediate meaningfulness in which a dense **head noun** SND was particularly detrimental to constructing meaning.
    - Recall that head nouns were the same across all meaningful groups, so it is the modifier that differs between groups. As such, it is the modifier that determines the meaningfulness of the overall phrase. Intermediate meaningful phrases are mediocre as a pair as they are somewhat familiar and somewhat plausible, and they would have some shared neighbours and some irrelevant neighbours. For these mediocre pairs, a semantically dense head noun may have too many irrelevant features to inhibit, resulting in slower meaning construction. In contrast, a sparse head noun has few irrelevant features to inhibit making it easier to identify the few shared neighbours with the modifier constituent, resulting in faster response latencies.
    - When intermediate meaningful phrases have dense semantic neighbourhoods for both adjective and head noun constituents, there are too many irrelevant features to search through to identify the few shared neighbours and integrate into a coherent meaning under time pressure, resulting in a faster rejection of sensicality.

The mechanism proposed by the predication algorithm (Kintsch, 2000) fits well with language-based co-occurrence models of conceptual representations, which are similarly derived through latent semantic analysis and propose a mechanism of spreading activation between semantic neighbours of a concept (Durda & Buchanan, 2008). In line with language-based models, the statistical distribution and co-occurrence of words can be aggregated across linguistic contexts to quantify a basis of how concepts and conceptual relations are mapped in the human mind (Buchanan et al., 2001; Lung & Burgess, 1996), and a spreading activation mechanism between conceptual representations could capture meaning construction as modelled by the predication algorithm (Kintsch, 2000). Likewise, other models of conceptual combination conceptualize representations as based in statistical use in language (ECCo; Lynott & Connell, 2011) or propose a role for statistically based relational information in processing, even if not tied to the conceptual representations (e.g., relational frequency; RICE, Gagné & Spalding, 2013). The proposed explanation can elucidate points made by other theories of conceptual combination, such as a prominent role of the modifier (CARIN; Gagné & Shoben, 1997), an interaction between modifier and head noun constituents (e.g., RICE, Gagné & Spalding, 2013; Concept Specialization Model, Cohen & Murphy, 1984; Interactive Property Attribution Model, Estes & Glucksberg, 2000; Interactional Hypothesis, Maguire et al., 2010), and involvement in both familiar and novel phrases (e.g., RICE, Gagné & Spalding, 2013). Other important predictions in theories of conceptual combination include competition among different interpretations that needs to be resolved for meaning construction (e.g., Interactional Hypothesis, Maguire et al., 2010; RICE, Gagné & Spalding, 2013; ECCo, Lynott & Conell, 2011) as well as incorporation of background knowledge (e.g., RICE, Gagné & Spalding, 2013; Concept Specialization Model, Cohen & Murphy, 1984). These latter predictions are better examined by Experiment 5, where participants were asked to provide an interpretation of novel adjective-noun phrases.

As mentioned, most low meaningful adjective-noun combinations were deemed non-sensical in Experiment 4 under timed task demands. However, on an offline judgment task with no time constraints imposed (Experiment 5), participants constructed numerous unique interpretations for low meaningful adjective-noun phrases. In further support of a unitary mechanism underlying semantic construction of simple and complex phrases, participants often composed sentences, sometimes elaborate, to convey the meaning of modifier-noun phrases. There were four main types of interpretations identified based on thematic analysis (refer back to Table 10). One primary interpretation type was labelled as “slot-filling” to be consistent with the field of conceptual combination. Recall that slot-filling is a mechanism describing schematic modification of the head noun by altering a feature denoted by the adjective (e.g., Smith et al., 1988) and is based on schematic conceptual representations. However, many models of conceptual combination propose that schematic representations consist of properties based on prior knowledge (e.g., Smith et al., 1988), and as demonstrated by Riordan and Jones (2011), distributional statistical models can account for feature-based information (see also Baroni & Lenci, 2010; Durda et al., 2009), with prior knowledge referring to exposure in language contexts. Therefore, a noun’s schema is thought to be captured within a noun’s semantic neighbourhood and can be modelled within a statistically derived language-based conceptual representation (e.g., such as WINDSORS; Durda & Buchanan, 2008).

Interpretations were categorized as “slot-filling” when the adjective modified an attribute or multiple attributes of a noun (Medin & Shoben 1988; Smith et al., 1988), and this was the most common approach employed across all adjective-noun phrases, regardless of constituent concreteness or semantic richness. This intuitively makes the most sense when considering adjectives and nouns linguistic purpose in language and consequently their semantic representations. An adjective’s role is to describe a noun, and adjective conceptual representations likely reflect this function. Based on syntax, adjective and nouns often co-occur in linguistic contexts. As such, adjective semantic representations likely have many nouns as semantic neighbours, and noun semantic representations also likely have many adjectives as semantic neighbours. For low meaningful or more novel adjective-noun phrases, the adjective is likely to be very distantly associated with the noun’s semantic neighbourhood and vice versa. However, low meaningful adjective-noun phrases can still be interpreted due to prior existing knowledge of how adjective-noun phrases function in language (i.e., knowing that adjectives are meant to describe or modify features of the noun).

This notion is similarly mirrored in the competition among relations in nominals (CARIN; Gagné & Shoben, 1997) which proposes that prior statistical distributional knowledge, or prior information about how nouns relate to one another, is accessed when constructing meaning of noun-noun combinations, which are syntactically incorrect in structure. Likewise, prior knowledge that adjectives typically modify an attribute of the noun can direct semantic processing. Further, Gagné and Shoben's (1997) priming experiments showed that nouns may have relational preferences because of the frequency of the relational occurrence in language (e.g., the <u>made of</u> relation occurs frequently when CHOCOLATE is the modifier noun), in turn making the relation easier to access during semantic processing. Likewise, nouns may have typical attributes that are preferably modified based on co-occurrence in language, such as the <u>colour</u> attribute may be more commonly modified in the concept CAR than <u>weight</u> (e.g., more likely to hear "check out that RED car" than "check out that HEAVY car"). This is in line with the Selective Modification Model's (SMM; Smith et al., 1988) proposal that noun conceptual representations contain information about diagnosticity of attributes (i.e., slots) and have associated weights or salience for each adjective feature (i.e., fillers or values) that reflect their prototypicality, based on subjective frequency and perceptibility. In both instances, prior exposure to the noun-noun relation or attribute modification facilitates processing.

Additionally, past research hypothesizes that spreading activation occurs more quickly for familiar adjective-noun phrases relative to novel adjective-noun phrases based on response times, in which the latter phrases are thought to require independent activation of constituent concepts (Potter & Falcouner, 1979). Thus, the frequency and familiarity of relations and attribute modifications in language may translate to the efficiency of spreading activation between modifier-noun concepts. Within the framework of the predication algorithm (Kintsch, 2000), individual constituents would be activated for both familiar and novel phrases, but the spread of activation would be more rapid in the former case, and this is supported by the meaningfulness effect observed in Experiment 4. Additionally, activation of individual constituents aligns with morphological decomposition or full-parsing models of compound word processing (Libben, 1998; Schreuder & Baayen, 1997; Taft, 2004).

Notably, a slot-filling approach was not only the most common interpretation strategy for adjectives paired with concrete nouns, but it was also the most common strategy for adjectives paired with abstract nouns. However, attributes of concrete and abstract concepts differed qualitatively. Similar to the nature of concrete and abstract conceptual representations (e.g., Kousta et al., 2011; Vigliocco et al., 2009), concrete attributes appeared more engrained in sensory-motor experiences (e.g., weight as it relates to size) whereas abstract attributes were more affective-based (e.g., weight as it relates to emotional intensity). To illustrate, for a concrete adjective-concrete head noun pair, one slot-filling interpretation for FLABBY BARN was "a barn which is fat" whereas for an abstract-abstract pair, one slot-filling interpretation of STEEP CRUSH was "heavy crushing on someone". Like concrete nouns, abstract nouns may also have certain attribution preferences. For example, <u>severity</u> may be more likely modified for BOREDOM than <u>purity</u> (e.g., more likely to hear "this boredom is EXTREME" than "this boredom is DIRTY"), but the opposite may be true for PRANK (e.g., "dirty prank" versus "extreme prank"). In addition, adjectives themselves also likely have preferences in how they modify a noun. For instance, GRAY probably most often modifies the <u>colour</u> of a noun, but it may also modify the <u>age</u> of hair or the <u>severity</u> of clouds (Medin and Shoben, 1988). As such, attribute preferences are likely a part of both adjective and noun conceptual representations based on their use in language and are interactively activated via spreading activation, consistent with theories of conceptual combination that propose that meaning is a function of the interaction between modifier and head noun constituents (e.g., Estes & Gluckberg, 2002; Gagné & Spalding, 2013; Maguire et al., 2010).

In the current study, slot-filling interpretations were diverse themselves, as a single adjective could modify different or multiple attributes of the noun. For example, VOCAL BLIZZARD may be interpreted by explicitly modifying the <u>sound</u> attribute (e.g., “a loud snowstorm”) or <u>strength</u> attribute (e.g., “strong snowstorm”) or both (e.g., “a very loud and stormy blizzard”), though as indicated by Medin and Shoben (1988), attributes are likely correlated with one another (e.g., a strong snowstorm is also likely to be loud). The variety of slot-filling interpretations produced by a single adjective-noun combination is in line with the proposal that numerous meanings of adjectives are automatically computed during adjective-noun processing (Mullaly et al., 2010), and suggests that multiple interpretations could be competing for selection.

Similarly, RICE (Gagné & Spalding 2013; Spalding et al., 2010) suggests that multiple relational structures may be evaluated in noun-noun processing, generating competition amongst relations and consequently, interpretation possibilities. The competition is assumed to be resolved by evaluating relational interpretations in the context of the semantic and relational availability within the noun. Similarly, the Concept Specialization Model (Cohen & Murphy, 1984) proposes that noun attributes are activated by context (i.e., in this case, the adjective) and values of the attributes are filled by an adjective based on prior background knowledge. Considering the predication algorithm (Kintsch, 2000), interpretation would proceed via spreading activation of adjective and noun constituent representations in which the algorithm searches the modifier for shared attributes of the head noun and inhibits irrelevant ones, producing a representation composed of the modifier, head noun, and shared attributes. Resolution of competition then, may be based on the frequency of attribute modification reflected in the co-occurrence of words and consequently reflected in the efficiency of spreading activation (and the search-and-inhibit process) between concepts. Additionally, the diversity in attribute modification is a testament to the individual variability in a concept’s representation, which language-based models acknowledge by speculating that conceptual representations are uniquely shaped by an individual’s linguistic environments (Buchanan et al., 2001). In light of Experiment 4 and 5 findings, the resolution of competition may take longer when there are multiple unique processing routes, such as for novel abstract adjective-noun phrases, which had slower processing times in Experiment 4 when accepted as sensical and more unique interpretations produced in Experiment 5 compared to concrete adjective-noun phrases. However, parallel findings with SND (i.e., more unique interpretations for dense-sparse SND pairs translating to slower processing in Experiment 4) were not found, though these phrase types may have been underrepresented in the smaller sample of low meaningful phrases that were deemed sensical by participants.

According to the RICE model, following resolution of the competition, the final interpretation is elaborated by recruiting extralinguistic knowledge (Gagné & Spalding 2013; Spalding et al., 2010). Many other theories also propose the recruitment of background knowledge (e.g., Cohen & Murphy, 1984; Maguire et al., 2010; Medin & Shoben, 1988). Relatedly, in the current study, world knowledge (e.g., a DENSE TOAD is “a hibernating toad, curled into a small ball to outlast the winter” or a JADED INSECT is “a bug that has learned that a specific situation may result in harm, so they learn to avoid that situation”) and personal experiences (e.g., CHATTY CIDER is “the right level of cider for having a good time” or ABSTRACT PASTRY is “The kind of pies I make when I’m bored of cooking, again”) and other elaborations were incorporated into interpretations and were categorized as a “noun elaboration” strategy. From a language-based lens, world knowledge plays a role in constructing individual conceptual representations (Borghi et al., 2013), and in essence, a concept’s semantic neighbourhood. Interestingly, information was also drawn from episodic memories (e.g., personal experiences). Though beyond the scope of the current study, this is consistent with literature suggesting interdependency between episodic and semantic memory systems (e.g., Graham et al., 2000; Greenberg & Verfaellie, 2010; Takashima et al., 2014), such as words with dense SNDs are remembered better on episodic memory tasks (Wong-Gonzalez, 2018). In comparison to a slot-filling approach, a noun elaboration strategy may intuitively take more time than a simpler modification of the noun. As such, an elaboration strategy may be a consequence of time spent conceptually combining concepts, and this is a testable hypothesis for future studies to explore.

"Abstraction" was a third interpretation strategy identified, where an entirely new conceptual representation was formed from the constituent concepts (e.g., "old person" for DUSTY HEADACHE). In the ECCo model (Lynott & Connell, 2010), this approach is similar to destructive processing when interpreting noun-noun combinations (e.g., property-mapping approach; Connell & Lynott, 2011a; Wisniewski, 1997; Wisniewski & Love, 1998), though in this instance, both the adjective and noun conceptual representations are deconstructed into a new representation. Similarly, adjective representations are deconstructed in both slot-filling and noun elaboration approaches. On the other hand, the fourth type of interpretation strategy, called "adjective reversal", can be viewed as a reversal slot-filling approach, as the adjective was converted into its noun form and functioned as a noun and vice versa. In other words, the noun representation was deconstructed to represent a feature. This strategy was rarely applied, but it was twice as common when the adjective was concrete (e.g., wormy to worm; solar to sun). For example, a reversal interpretation for ITCHY SKELETON was "an itch so bad it feels like it goes down to your bones".

Aside from time spent combining the concepts and constructing meaning, the four interpretation strategies may be differentiated by the intactness of the constituent representations in the final representation based on Kintch's (2000) model (similar to destructive and non-destructive processing in ECCo; Lynott & Connell, 2010). Nouns are a broad syntactic class that arguably dominate language contexts, and other syntactic classes typically function in relation to nouns, such as to describe features (i.e., adjective) or actions (i.e., verbs) of the noun. Consequently, noun constituent representations are often left intact in the final combined representation, and this is supported by Experiment 5 findings, in which the noun dominates the final representation of slot-filling and noun elaboration interpretations, making up the vast majority of overall interpretations. Similarly, in noun-noun combinations, both noun constituent representations are left intact in relation linking interpretations (Gagné & Shoben,1997; Connell & Lynott, 2011a), which is also the dominant approach when interpreting noun-noun combinations (Gagné, 2000). In contrast, when slot-filing was reversed between the adjective and noun (i.e., adjective reversal interpretations), the adjective representation dominated in the final combined representation.

Recall that the predication algorithm suggests that meaning is a result of the combined activation of the modifier concept, noun concept, and their shared attributes. Perhaps the distribution of activation differs amongst the adjective and noun concepts and is reflective of their overall contribution to the final combined representation, such that the noun concept is the default dominant activation pattern as exemplified in slot-filling and noun elaboration approaches, whereas in a reversal the adjective concept has the largest proportion of activation and is consequently most intact in the final interpretation. An important clarification to make is that shared attributes are not viewed as separate from the conceptual representations themselves, but rather as represented in both adjective and noun representations (e.g., STEEP and CRUSH are associated with *falling* leading to interpretations like “when you fall for someone very quickly”). As novel adjective-noun phrases likely have very few shared neighbours, other potential semantic neighbours may be activated to search for meaning (i.e., irrelevant features could be reactivated). For abstraction interpretations, an entirely new conceptual representation is constructed suggesting that both adjective and noun concepts are minimized in the final representation, and instead their semantic neighbours may be used to construct a new representation and dominate the distribution pattern. For example, DUSTY may be related to concepts such as *old* and *gray* and the concept HEADACHE can be associated with *people* (e.g., “you are giving me a headache!”), resulting in the (rather offensive) abstraction “old person”. Whether there is a processing cost associated with an abstraction strategy is a fruitful question for future research to test out this hypothesis.

As illustrated, any of the four interpretation strategies could be applied to interpret a single adjective-noun phrase. Interpretation approach as a function of processing time was proposed as one explanation, potentially leading to differences in the distribution pattern of activation among concepts. Additionally, individual differences likely contribute as well (e.g., creativity), especially considering participants were provided no context to facilitate an interpretation. Furthermore, in the current study, fine-grained semantic features (i.e., concreteness and SND) were shown to have some influence on the interpretation approach. Aligned with predictions (Hypothesis 6), more unique interpretations were produced for adjective-noun phrases with an abstract head noun compared to a concrete one. As mentioned, concrete concepts are thought to have a more stable referent and less ambiguity (Borghi et al., 2017; Crutch & Warrington, 2005), potentially leading to circumscribed processing routes. In contrast, more diverse interpretations for abstract head nouns concepts may translate to a variety of potential processing routes, which may be related to the slower response times observed for abstract adjective-noun phrases in Experiment 4 as mentioned (i.e., it takes longer to resolve the competition between different processing routes).

Interindividual variability in abstract concept representation (Borghi & Binkofski, 2014; Borghi et al., 2017) also likely accounts for the diversity in interpretations. Relative to adjective-nouns with abstract head nouns, those with concrete head nouns had a higher proportion of interpretations categorized as slot-filling. This is reflected in the literature, as slot-filling approaches have dominated the research scope for concrete modifier-noun phrases (e.g., Smith et al., 1988), and as mentioned, concrete concepts are rich in sensorimotor attributes (Kousta et al., 2011; Vigliocco et al., 2009) which are adept to be modified via a slot-filling approach. In contrast, complex linguistic associations define abstract conceptual representations (Borghi et al., 2013; Crutch & Warrington, 2005; Dove, 2011) and this is reflected in proportionately higher elaborate interpretation strategies, such as noun elaboration and abstraction approaches. If noun elaboration and abstraction are associated with a processing cost, this could be another potential reason that slower processing times were observed for novel abstract phrases in Experiment 4. Interestingly, the interpretation type for abstract noun phrases was dependent on adjective concreteness, in which concrete-abstract phrases had more interpretations categorized as abstraction relative to abstract-abstract phrases which had more noun elaboration interpretations. Perhaps the rich imagistic features contained in the concrete adjective concept facilitates a basis for an entirely new conceptual representation, whereas when the adjective is abstract it provides more complex associations to elaborate on the abstract head noun.

Constituent semantic neighbourhood density was also examined in this context. It was predicted that dense SND constituents would provide a rich network of information to incorporate into interpretations, leading to more unique interpretations. This hypothesis was partially supported (Hypothesis 6), as a semantically *dense* adjective paired with a semantically *sparse* head noun produced the most unique interpretations. Relative to other adjective-noun phrases, this semantic composition of adjective-noun phrases (i.e., dense SND-sparse SND combination) also had the least proportion of slot-filling interpretations and the most abstraction interpretations. In these instances, a semantically sparse head noun may be less favoured to conceptually dominate the final representation, and conceptual change may be promoted when paired with a semantically dense adjective constituent. When both constituents were sparse, however, the default may still favour the noun representation. With only sparsely associated information to incorporate into interpretations, there may be a reduced potential for elaboration, and this is supported by the lower proportion of noun elaboration interpretations with semantically sparse-sparse phrases. In both instances, when the head noun semantic neighbourhood was sparse, there were apparent differences in the preferred interpretation strategy based on modifier SND, suggesting that the adjective plays a larger role in meaning construction when the head noun is semantically poor. In contrast, when the head noun was semantically dense, the proportion of interpretations across the four types was nearly identical regardless of modifier SND. This is opposite to what was observed in Experiment 4 when time pressures were imposed, where novel adjective-noun phrases with a sparse SND head noun yielded no differences in response latency based on modifier SND, though differences were observed when the head noun was semantically dense. As such, Experiment 5 SND findings may reflect downstream differences in processing that are not captured under imposed time pressures. Taken together, the semantic composition of adjective-noun phrases impacted the preferred interpretation approach for novel meaning construction and is a fruitful area for further investigation.

Incidentally, the intended adjective-noun phrase MOBILE CHAOS was interpreted as a noun-noun combination (e.g., "a mobile phone in which the apps are completely disorganized") in a third of the interpretations (*N*=20). This allowed a unique glimpse of the interpretation approaches when both nouns are abstract in a noun-noun combination. Firstly, nineteen of the interpretations were unique, which is another testament to the creativity with abstract concepts. One interpretation could be classified as relation-linking (e.g., "chaos on the phone"), and some could be classified as property-mapping, though it was reversed in that a salient feature of the head noun was mapped onto the modifier (e.g., "a mobile phone in which the apps are completely disorganised"). The former interpretation example could also be classified as slot-filling (e.g., modify the <u>location</u> attribute of chaos) whereas the latter could be a reversed slot-filling (e.g., modify the <u>arrangement</u> attribute of a phone). A noun elaboration approach was also identified, but interestingly it was often an elaboration of the modifier (e.g., "a car jam on the intersection", "a vehicle that has gone crazy and caused an accident") rather than the head noun. Although some slot-filling reversals were elaborate in the current study, there were too few reversal interpretations to meaningfully distinguish among non-elaborative and elaborative types, but it appears this may be more meaningful to differentiate in abstract noun-noun combinations. Abstraction interpretations were also identified (e.g., "network problems", "when too many people convert vans into homes"). Collectively, abstractness in noun-noun combinations warrants a thorough investigation for a complete theory of conceptual combination in noun-noun phrases.

Conclusion

The semantic processing of modifier-noun phrases involves a cognitive process called conceptual combination, which constructs an integrated meaning of the two constituent concepts (Murphy, 1988; Wisniewski, 1996). The nature of the conceptual representations themselves is relevant to understanding the underlying mechanism(s) involved in conceptual combination. Schema-based conceptual models are prevalent in theories of conceptual combination (Cohen & Murphy, 1984; Estes & Glucksberg, 2000; Smith et al., 1988; Wisniewski, 1997) though they are largely circumscribed to concrete conceptual representations. This dissertation holistically examined semantic processing of adjective-noun phrases varying in semantic properties including meaningfulness, abstractness, and semantic richness, from a language-based model of concepts. Language-based models assume that semantic knowledge of concrete and abstract concepts is derived and organized through the co-occurrence patterns in linguistic contexts, and that sensory-motor and feature-based information that compose a concepts schema are acquired through language environments and in turn, subsumed within a concepts semantic neighbourhood (Baroni & Lenci, 2010; Durda et al., 2009; Riordan & Jones, 2011). Therefore, adjective and noun conceptual representations would be reflective of their use in language (i.e., adjectives often co-occur with nouns and are meant to describe nouns) and their representations would instantiate common co-occurrences in linguistic environments (e.g., a car is more likely to described by its colour than weight). A language-based model of concepts is supported in the current study by the SND effects observed in Experiments 4 and 5 and adds to the rich empirical support for these models (e.g., see Günther et al., 2019, and Jones et al., 2015 for reviews).

Other theories of conceptual combination similarly emphasize co-occurrence patterns in language as a means for interpreting familiar and novel modifier-noun phrases (e.g., relational frequency; Gagné & Shoben, 1997; Gagné & Spalding, 2013). In adjective-noun phrases, Smith et al. (1988) suggested that noun and adjective conceptual representations contain information reflecting the salience of prototypicality (e.g., *red* is a more salient value for APPLE than *round*) and diagnosticity (e.g., shape is more relevant to BANANA than APPLE) of features and attributes, based on their subjective frequency and perceptibility. From a language-based lens, these feature and attribute preferences are thought to be inherent to both concrete and abstract conceptual representations themselves and a consequence of their co-occurrence in language. Within these models, concepts are thought to be activated via a spreading activation mechanism resulting in activation of a target concept and its nearby semantic neighbors (i.e., semantic neighbourhood; Buchanan et al., 2001). In this case, the efficiency of the spreading activation mechanism may reflect the familiarity and plausibility between adjective and noun concepts, as purported by other theories and empirical studies (e.g., Buchanan et al., 2001; Collins & Loftus, 1975; Potter & Falcouner, 1979). The current study further supports this notion, as a meaningfulness effect was observed in a graded fashion across tasks that increasingly recruited deeper semantic processing.

Kintsch's (2000) computational model, which has been applied to understand meaning construction of sentences (Kintsch, 2000), metaphors (Al-Azary & Buchanan, 2017), and modifier-noun metaphors (Al-Azary et al., 2021), was used as a basis for modelling meaning construction in adjective-noun phrases in the current study. To interpret adjective-noun phrases in the framework of his model, adjective and noun conceptual representations (i.e., semantic neighbourhoods) would be co-activated, and a search-and-inhibit process would pursue to activate shared neighbours and inhibit irrelevant neighbours, with the final integrated conceptual representation resulting from the combined activation of the noun, adjective, and shared neighbours (Kintsch, 2000). In line with other studies in metaphors, including adjective-noun metaphors (Al-Azary et al., 2021; Al-Azary & Buchanan, 2017), the current study found that a high degree of semantic richness of the modifier constituent was detrimental to meaning construction of novel adjective-noun phrases under time constraints (presumably via slowing the search-and-inhibit process), though some degree of semantic richness was preferred. Extending previous findings, the current study found asymmetrical SND effects dependent on the meaningfulness and abstractness of the adjective-noun phrase and applied Kintsch's (2000) computational model as a mechanism for conceptual combination in adjective-noun phrases.

Applying Kintsch's (2000) model to understand Experiment 4 and 5 findings can illustrate important points made by other theories of conceptual combination, such as a prominent role of the modifier noun (Gagné & Shoben, 1997), an interaction between modifier and head noun constituents (Cohen & Murphy, 1984; Estes & Glucksberg, 2000; Gagné & Spalding, 2013; Maguire et al., 2010), involvement in both familiar and novel phrases (Gagné & Spalding, 2013), competition among different interpretations that needs to be resolved for meaning construction (Gagné & Spalding, 2013; Lynott & Conell, 2011a; Maguire et al., 2010) and incorporation of background knowledge into final representations (Cohen & Murphy, 1984; Gagné & Spalding, 2013). Based on the numerous unique interpretations produced for novel phrases in Experiment 5 consisting of four primary themes, it was hypothesized that the final combined conceptual representation may have a different pattern of distribution among the component parts. For instance, the activation pattern may be preferentially dominated by the noun, consistent with a nouns prominent use in language and exemplified by the greater proportion of slot-filling and noun elaboration approaches in Experiment 5. Notably, the semantic composition of adjective-noun phrases (i.e., their concreteness and semantic neighbourhood density) impacted early semantic processing under time constraints (Experiment 4) as well as revealed nuances in downstream semantic processing without imposed time constraints (Experiment 5). As an example, novel abstract phrases were processed slower than concrete phrases when making sensicality decisions in Experiment 4, and these phrase types also yielded more unique interpretations in Experiment 5, primarily of a noun elaboration or abstraction type, which may partly explain why abstract phrases have a processing cost early on (i.e., more processing routes, more elaborate processing).

In summary, the current dissertation studied a conceptual combinatory process in familiar and novel adjective-noun phrases, while considering the abstractness of the phrase and semantic density of the constituents to better understand the conceptual representations and mechanism involved in this process. This dissertation provides additional support for a language-based model of concepts and extends the computation model proposed by Kintsch (2000) as a model to describe a conceptual combination process within adjective-noun phrases varying in semantic properties. This dissertation highlights early and downstream semantic effects during semantic processing of adjective-noun phrases that were not examined previously, and it proposes four primary interpretation strategies to interpreting adjective-noun phrases, including slot-filling, noun elaboration, abstraction, and adjective-reversal.

Limitations and Future Directions

This dissertation sought to comprehensively examine semantic processing in adjective-noun phrases, and some limitations and future directions of this research are outlined below. Although the current study was designed to have a well-controlled stimulus set and robust analyses accounting for orthographic variables, there are many other orthographic, phonological, and semantic variables that have been shown to effect processing times (e.g., imageability, age of acquisition; Khanna & Cortese, 2021; Morrison & Ellis, 2010) that were not examined in this study. This is a common consideration across all psycholinguistic research given the number of influential word-level variables, and this may be a relevant factor that is contributing to the considerable amount of unexplained variance in Experiments 2 to 4 linear mixed models. Similarly, different types of adjective-noun phrases were not distinguished in this study (e.g., intersective versus subjective phrases; Drašković et al., 2013; Kamp & Partee, 1995), although it is reasonable to assume conceptual qualitative differences exist between phrases and would be a fruitful area for future research to explore. Relatedly, variables relevant to embodied cognition (e.g., body-object interaction) were not included and were beyond the scope for the current study but are certainly worthy of future investigations given the growing support for mixed language-embodiment accounts (Barsalou et al., 2008; Louwerse, 2007) as well as the incorporation of sensory-motor information in participant interpretations in Experiment 5. Neuroimaging investigations may be worthwhile in this context to examine whether perceptual-based inputs are activated during meaning construction, particularly on an explicit judgment task where downstream processing can be captured.

This dissertation attempted to capture an array of tasks that differed in their level of semantic processing. Other tasks could have been included to better capture this semantic gradient, particularly between Experiments 3 and 4, such as a word/pseudohomophone DLDT or DLDT with word-nonword distract pairs. This is another direction for future investigations to enable a complete examination of semantic processing in adjective-noun phrases. In addition, all experimental data (i.e., Experiments 1 to 5) were gathered on Prolific, which is a United Kingdom English-speaking sample of the population, though the stimulus set was constructed based on Canadian English and co-occurrence models were derived from language contexts in North America (Durda & Buchanan, 2008). This is primarily due to concerns raised about the quality of data collection on a well-known North American platform (i.e., Amazon Mechanical Turk; mturk.com) and COVID-19 restrictions preventing in-person data collection on a Canadian university campus. As such, data collection in a North American population to examine the generalizability of the findings would be another research endeavour. Worth noting, although demographics were not collected for Experiment 1, all meaningfulness and concreteness ratings were collected from Prolific and the participants recruited in Experiments 2 to 5 were drawn from the Prolific population as well. With regard to Experiment 5, an abundant number of interpretations were gathered (i.e., 3780 interpretations) and coded independently by the researcher. It would have been more rigorous to have a second blind coder for all interpretations, but instead a small random sample of interpretations were distributed and coded by four undergraduate researchers to examine interrater reliability given the daunting amount of data collected for this task.

This dissertation provided many testable hypotheses for future directions in the field of conceptual combination. For instance, whether the framework provided by the predication algorithm (Kintsch, 2000) can be used to model meaning construction in other modifier-noun phrases (e.g., noun-noun, verb-nouns) and provide a parsimonious explanation of semantic processing is a conducive area warranting further research. Another fruitful direction is whether the mechanism of conceptual combination in adjective-noun phrases is a function of the efficiency of spreading activation among constituent concepts, based on frequency of prior exposure to constituent attribute modifications in language (similar to relational frequency in noun-noun phrases as proposed by Gagné & Spalding, 2013). Relatedly, investigating whether processing time results in different interpretation approaches (e.g., slot-filling versus abstraction) is of interest, and if processing costs are associated with different interpretation strategies as would be predicted (e.g., slower processing involved in a noun elaboration interpretation relative to a slot-filling interpretation). Lastly, this dissertation highlighted the relevance of considering conceptual abstractness in models of conceptual combination, in line with other fields of psycholinguistics that are increasingly turning attention to qualitative differences in abstract conceptual representations.

REFERENCES

Al-Azary, H., & Buchanan, L. (2017). Novel metaphor comprehension: Semantic neighbourhood density interacts with concreteness. *Memory & Cognition*, *45*(2), 296–307. https://doi.org/10.3758/s13421-016-0650-7

Al-Azary, H., Gagné, C. L., & Spalding, T. L. (2021). Flute birds and creamy skies: The metaphor interference effect in modifier–noun phrases. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 75*(2), 175–181. https://doi.org/10.1037/cep0000251

Al-Azary, H., McAuley, T., Buchanan, L., & Katz, A. N. (2019). Semantic processing of metaphor: A case-study of deep dyslexia. *Journal of Neurolinguistics*, *51*, 297-308. https://doi.org/10.1016/j.jneuroling.2019.04.003

Andrews, M., Frank, S., & Vigliocco, G. (2014). Reconciling embodied and distributional accounts of meaning in language. *Topics in Cognitive Science*, *6*(3), 359–370. https://doi.org/10.1111/tops.12096

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. https://doi.org/10.1080/09541440340000123

Arcara, G., Marelli, M., Buodo, G., & Mondini, S. (2014). Compound headedness in the mental lexicon: An event-related potential study. *Cognitive Neuropsychology*, *31*(1-2), 164-183. https://doi.org/10.1080/02643294.2013.847076

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. https://doi.org/10.1006/jmla.1997.2509

Bak, T. H., O'Donovan, D. G., Xuereb, J. H., Boniface, S., & Hodges, J. R. (2001). Selective impairment of verb processing associated with pathological changes in Brodmann areas 44 and 45 in the motor neurone disease–dementia–aphasia syndrome. *Brain*, *124*(1), 103-120. https://doi.org/10.1093/brain/124.1.103

Baron, S. G., & Osherson, D. (2011). Evidence for conceptual combination in the left anterior temporal lobe. *Neuroimage*, *55*(4), 1847-1852. https://doi.org/10.1016/j.neuroimage.2011.01.066

Baron, S. G., Thompson-Schill, S. L., Weber, M., & Osherson, D. (2010). An early stage of conceptual combination: Superimposition of constituent concepts in left anterolateral temporal lobe. *Cognitive Neuroscience, 1*(1), 44-51. https://doi.org/10.1080/17588920903548751

Baroni, M., & Lenci, A. (2010). Distributional memory: A general framework for corpus-based semantics. *Computational Linguistics*, *36*(4), 673-721. https://doi.org/10.1162/coli_a_00016

Baroni, M., & Zamparelli, R. (2010, October). Nouns are vectors, adjectives are matrices: Representing adjective-noun constructions in semantic space. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing* (pp. 1183-1193). Association for Computational Linguistics.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577–660. https://doi.org/10.1017/S0140525X99002149

Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. In M. De Vega, A. M. Glenberg & A. C. Graesser (Eds.)., *Symbols, embodiment, and meaning* (pp. 245-283). Oxford University Press.

Barsalou, L. W., Solomon, K. O., & Wu, K. (1999). Perceptual simulation in conceptual tasks. In M. K. Hiraga, C. Sinha, & S. Wilcox (Eds.)., *Cultural, psychological and typological issues in cognitive linguistics* (209-228). John Benjamins Publishing Company.

Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating abstract concepts. In D. Pecher & R. A. Zwaan (Eds.)., *Grounding cognition: The role of perception and action in memory, language, and thought* (pp. 129-163). Cambridge University Press.

Bartoń, K. (2020). *MuMIn: Multi-model inference*. https://CRAN.R-project.org/package=MuMIn

Bates, D., Maechler, M., Bolker, B., Walker, S., & Haubo Bojesen Christensen, R. (2015). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7. 2014.

Bemis, D. K., & Pylkkänen, L. (2013). Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. *Cerebral Cortex*, *23*(8), 1859-1873. https://doi.org/10.1093/cercor/bhs170

Ben-Shachar M, Lüdecke D, Makowski D (2020). effectsize: Estimation of Effect Size Indices and Standardized Parameters. Journal of Open Source Software, 5(56), 2815. doi: 10.21105/joss.02815

Bergen, B. (2015). Embodiment, simulation and meaning. In N. Riemer (Ed.), The Routledge handbook of semantics (pp. 158-173). Routledge.

Binder, J. R. (2016). In defense of abstract conceptual representations. *Psychonomic Bulletin & Review*, *23*(4), 1096–1108. https://doi.org/10.3758/s13423-015-0909-1

Binder, J. R., Conant, L. L., Humphries, C. J., Fernandino, L., Simons, S. B., Aguilar, M., & Desai, R. H. (2016). Toward a brain-based componential semantic representation. *Cognitive Neuropsychology*, *33*(3–4), 130–174. https://doi.org/10.1080/02643294.2016.1147426

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767-2796. https://doi.org/10.1093/cercor/bhp055

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993-1022.

Bock, J. S., & Clifton, C. (2000). The role of salience in conceptual combination. *Memory & Cognition*, *28*(8), 1378-1386. https://doi.org/10.3758/BF03211838

Borghi, A. M., & Binkofski, F. (2014). *Words as social tools: An embodied view on abstract concepts* (Vol. 2). Springer.

Borghi, A. M., Binkofski, F., Castelfranchi, C., Cimatti, F., Scorolli, C., & Tummolini, L. (2017). The challenge of abstract concepts. *Psychological Bulletin*, *143*(3), 263–292. https://doi.org/10.1037/bul0000089

Borghi, A. M., Scorolli, C., Caligiore, D., Baldassarre, G., & Tummolini, L. (2013). The embodied mind extended: using words as social tools. *Frontiers in Psychology*, *4*(214), 1-10. https://doi.org/10.3389/fpsyg.2013.00214

Brooks, T. L., & Cid de Garcia, D. (2015). Evidence for morphological composition in compound words using MEG. *Frontiers in Human Neuroscience*, *9*(215), 1-8. https://doi.org/10.3389/fnhum.2015.00215

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904-911. https://doi.org/10.3758/s13428-013-0403-5

Buchanan, L., Hildebrandt, N., & MacKinnon, G. E. (1996). Phonological processing of nonwords in deep dyslexia: Typical and independent?. *Journal of Neurolinguistics*, *9*(2), 113-133. https://doi.org/10.1016/0911-6044(96)00001-2

Buchanan, L., Westbury, C., & Burgess, C. (2001). Characterizing semantic space: Neighborhood effects in word recognition. *Psychonomic Bulletin & Review*, *8*(3), 531–544. https://doi.org/10.3758/BF03196189

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.)., *Language production* (Volume 2; pp. 257-294). Academic Press.

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, *10*(5), 425-455. https://doi.org/10.1080/01690969508407111

Byrne, B. M. (2010). Structural equation modeling with AMOS: Basic concepts, applications, and programming. New York: Routledge.

Cohen, B., & Murphy, G. L. (1984). Models of concepts. *Cognitive Science*, *8*(1), 27-58. https://doi.org/10.1016/S0364-0213(84)80024-5

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407 –428. https://doi.org/10.1037/0033-295X.82.6.407

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. In D. A. Balota & E. J. Marsh, *Cognitive psychology*. Psychology Press.

Connell, L., & Lynott, D. (2011a). Interpretation and representation: Testing the Embodied Conceptual Combination (ECCo) theory. In B. Kokinov, A. Karmiloff-Smith, & N. J. Nersessian (Eds.), *European Perspectives on Cognitive Science*. New Bulgarian University Press.

Connell, L., & Lynott, D. (2011b). Modality switching costs emerge in concept creation as well as retrieval. *Cognitive Science*, *35*(4), 763–778. https://doi.org/10.1111/j.1551-6709.2010.01168.x

Connell, L., & Lynott, D. (2013). Flexible and fast: Linguistic shortcut affects both shallow and deep conceptual processing. *Psychonomic Bulletin & Review*, *20*(3), 542–550. https://doi.org/10.3758/s13423-012-0368-x

Cohen, B., & Murphy, G. L. (1984). Models of Concepts. *Cognitive Science*, *8*(1), 27–58. https://doi.org/10.1207/s15516709cog0801_2

Costello, F., & Keane, M. T. (1997). Testing the Constraint Theory of Combination. In *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (Vol. 19, pp. 137-142). Psychology Press.

Costello, F. J., & Keane, M. T. (2000). Efficient creativity: Constraint-guided conceptual combination. *Cognitive Science*, *24*(2), 299–349. https://doi.org/10.1207/s15516709cog2402_4

Costello, F. J., & Keane, M. T. (2001). Testing two theories of conceptual combination: Alignment versus diagnosticity in the comprehension and production of combined concepts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 255-271. https://doi.org/10.1037/0278-7393.27.1.255

Coutanche, M.N., Solomon, S.H., and Thompson-Schill, S.L. (2019). Conceptual Combination. In D. Poeppel, G.R. Mangun and M.S. Gazzaniga (Eds.), *The cognitive neurosciences*, (6th ed.). MIT Press.

Crutch, S. J., Connell, S., & Warrington, E. K. (2009). The different representational frameworks underpinning abstract and concrete knowledge: Evidence from odd-one-out judgements. *The Quarterly Journal of Experimental Psychology*, *62*(7), 1377-1390. https://doi.org/10.1080/17470210802483834

Crutch, S. J., & Jackson, E. C. (2011). Contrasting graded effects of semantic similarity and association across the concreteness spectrum. *The Quarterly Journal of Experimental Psychology*, *64*(7), 1388-1408. https://doi.org/10.1080/17470218.2010.543285

Crutch, S. J., Ridha, B. H., & Warrington, E. K. (2006). The different frameworks underlying abstract and concrete knowledge: Evidence from a bilingual patient with a semantic refractory access dysphasia. *Neurocase*, *12*(3), 151-163. https://doi.org/10.1080/13554790600598832

Crutch, S. J. & Warrington, E. K. (2005). Abstract and concrete concepts have structurally different representational frameworks. *Brain*, *128*(3), 615–627. https://doi.org/10.1093/brain/awh349

Danguecan, A. N., & Buchanan, L. (2016). Semantic neighborhood effects for abstract versus concrete words. *Frontiers in Psychology*, *7*(1034), 1-15. https://doi.org/10.3389/fpsyg.2016.01034

De Jong, N. H., Feldman, L. B., Schreuder, R., Pastizzo, M., & Baayen, R. H. (2002). The processing and representation of Dutch and English compounds: Peripheral morphological and central orthographic effects. *Brain and Language*, *81*(1-3), 555-567. https://doi.org/10.1006/brln.2001.2547

Dove, G. (2009). Beyond perceptual symbols: A call for representational pluralism. *Cognition*, *110*(3), 412-431. https://doi.org/10.1016/j.cognition.2008.11.016

Dove, G. (2011). On the need for embodied and dis-embodied cognition. *Frontiers in Psychology*, *1*(242), 1-13. https://doi.org/10.3389/fpsyg.2010.00242

Dove, G. (2014). Thinking in words: Language as an embodied medium of thought. *Topics in Cognitive Science*, *6*(3), 371–389. https://doi.org/10.1111/tops.12102

Downing, P. (1977). On the creation and use of English compound nouns. *Language*, *53*(4), 810-842. https://doi.org/10.2307/412913

Drašković, I., Pustejovsky, J., & Schreuder, R. (2013). Adjective-noun combinations and the generative lexicon. In J. Pustejovsky, P. Bouillon, H. Isahara, K. Kanzaki, & C. Lee (Eds.)., *Advances in generative lexicon theory* (pp. 181-202). Springer.

Duñabeitia, J. A., Avilés, A., Afonso, O., Scheepers, C., & Carreiras, M. (2009). Qualitative differences in the representation of abstract versus concrete words: Evidence from the visual-world paradigm. *Cognition*, *110*(2), 284–292. https://doi.org/10.1016/j.cognition.2008.11.012

Durda, K., & Buchanan, L. (2008). WINDSORS: Windsor improved norms of distance and similarity of representations of semantics. *Behavior Research Methods*, *40*(3), 705–712. https://doi.org/10.3758/BRM.40.3.705

Durda, K., Buchanan, L., & Caron, R. (2009). Grounding co-occurrence: Identifying features in a lexical co-occurrence model of semantic memory. *Behavior Research Methods*, *41*(4), 1210-1223. https://doi.org/10.3758/BRM.41.4.1210

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. https://doi.org/10.1075/ml.8.1.04elb

El Yagoubi, R., Chiarelli, V., Mondini, S., Perrone, G., Danieli, M., & Semenza, C. (2008). Neural correlates of Italian nominal compounds and potential impact of headedness effect: An ERP study. *Cognitive Neuropsychology*, *25*(4), 559-581. https://doi.org/10.1080/02643290801900941

Esopenko, C., Gould, L., Cummine, J., Sarty, G., Kuhlmann, N., & Borowsky, R. (2012). A neuroanatomical examination of embodied cognition: semantic generation to action-related stimuli. *Frontiers in Human Neuroscience*, 6(84), 1-12. https://doi.org/10.3389/fnhum.2012.00084

Estes, Z. (2003). A tale of two similarities: Comparison and integration in conceptual combination. *Cognitive Science*, *27*(6), 911-921. https://doi.org/10.1207/s15516709cog2706_4

Estes, Z., & Glucksberg, S. (2000). Interactive property attribution in concept combination. *Memory & Cognition*, *28*(1), 28–34. https://doi.org/10.3758/BF03211572

Estes, Z., Jones, L. L., Leech, R., Mareschal, D., & Cooper, R. P. (2008). Relational processing in conceptual combination and analogy. *Behavioral and Brain Sciences*, *31*(4), 385-386. https://doi.org/10.1017/S0140525X08004548

Flick, G., Oseki, Y., Kaczmarek, A. R., Al Kaabi, M., Marantz, A., & Pylkkänen, L. (2018). Building words and phrases in the left temporal lobe. *Cortex*, *106*, 213-236. https://doi.org/10.1016/j.cortex.2018.06.004

Fodor, J. A. (1975). *The language of thought* (5th ed.). Harvard University Press.

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression*, Third edition. Sage, Thousand Oaks CA. https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

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. https://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. https://doi.org/10.1037/0278-7393.27.1.236

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. https://doi.org/10.1037/0278-7393.23.1.71

Gagné, C. L., & Shoben, E. J. (2002). Priming relations in ambiguous noun-noun combinations. *Memory & Cognition*, *30*(4), 637–646. https://doi.org/10.3758/BF03194965

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. https://doi.org/10.1016/S0093-934X(03)00459-0

Gagné, C. L., & Spalding, T. L. (2006). Using conceptual combination research to better understand novel compound words. *SKASE Journal of Theoretical Linguistics*, *3*(2), 9-16.

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. https://doi.org/10.1016/j.jml.2008.07.003

Gagné, C. L., & Spalding, T. L. (2013). Conceptual composition: The role of relational competition in the comprehension of modifier-noun phrases and noun–noun compounds. *Psychology of Learning and Motivation,* 59, 97-130. https://doi.org/10.1016/B978-0-12-407187-2.00003-4

Gallant, J., & Libben, G. (2019). No lab, no problem: Designing lexical comprehension and production experiments using PsychoPy3. *The Mental Lexicon*, *14*(1), 152-168. https://doi.org/10.1075/ml.00002.gal

Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), Perceiving, acting, and knowing: Toward an ecological psychology (pp. 67-82). Hillsdale, NJ: Erlbaum.

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development*, *1*(1), 23-64. https://doi.org/10.1207/s15473341lld0101_4

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*(3), 558-565. https://doi.org/10.3758/BF03196313

Glenberg, A. M., & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse Processes*, *28*(1), 1–26. https://doi.org/10.1080/01638539909545067

Glucksberg, S., Gildea, P., & Bookin, H. B. (1982). On understanding nonliteral speech: Can people ignore metaphors?. *Journal of Verbal Learning and Verbal Behavior,* 21(1), 85-98. https://doi.org/10.1016/S0022-5371(82)90467-4

Goldinger, S. D., Papesh, M. H., Barnhart, A. S., Hansen, W. A., & Hout, M. C. (2016). The poverty of embodied cognition. *Psychonomic Bulletin & Review*, *23*(4), 959–978. https://doi.org/10.3758/s13423-015-0860-

Graham, K. S., Patterson, K., Powis, J., Drake, J., & Hodges, J. R. (2002). Multiple inputs to episodic memory: Words tell another story. *Neuropsychology, 16*(3), 380–389. https://doi.org/10.1037/0894-4105.16.3.380

Greenberg, D. L., & Verfaellie, M. (2010). Interdependence of episodic and semantic memory: Evidence from neuropsychology. *Journal of the International Neuropsychological Society, 16*(5), 748-753. https://doi.org/10.1017/S1355617710000676

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review, 114*(2), 211–244. https://doi.org/10.1037/0033-295X.114.2.211

Guevara, E. (2011, January). Computing semantic compositionality in distributional semantics. In *Proceedings of the Ninth International Conference on Computational Semantics* (pp. 135-144). Association for Computational Linguistics.

Günther, F., & Marelli, M. (2016). Understanding karma police: The perceived plausibility of noun compounds as predicted by distributional models of semantic representation. *Plos One*, *11*(10), e0163200. https://doi.org/10.1371/journal.pone.0163200

Günther, F., & Marelli, M. (2019). Enter sandman: Compound processing and semantic transparency in a compositional perspective. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(10), 1872–1882. https://doi.org/10.1037/xlm0000677

Günther, F., Petilli, M. A., & Marelli, M. (2020). Semantic transparency is not invisibility: A computational model of perceptually-grounded conceptual combination in word processing. *Journal of Memory and Language*, *112*, 104104. https://doi.org/10.1016/j.jml.2020.104104

Günther, F., Rinaldi, L., & Marelli, M. (2019). Vector-space models of semantic representation from a cognitive perspective: A discussion of common misconceptions. *Perspectives on Psychological Science*, *14*(6), 1006-10. https://doi.org/10.1177/1745691619861372

Hair, J., Black, W. C., Babin, B. J. & Anderson, R. E. (2010) Multivariate data analysis (7th ed.). Upper Saddle River, New Jersey: Pearson Educational International.

Hamilton, A. C., & Martin, R. C. (2010). Inferring semantic organization from refractory access dysphasia: Further replication in the domains of geography and proper nouns but not concrete and abstract concepts. *Cognitive Neuropsychology*, *27*(8), 614-635. https://doi.org/10.1080/02643294.2011.609541

Hampton, J. A. (1995). Testing the prototype theory of concepts. *Journal of Memory and Language*, *34*(5), 686-708. https://doi.org/10.1006/jmla.1995.1031

Hampton, J. A. (1987). Inheritance of attributes in natural concept conjunctions. *Memory & Cognition*, *15*(1), 55–71. https://doi.org/10.3758/BF03197712

Hampton, J. A. (1988). Overextension of conjunctive concepts: Evidence for a unitary model of concept typicality and class inclusion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(1), 12 –32. https://doi.org/10.1037/0278-7393.14.1.12

Hampton, J. A. (1997). Conceptual combination. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts, and categories* (pp. 133-160). MIT Press.

Harris, Z. S. (1954). Distributional structure. *Word*, *10*(2-3), 146-162. https://doi.org/10.1080/00437956.1954.11659520

Holcomb, P. J., Kounios, J., Anderson, J. E., & West, W. C. (1999). Dual-coding, context-availability, and concreteness effects in sentence comprehension: An electrophysiological investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(3), 721–742. https://doi.org/10.1037/0278-7393.25.3.721

Huang, H. W., & Federmeier, K. D. (2015). Imaginative language: What event-related potentials have revealed about the nature and source of concreteness effects. *Language and Linguistics*, *16*(4), 503-515. https://doi.org/10.1177/1606822X15583233

Huang, H.-W., Lee, C.-L., & Federmeier, K. D. (2010). Imagine that! ERPs provide evidence for distinct hemispheric contributions to the processing of concrete and abstract concepts. *NeuroImage*, *49*(1), 1116–1123. https://doi.org/10.1016/j.neuroimage.2009.07.031

Hutchinson, S., & Louwerse, M. M. (2014). Language statistics explain the spatial–numerical association of response codes. *Psychonomic bulletin & review*, *21*(2), 470-478. https://doi.org/10.3758/s13423-013-0492-2

Inhoff, A. W., Radach, R., & Heller, D. (2000). Complex compounds in German: Interword spaces facilitate segmentation but hinder assignment of meaning. *Journal of Memory and Language*, *42*(1), 23-50. https://doi.org/10.1006/jmla.1999.2666

Isel, F., Gunter, T. C., & Friederici, A. D. (2003). Prosody-assisted head-driven access to spoken German compounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(2), 27 – 288. https://doi.org/10.1037/0278-7393.29.2.277

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434-446. https://doi.org/10.1016/j.jml.2007.11.007

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance*, *1*(2), 130-136. https://doi.org/10.1037/0096-1523.1.2.130

Jefferies, E., Patterson, K., Jones, R. W., & Lambon Ralph, M. A. (2009). Comprehension of concrete and abstract words in semantic dementia. *Neuropsychology, 23*(4), 492–499. https://doi.org/10.1037/a0015452

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. https://doi.org/10.1016/j.jml.2011.07.003

Jones, M. N., Kintsch, W., & Mewhort, D. J. (2006). High-dimensional semantic space accounts of priming. *Journal of Memory and Language*, *55*(4), 534-552. https://doi.org/10.1016/j.jml.2006.07.003

Jones, M. N., & Mewhort, D. J. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological Review*, *114*(1), 1-37.

Jones, M. N., Willits, J., & Dennis, S. (2015). Models of semantic memory. In J. R. Busemeyer (Ed.)., *The oxford handbook of computational and mathematical psychology* (pp. 232-254). Oxford University Press.

Juhasz, B. J., Inhoff, A. W., & Rayner, K. (2005). The role of interword spaces in the processing of English compound words. *Language and Cognitive Processes*, *20*(1-2), 291-316. https://doi.org/10.1080/01690960444000133

Kamp, H. & Partee, B. (1995). Prototype theory and compositionality. *Cognition*, *57*(2), 129–191. https://doi.org/10.1016/0010-0277(94)00659-9

Katz, R. B., & Goodglass, H. (1990). Deep dysphasia: Analysis of a rare form of repetition disorder. *Brain and Language*, *39*(1), 153-185. https://doi.org/10.1016/0093-934X(90)90009-6

Keene, O. N. (1995). The log transformation is special. *Statistics in Medicine*, *14*(8), 811-819. https://doi.org/10.1002/sim.4780140810.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*(3), 627-633. https://doi.org/10.3758/BRM.42.3.627

Khanna, M. M., & Cortese, M. J. (2021). How well imageability, concreteness, perceptual strength, and action strength predict recognition memory, lexical decision, and reading aloud performance. *Memory, 29*(5), 622-636. https://doi.org/10.1080/09658211.2021.1924789

Kintsch, W. (2000). Metaphor comprehension: A computational theory. *Psychonomic Bulletin & Review*, *7*(2), 257–266. https://doi.org/10.3758/BF03212981

Koester, D., Holle, H., & Gunter, T. C. (2009). Electrophysiological evidence for incremental lexical-semantic integration in auditory compound comprehension. *Neuropsychologia*, *47*(8-9), 1854-1864. https://doi.org/10.1016/j.neuropsychologia.2009.02.027

Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General*, *140*(1), 14–34. https://doi.org/10.1037/a0021446

Kroll, J. F., & Merves, J. S. (1986). Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(1), 92 –107. https://doi.org/10.1037/0278-7393.12.1.92

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. https://doi.org/10.1037/a0013484

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1-26. 10.18637/jss.v082.i13

Lakoff, G., & Johnson, M. (1980). The metaphorical structure of the human conceptual system. *Cognitive Science*, *4*(2), 195-208. https://doi.org/10.1207/s15516709cog0402_4

Lambon Ralph, M. A., Ehsan, S., Baker, G. A., & Rogers, T. T. (2012). Semantic memory is impaired in patients with unilateral anterior temporal lobe resection for temporal lobe epilepsy. *Brain, 135*(1), 242-258. https://doi.org/10.1093/brain/awr325

Landauer, T. K., & Dumais, S. (2008). Latent semantic analysis. *Scholarpedia*, *3*(11), 4356, doi:10.4249/scholarpedia.4356

Lenth, R. V. (2021). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.5.4. https://CRAN.R-project.org/package=emmeans

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

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

Libben, G. (2003). Morphological parsing and morphological structure. In E. M. H. Assink & D. Sandra (Eds.)., *Reading complex words* (pp. 221-239). Springer.

Libben, G. (2014). The nature of compounds: A psychocentric perspective. *Cognitive Neuropsychology*, *31*(1–2), 8–25. https://doi.org/10.1080/02643294.2013.874994

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. https://doi.org/10.1016/S0093-934X(02)00520-5

Libben, G., & Jarema, G. (2006). The representation and processing of compound words. *Word Knowledge and Word Usage*, 336-352. https://doi.org/10.1515/9783110440577-009

Louwerse, M. M. (2007). Symbolic or embodied representations: A case for symbol interdependency. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.)., Handbook of latent semantic analysis (pp. 107-120). Lawrence Erlbaum Associates, Inc.

Louwerse, M. M. (2011). Symbol interdependency in symbolic and embodied cognition. *Topics in Cognitive Science*, *3*(2), 273-302. https://doi.org/10.1111/j.1756-8765.2010.01106.x

Louwerse, M. M., & Jeuniaux, P. (2008). Language comprehension is both embodied and symbolic. In M. De Vega, A. M. Glenberg & A. C. Graesser (Eds.)., *Symbols, embodiment, and meaning* (pp. 309-326). Oxford University Press.

Louwerse, M. M., & Jeuniaux, P. (2010). The linguistic and embodied nature of conceptual processing. *Cognition*, *114*(1), 96-104. https://doi.org/10.1016/j.cognition.2009.09.002

Lucas, H. D., Hubbard, R. J., & Federmeier, K. D. (2017). Flexible conceptual combination: Electrophysiological correlates and consequences for associative memory: ERP correlates of conceptual combination. *Psychophysiology*, *54*(6), 833–847. https://doi.org/10.1111/psyp.12840

Lukasz, K., & Novomsetsky, F. (2015). *Moments, cumulants, skewness, kurtosis and related tests.* http://www.r-project.org, http://www.komsta.net/

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, *28*(2), 203–208. https://doi.org/10.3758/BF03204766

Lutfallah, S. & Buchanan, L. (2019). Quantifying subjective data using online Q-methodology software. *The Mental Lexicon, 14*(3), 415-423.https://doi.org/10.1075/ml.20002.lut

Lutfallah, S., Fast, C., Rangan, C., & Buchanan, L. (2018). Semantic neighbourhoods: There's an app for that. *The Mental Lexicon*, *13*(3), 388-393. https://doi.org/10.1075/ml.18015.lut

Lynott, D., & Connell, L. (2010). Embodied Conceptual Combination. *Frontiers in Psychology*, *1*(212), 1-14. https://doi.org/10.3389/fpsyg.2010.00212

MacGregor, L. J., & Shtyrov, Y. (2013). Multiple routes for compound word processing in the brain: Evidence from EEG. *Brain and Language*, *126*(2), 217-229. https://doi.org/10.1016/j.bandl.2013.04.002

Magnusdottir, S., Fillmore, P., Den Ouden, D. B., Hjaltason, H., Rorden, C., Kjartansson, O., ... & Fridriksson, J. (2013). Damage to left anterior temporal cortex predicts impairment of complex syntactic processing: A lesion-symptom mapping study. *Human Brain Mapping, 34*(10), 2715-2723. https://doi.org/10.1002/hbm.22096

Maguire, P., Devereux, B., Costello, F., & Cater, A. (2007). A reanalysis of the CARIN theory of conceptual combination. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 811-821. https://doi.org/10.1037/0278-7393.33.4.811

Maguire, P., Maguire, R., & Cater, A. W. S. (2010). The influence of interactional semantic patterns on the interpretation of noun–noun compounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(2), 288–297. https://doi.org/10.1037/a0018687

Malhi, S. K., & Buchanan, L. (2018). A test of the symbol interdependency hypothesis with both concrete and abstract stimuli. *Plos One*, *13*(3), e0192719. 10.1371/journal.pone.0192719

Malhi, S. K., McAuley, T. L., Lansue, B., & Buchanan, L. (2019). Concrete and abstract word processing in deep dyslexia. *Journal of Neurolinguistics*, *51*, 309-323. https://doi.org/10.1016/j.jneuroling.2018.11.001

Mandera, P., Keuleers, E., & Brysbaert, M. (2017). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language*, *92*, 57-78. https://doi.org/10.1016/j.jml.2016.04.001

Marelli, M., & Luzzatti, C. (2012). Frequency effects in the processing of Italian nominal compounds: Modulation of headedness and semantic transparency. *Journal of Memory and Language*, *66*(4), 644-664. https://doi.org/10.1016/j.jml.2012.01.003

McAuley, T. (2018) Examining semantic effects in conceptual combination (unpublished Master's thesis). University of Windsor, Windsor, Ontario.

Medin, D. L. (1975). A theory of context in discrimination learning. *Psychology of Learning and Motivation*, *9*, 263-314. https://doi.org/10.1016/S0079-7421(08)60273-X

Medin, D. L., & Coley, J. D. (1998). Concepts and categorization. In J. Hochberg (Ed.), *Handbook of perception and cognition (*2nd ed.). Academic Press.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*(3), 207 –238. https://doi.org/10.1037/0033-295X.85.3.207

Medin, D. L., & Shoben, E. J. (1988). Context and structure in conceptual combination. *Cognitive Psychology*, *20*(2), 158–190. https://doi.org/10.1016/0010-0285(88)90018-7

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology*, *32*(1), 89-115.

Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex*, *48*(7), 788–804. https://doi.org/10.1016/j.cortex.2010.11.002

Meteyard, L., & Davies, R. A. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language*, *112*, 104092. https://doi.org/10.1016/j.jml.2020.104092

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*(2), 227–234. https://doi.org/10.1037/h0031564

Minsky, M. (1975). A framework for representing knowledge. In P.H. Winston (Ed.), *The psychology of computer vision*. McGraw-Hill Companies.

Mitchell, J., & Lapata, M. (2010). Composition in distributional models of semantics. *Cognitive Science*, *34*(8), 1388–1429. https://doi.org/10.1111/j.1551-6709.2010.01106.x

Morrison, C. M., & Ellis, A. W. (2000). Real age of acquisition effects in word naming and lexical decision. *British Journal of Psychology, 91*(2), 167-180. https://doi.org/10.1348/000712600161763

Mullaly, A. C., Gagné, C. L., Spalding, T. L., & Marchak, K. A. (2010). Examining ambiguous adjectives in adjective-noun phrases: Evidence for representation as a shared core-meaning with sense specialization. *The Mental Lexicon, 5*(1), 87-114. https://doi.org/10.1075/ml.5.1.04mul

Murphy, G. L. (1988). Comprehending complex concepts. *Cognitive Science*, *12*(4), 529-562. https://doi.org/10.1207/s15516709cog1204_2

Murphy, G. L. (1990). Noun phrase interpretation and conceptual combination. *Journal of Memory and Language*, *29*(3), 259–288. https://doi.org/10.1016/0749-596X(90)90001-G

Murphy, G. L. (2016). Is there an exemplar theory of concepts? *Psychonomic Bulletin & Review*, *23*(4), 1035–1042. https://doi.org/10.3758/s13423-015-0834-3

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*(3), 289 –316. https://doi.org/10.1037/0033-295X.92.3.289

Nagle, C. (2019). An introduction to fitting and evaluating mixed-effects models in R. In J. Levis, C. Nagle, & E. Todey (Eds.), Proceedings of the 10th Pronunciation in Second Language Learning and Teaching Conference, ISSN 2380-9566, Ames, IA, September 2018 (pp. 82-105). Ames, IA: Iowa State University.

Nakagawa, S., Johnson, P. C., & Schielzeth, H. (2017). The coefficient of determination R 2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface*, *14*(134), 20170213. https://doi.org/10.1098/rsif.2017.0213

Nelson, D.L., Mcevoy, C.L. & Dennis, S. (2000). What is free association and what does it measure?. *Memory & Cognition, 28,* 887–899. https://doi.org/10.3758/BF03209337

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402-407. https://doi.org/10.3758/BF03195588

Oden, G. C. (1977). Integration of fuzzy logical information. *Journal of Experimental Psychology: Human Perception and Performance, 3*(4), 565–575. https://doi.org/10.1037/0096-1523.3.4.565

Osherson, D. N., & Smith, E. E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition*, *9*(1), 35-58. https://doi.org/10.1016/0010-0277(81)90013-5

Paivio, A. (1971). Imagery and language. In S. J. Segal (Ed.)., *Imagery* (pp. 7-32). Academic Press. https://doi.org/10.1016/B978-0-12-635450-8.50008-X

Parrish, A., & Pylkkänen, L. (2022). Conceptual combination in the LATL with and without syntactic composition. *Neurobiology of Language, 3*(1), 46–66. https://doi.org /10.1162/nol_a_00048

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*. 10.3758/s13428-018-01193-y

Pexman, P. M., Hargreaves, I. S., Edwards, J. D., Henry, L. C., & Goodyear, B. G. (2007). Neural correlates of concreteness in semantic categorization. *Journal of Cognitive Neuroscience*, *19*(8), 1407–1419. https://doi.org/10.1162/jocn.2007.19.8.1407

Pexman, P. M., Hargreaves, I. S., Siakaluk, P. D., Bodner, G. E., & Pope, J. (2008). There are many ways to be rich: Effects of three measures of semantic richness on visual word recognition. *Psychonomic Bulletin & Review*, *15*(1), 161-167. https://doi.org/10.3758/PBR.15.1.161

Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology*, *10*(5), 377–500. https://doi.org/10.1080/02643299308253469

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*(3, Pt.1), 353–363. https://doi.org/10.1037/h0025953

Potter, M. C., & Faulconer, B. A. (1979). Understanding noun phrases. *Journal of Verbal Learning and Verbal Behavior*, *18*(5), 509–521. https://doi.org/10.1016/S0022-5371(79)90274-3

Pylkkänen, L. (2016). Composition of complex meaning: Interdisciplinary perspectives on the left anterior temporal lobe. In G Hickock and S. Small (Eds.), *Neurobiology of language* (pp.621-631). London: Elsevier.

Quené, H., & Van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language*, *59*(4), 413-425. https://doi.org/10.1016/j.jml.2008.02.002

Raffray, C. N., Pickering, M. J., & Branigan, H. P. (2007). Priming the interpretation of noun–noun combinations. *Journal of Memory and Language*, *57*(3), 380-395. https://doi.org/10.1016/j.jml.2007.06.009

Ran, B., & Duimering, P. R. (2009). Conceptual combination: Models, theories, & controversies. In S. P. Weingarten & H. O. Penat (Eds.), *Cognitive psychology research developments* (pp. 65-90). Nova Science Publishers, Inc.

Recchia, G., & Jones, M. N. (2012). The semantic richness of abstract concepts. *Frontiers in Human Neuroscience*, *6*(315), 1-16. https://doi.org/10.3389/fnhum.2012.00315

Riordan, B., & Jones, M. N. (2010). Redundancy in perceptual and linguistic experience: Comparing feature-based and distributional models of semantic information. *Topics in Cognitive Science*, *3*, 303–345. https://doi.org/10.1111/j.1756-8765.2010.01111.x

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, *12*(1), 1-20. https://doi.org/10.1016/S0022-5371(73)80056-8

Rips, L. J., Smith, E. E., & Medin, D. L. (2012). Concepts and categories: Memory, meaning, and metaphysics. In K. Holyoak & R. Morrison (Eds.), *The oxford handbook of thinking and reasoning* (pp 177-209). OUP USA. 10.1093/oxfordhb/9780199734689.001.0001

Robinson, D., Hayes, A. & Couch, S (2022). broom: Convert Statistical Objects into Tidy Tibbles. R package version 0.7.12. https://CRAN.R-project.org/package=broom

Rodríguez-Ferreiro, J., Gennari, S. P., Davies, R., & Cuetos, F. (2011). Neural correlates of abstract verb processing. *Journal of Cognitive Neuroscience*, *23*(1), 106-118. https://doi.org/10.1162/jocn.2010.21414

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573-605. https://doi.org/10.1016/0010-0285(75)90024-9

Rumelhart, D. E. (1980). On evaluating story grammars. *Cognitive Science*, *4*(3), 313-316. 10.1207/s15516709cog0403_5

Schmidt, G. L., & Seger, C. A. (2009). Neural correlates of metaphor processing: The roles of figurativeness, familiarity and difficulty. *Brain and Cognition, 71*(3), 375-386. https://doi.org/10.1016/j.bandc.2009.06.001

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. https://doi.org/10.3758/s13423-015-0926-0

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.)., *Morphological aspects of language processing* (pp. 257-294). Lawrence Erlbaum Associates, Inc.

Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language*, *37*(1), 118-139. https://doi.org/10.1006/jmla.1997.2510

Schwanenflugel, P. J. (1991). Contextual constraint and lexical processing. In G. B. Simpson (Ed.)., *Understanding word and sentence* (pp. 23-45). Elsevier Science Publishing, Inc.

Schwanenflugel, P. J. (2013). *The psychology of word meanings*. Psychology Press.

Schwanenflugel, P. J., & Shoben, E. J. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*(1), 82–102. https://doi.org/10.1037/0278-7393.9.1.82

Shaoul, C., & Westbury, C. (2006). Word frequency effects in high-dimensional co-occurrence models: A new approach. *Behavior Research Methods, 38,* 190–195. https://doi.org/10.3758/BF03192768

Shulman, H. G., & Davidson, T. C. (1977). Control properties of semantic coding in a lexical decision task. *Journal of Verbal Learning and Verbal Behavior*, *16*(1), 91-98. https://doi.org/10.1016/S0022-5371(77)80010-8

Siakaluk, P. D., Pexman, P. M., Sears, C. R., Wilson, K., Locheed, K., & Owen, W. J. (2008). The benefits of sensorimotor knowledge: Body–object interaction facilitates semantic processing. *Cognitive Science*, *32*(3), 591-605. https://doi.org/10.1080/03640210802035399

Smith, E. E., & Osherson, D. N. (1984). Conceptual combination with prototype concepts. *Cognitive Science*, *8*(4), 337–361. https://doi.org/10.1207/s15516709cog0804_2

Smith, E. E., Osherson, D. N., Rips, L. J., & Keane, M. (1988). Combining prototypes: A selective modification model. *Cognitive Science*, *12*(4), 485-527. https://doi.org/10.1207/s15516709cog1204_1

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts* (Volume 9.). Harvard University Press.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, *81*(3), 214–241. https://doi.org/10.1037/h0036351

Spalding, T. L., & Gagné, C. L. (2008). CARIN theory reanalysis reanalyzed: A comment on Maguire, Devereux, Costello, and Cater (2007). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(6), 1573–1578. https://doi.org/10.1037/a0013120

Spalding, T. L., & Gagné, C. L. (2014). Relational diversity affects ease of processing even for opaque English compounds. *The Mental Lexicon*, *9*(1), 48-66. https://doi.org/10.1075/ml.9.1.03spa

Spalding, T. L., Gagné, C. L., Mullaly, A., & Ji, H. (2010). Relation-based interpretation of noun-noun phrases: A new theoretical approach. In S. Olsen (Ed.)., *New impulses in word formation* (pp. 283-316). Buske.

Spencer, A. (1991). *Morphological theory: An introduction to word structure in generative grammar* (Vol. 2). Oxford: Basil Blackwell.

Springer, K., & Murphy, G. L. (1992). Feature Availability in Conceptual Combination. *Psychological Science*, *3*(2), 111–117. https://doi.org/10.1111/j.1467-9280.1992.tb00008.x

Takashima, A., Bakker, I., Van Hell, J. G., Janzen, G., & McQueen, J. M. (2014). Richness of information about novel words influences how episodic and semantic memory networks interact during lexicalization. *NeuroImage, 84,* 265-278. https://doi.org/10.1016/j.neuroimage.2013.08.023

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology Section A*, *57*(4), 745-765. https://doi.org/10.1080/02724980343000477

Thagard, P. (1984). Conceptual combination and scientific discovery. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, *1984*(1), 3–12. https://doi.org/10.1086/psaprocbienmeetp.1984.1.192323

Thagard, P. (1997). *Coherent and creative conceptual combinations.* In T. B. Ward, S. M. Smith, & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes* (p. 129–141). American Psychological Association. https://doi.org/10.1037/10227-006

Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*(4), 327 –352. https://doi.org/10.1037/0033-295X.84.4.327

Vandenberghe, R., Nobre, A. C., & Price, C. J. (2002). The response of left temporal cortex to sentences. *Journal of Cognitive Neuroscience, 14*(4), 550-560. 10.1162/08989290260045800

Vecchi, E. M., Baroni, M., & Zamparelli, R. (2011, June). (Linear) maps of the impossible: capturing semantic anomalies in distributional space. In *Proceedings of the Workshop on Distributional Semantics and Compositionality* (pp. 1-9). Association for Computational Linguistics.

Vecchi, E. M., Marelli, M., Zamparelli, R., & Baroni, M. (2017). Spicy adjectives and nominal donkeys: Capturing semantic deviance using compositionality in distributional spaces. *Cognitive Science*, *41*(1), 102–136. https://doi.org/10.1111/cogs.12330

Vigliocco, G., Meteyard, L., Andrews, M., & Kousta, S. (2009). Toward a theory of semantic representation. *Language and Cognition*, *1*(2), 219-247. https://doi.org/10.1515/LANGCOG.2009.011

Wang, J., Conder, J. A., Blitzer, D. N., & Shinkareva, S. V. (2010). Neural representation of abstract and concrete concepts: A meta-analysis of neuroimaging studies. *Human Brain Mapping*, *31*(10), 1459-1468. https://doi.org/10.1002/hbm.20950

Waxman, S. R., & Markow, D. B. (1998). Object properties and object kind: Twenty-one-month-old infants' extension of novel adjectives. *Child Development*, *69*(5), 1313-1329. https://doi.org/10.1111/j.1467-8624.1998.tb06214.x

Weiskopf, D. A. (2010). Embodied cognition and linguistic comprehension. *Studies in History and Philosophy of Science Part A*, *41*(3), 294–304. https://doi.org/10.1016/j.shpsa.2010.07.005

West, W. C., & Holcomb, P. J. (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: an electrophysiological investigation. *Journal of Cognitive Neuroscience*, *12*(6), 1024-1037. https://doi.org/10.1162/08989290051137558

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

Wickham et al., (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686, https://doi.org/10.21105/joss.01686

Wiemer-Hastings, K., & Xu, X. (2005). Content differences for abstract and concrete concepts. *Cognitive Science*, *29*(5), 719-736. https://doi.org/10.1207/s15516709cog0000_33

Wilkenfeld, M. J., & Ward, T. B. (2001). Similarity and emergence in conceptual combination. *Journal of Memory and Language*, *45*(1), 21-38. https://doi.org/10.1006/jmla.2000.2772

Wisniewski, E. J. (1996). Construal and similarity in conceptual combination. *Journal of Memory and Language*, *35*(3), 434–453. https://doi.org/10.1006/jmla.1996.0024

Wisniewski, E. J. (1997). When concepts combine. *Psychonomic Bulletin & Review*, *4*(2), 167–183. https://doi.org/10.3758/BF03209392

Wisniewski, E. J. (2000). Similarity, alignment, and conceptual combination: Comment on Estes and Glucksberg. *Memory & Cognition*, *28*(1), 35–38. https://doi.org/10.3758/BF03211573

Wisniewski, E. J., & Love, B. C. (1998). Relations versus properties in conceptual combination. *Journal of Memory and Language*, *38*(2), 177–202. https://doi.org/10.1006/jmla.1997.2550

Wisniewski, E. J., & Markman, A. B. (1993). The role of structural alignment in conceptual combination. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 1083-1086). Erlbaum.

Wisniewski, E. J., & Murphy, G. L. (2005). Frequency of relation type as a determinant of conceptual combination: A reanalysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(1), 169–174. https://doi.org/10.1037/0278-7393.31.1.169

Wong-Gonzalez, D. (2018). The relationship between semantic and episodic memory: Exploring the effect of semantic neighbourhood density on episodic memory. Electronic Theses and Dissertations. 7585. https://scholar.uwindsor.ca/etd/7585

Wu, L., & Barsalou, L. W. (2009). Perceptual simulation in conceptual combination: Evidence from property generation. *Acta Psychologica*, *132*(2), 173–189. https://doi.org/10.1016/j.actpsy.2009.02.002

Yap, M. J., Tan, S. E., Pexman, P. M., & Hargreaves, I. S. (2011). Is more always better? Effects of semantic richness on lexical decision, speeded pronunciation, and semantic classification. *Psychonomic Bulletin & Review*, *18*(4), 742-750. https://doi.org/10.3758/s13423-011-0092-y

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, *8*(3), 338-353. https://doi.org/10.1016/S0019-9958(65)90241-X

Zhang, L., & Pylkkänen, L. (2015). The interplay of composition and concept specificity in the left anterior temporal lobe: An MEG study. *NeuroImage*, *111*, 228-240. https://doi.org/10.1016/j.neuroimage.2015.02.028

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

APPENDICES

Appendix A

Stimulus Set with 117 Adjective-Noun Combinations and Meaningfulness Values

Noun	Low Meaningful Adjective	*M* (*SD*)	Intermediate Meaningful Adjective	*M* (*SD*)	High Meaningful Adjective	*M* (*SD*)
boredom	illegal	-1.44 (.87)	gloomy	-.33 (.95)	sheer	1.51 (.66)
cacti	starry	-1.63 (.71)	fuzzy	.04 (1.14)	prickly	1.13 (1.12)
canvas	hasty	-1.49 (.87)	unique	.11 (.90)	blank	1.66 (.77)
chaos	mobile	-1.06 (.84)	raw	.24 (1.34)	utter	1.44 (.82)
cider	chatty	-1.78 (.55)	tangy	.72 (.88)	fruity	1.24 (.96)
confession	faded	-1.07 (.98)	modest	.21 (1.01)	sincere	1.43 (.83)
conspiracy	lyrical	-.91 (1.02)	toxic	.73 (1.16)	wacky	.83 (1.14)
crush	steep	-1.30 (.78)	desirous	.05 (1.00	secretive	.98 (.98)
delusion	starchy	-1.54 (.87)	morbid	.27 (1.18)	paranoid	1.34 (.89)
denim	brutal	-1.48 (.71)	sleek	.11 (.96)	vintage	1.02 (1.14)
dessert	graceful	-1.07 (1.08)	salty	-.10 (1.22)	tasty	1.56 (1.03)
devotion	sassy	-.98 (1.23)	affectionate	.73 (1.05)	loyal	1.51 (.75)
envelope	rowdy	-1.70 (.55)	bulky	.74 (1.08)	torn	1.27 (1.05)
fraud	corny	-1.40 (.89)	crooked	.15 (1.33)	deceptive	.95 (1.05)
hike	ripe	-1.85 (.47)	slippery	.18 (.81)	tedious	.76 (.92)
hobby	strict	-.80 (1.19)	bizarre	.65 (1.14)	enjoyable	1.38 (.80)
insect	jaded	-1.22 (.59)	aquatic	.46 (1.09)	venomous	1.44 (.94)
lava	foul	-1.24 (.85)	viscous	.34 (1.02)	fiery	1.31 (.67)
legend	tidy	-1.13 (1.02)	grim	-.02 (.92)	urban	1.54 (.78)
luggage	risky	-.89 (.82)	trendy	.18 (1.00)	hefty	.96 (.93)
nest	devious	-1.33 (.87)	shallow	.06 (.93)	cosy	.93 (1.07)
pants	cursive	-1.60 (.76)	classy	.11 (.75)	stylish	1.17 (.95)
pastry	abstract	-1.20 (.73)	crunchy	.52 (1.01)	flaky	1.42 (.99)
pebble	lethal	-.87 (.97)	stray	-.04 (.99)	shiny	.92 (.90)
picnic	drastic	-1.26 (.85)	hearty	.27 (.96)	romantic	1.39 (.74)
poster	livid	-.133 (.90)	theatrical	.46 (.96)	graphic	1.10 (.95)
realm	clumsy	-1.12 (1.00)	spacious	.42 (.94)	magical	1.45 (.75)
revenge	obtuse	-1.25 (.91)	gutsy	.02 (1.03)	gory	.98 (1.17)
robe	speedy	-1.46 (.95)	snug	.67 (.90)	silky	1.46 (.75)
rumor	spongy	-1.64 (.74)	ominous	.59 (.97)	scandalous	1.33 (1.02)
sarcasm	earthy	-1.34 (.89)	slick	.02 (.96)	witty	.94 (1.16)
skeleton	itchy	-1.31 (.90)	frail	.63 (1.06)	creepy	1.36 (.98)
soda	tense	-1.70 (.81)	dilute	-.24 (.94)	fizzy	1.73 (.81)
spree	wobbly	-.35 (.91)	costly	.66 (.96)	impulsive	1.15 (.79)
theater	waxy	-1.56 (.73)	amateur	.64 (.86)	musical	1.37 (.93)
theft	festive	-1.23 (.90)	vulgar	-.09 (1.02)	petty	1.27 (1.03)
trance	slimy	-1.50 (.88)	mute	-.34 (1.07)	hypnotic	1.24 (.97)
umbrella	clingy	-1.17 (.97)	shady	.42 (1.09)	damp	.82 (1.19)
urge	stoic	-.77 (.94)	stubborn	.37 (1.04)	bodily	.75 (1.26)

Appendix B

Novel Adjective-Noun Combinations

Adjective-noun	Adjective-noun	Adjective-noun
gassy fantasy	clumsy realm	brutal denim
wormy solo	robust chill	speedy robe
spongy rumor	faded confession	solar mold
glossy bias	mobile chaos	livid poster
starchy delusion	sassy devotion	devious nest
slimy trance	lyrical conspiracy	itchy skeleton
stocky legacy	cloudy dilemma	sleepy vaccine
illegal boredom	strict hobby	drastic picnic
bald format	stoic urge	foul lava
corny fraud	ripe hike	jaded insect
grilled accuracy	chatty cider	vocal blizzard
wobbly spree	sappy pyramid	abstract pastry
earthy sarcasm	rowdy envelope	frigid stump
stale cognition	tense soda	clingy umbrella
sparse craze	flabby barn	rubbery orchid
steep crush	hilly banner	graceful dessert
dusty headache	starry cacti	mystical stair
obtuse revenge	cursive pants	greasy raft
festive theft	beefy fountain	dense toad
tidy legend	waxy theater	risky luggage
eerie tangent	hasty canvas	lethal pebble

Appendix C

Summary of Lexical and Semantic Variables for Adjective-Noun Combinations

Adjective	Noun	Meaningful group	Combined Letter Length	Mean OF	Adjective SND	Noun SND	Concreteness	
							M	*SD*
sheer	boredom	high	12	12.29	0.49	0.59	-1.07	1.02
prickly	cacti	high	12	1.59	0.65	0.68	1.49	0.85
blank	canvas	high	6	7.67	0.38	0.53	0.68	1.10
utter	chaos	high	10	19.91	0.52	0.43	-1.26	0.91
fruity	cider	high	10	2.14	0.68	0.80	1.03	0.94
sincere	confession	high	17	15.46	0.55	0.55	-0.85	1.17
wacky	conspiracy	high	15	5.25	0.60	0.52	-1.27	0.89
secretive	crush	high	13	6.88	0.43	0.45	-1.05	0.94
paranoid	delusion	high	16	3.47	0.63	0.50	-1.22	0.93
vintage	denim	high	12	2.77	0.47	0.75	0.59	1.19
tasty	dessert	high	15	10.24	0.55	0.77	0.99	0.95
loyal	devotion	high	13	18.55	0.43	0.58	-1.18	1.13
torn	envelope	high	13	10.47	0.38	0.43	1.47	0.91
deceptive	fraud	high	14	7.20	0.56	0.61	-0.89	1.09
tedious	hike	high	12	3.68	0.42	0.73	-0.04	1.05
enjoyable	hobby	high	14	4.88	0.59	0.47	-0.30	1.14
venomous	insect	high	14	5.47	0.60	0.72	1.01	1.15
fiery	lava	high	8	11.49	0.51	0.73	0.81	1.32
urban	legend	high	11	21.22	0.50	0.43	-1.10	1.00
hefty	luggage	high	7	4.13	0.51	0.52	1.04	1.03
cosy	nest	high	8	15.03	0.39	0.67	0.33	1.11
stylish	pants	high	12	5.83	0.51	0.82	0.58	1.07
flaky	pastry	high	14	11.53	0.50	0.80	1.33	0.99
shiny	pebble	high	11	4.95	0.61	0.42	1.23	1.05
romantic	picnic	high	14	15.56	0.43	0.67	0.01	1.05
graphic	poster	high	11	4.38	0.45	0.53	0.73	1.25
magical	realm	high	12	11.53	0.67	0.53	-1.22	1.04
gory	revenge	high	11	8.68	0.52	0.46	-0.99	1.07
silky	robe	high	4	6.76	0.41	0.63	1.36	0.87
scandalous	rumor	high	15	3.28	0.53	0.54	-1.15	0.88
witty	sarcasm	high	12	5.16	0.68	0.67	-1.23	0.87
creepy	skeleton	high	8	4.05	0.59	0.62	0.34	1.07
fizzy	soda	high	9	9.00	0.80	0.61	1.29	0.95
impulsive	spree	high	14	3.17	0.66	0.53	-0.96	0.93
musical	theater	high	14	8.27	0.44	0.54	0.41	1.05
petty	theft	high	10	11.82	0.47	0.60	-0.51	1.03
hypnotic	trance	high	14	5.54	0.51	0.53	-0.90	1.13
damp	umbrella	high	13	5.54	0.43	0.42	1.21	1.05
bodily	urge	high	10	12.98	0.41	0.46	-0.73	1.12
gloomy	boredom	intermediate	13	11.07	0.59	0.59	-1.01	1.05
fuzzy	cacti	intermediate	10	1.34	0.47	0.68	1.07	1.28
unique	canvas	intermediate	12	20.84	0.38	0.53	0.17	1.08
raw	chaos	intermediate	8	17.93	0.56	0.43	-1.32	0.90
tangy	cider	intermediate	11	2.33	0.75	0.80	1.16	1.15
modest	confession	intermediate	19	9.21	0.42	0.55	-1.00	0.97
toxic	conspiracy	intermediate	15	7.99	0.65	0.52	-1.33	1.00
desirous	crush	intermediate	5	4.20	0.66	0.45	-1.07	1.06
morbid	delusion	intermediate	14	5.86	0.59	0.50	-1.27	0.99
sleek	denim	intermediate	10	2.83	0.49	0.75	0.57	1.21
salty	dessert	intermediate	12	2.79	0.63	0.77	1.12	1.01
affectionate	devotion	intermediate	20	16.00	0.60	0.58	-1.05	1.10
bulky	envelope	intermediate	8	8.84	0.47	0.43	1.37	1.09
crooked	fraud	intermediate	12	10.40	0.40	0.61	-0.81	1.10
slippery	hike	intermediate	11	4.15	0.58	0.73	0.52	1.02
bizarre	hobby	intermediate	12	5.14	0.53	0.47	-0.61	0.98

aquatic	insect	intermediate	11	5.30	0.47	0.72	1.20	1.08
viscous	lava	intermediate	11	4.46	0.53	0.73	0.83	1.20
grim	legend	intermediate	10	19.93	0.48	0.43	-1.01	0.98
trendy	luggage	intermediate	13	4.61	0.57	0.52	0.75	0.97
shallow	nest	intermediate	11	20.71	0.53	0.67	0.85	1.07
classy	pants	intermediate	11	4.69	0.47	0.82	0.65	1.07
crunchy	pastry	intermediate	13	1.82	0.71	0.80	1.60	0.74
stray	pebble	intermediate	11	5.56	0.39	0.42	0.55	1.18
hearty	picnic	intermediate	12	9.03	0.58	0.67	0.63	1.02
theatrical	poster	intermediate	16	5.86	0.54	0.53	0.61	1.25
spacious	realm	intermediate	13	9.96	0.58	0.53	-0.53	1.24
gutsy	revenge	intermediate	12	8.38	0.47	0.46	-1.19	1.02
snug	robe	intermediate	8	8.98	0.50	0.63	1.09	1.08
ominous	rumor	intermediate	12	4.89	0.54	0.54	-0.91	1.23
slick	sarcasm	intermediate	12	4.19	0.45	0.67	-1.24	0.82
frail	skeleton	intermediate	13	8.63	0.46	0.62	0.85	1.04
dilute	soda	intermediate	9	3.59	0.67	0.61	1.01	1.08
costly	spree	intermediate	11	6.39	0.46	0.53	-0.67	1.06
amateur	theater	intermediate	11	3.07	0.65	0.54	0.05	1.11
vulgar	theft	intermediate	11	12.77	0.55	0.60	-0.57	1.00
mute	trance	intermediate	10	7.48	0.41	0.53	-0.80	1.15
shady	umbrella	intermediate	13	8.59	0.47	0.42	0.88	1.11
stubborn	urge	intermediate	12	9.67	0.65	0.46	-1.15	1.11
illegal	boredom	low	14	7.83	0.52	0.59	-0.90	1.30
starry	cacti	low	12	1.80	0.45	0.68	0.93	1.19
hasty	canvas	low	11	19.52	0.40	0.53	0.41	1.26
mobile	chaos	low	11	11.41	0.64	0.43	-0.79	1.21
chatty	cider	low	5	3.99	0.47	0.80	0.45	1.31
faded	confession	low	15	19.78	0.45	0.55	-0.76	1.21
lyrical	conspiracy	low	17	6.08	0.54	0.52	-0.96	1.22
steep	crush	low	14	5.10	0.44	0.45	-0.49	1.19
starchy	delusion	low	15	2.95	0.66	0.50	-1.18	0.90
brutal	denim	low	5	0.64	0.56	0.75	0.87	1.07
graceful	dessert	low	12	2.44	0.52	0.77	0.77	1.24
sassy	devotion	low	13	10.54	0.59	0.58	-1.13	1.23
rowdy	envelope	low	12	22.93	0.48	0.43	0.66	1.18
corny	fraud	low	10	6.36	0.65	0.61	-0.85	1.32
ripe	hike	low	4	0.54	0.63	0.73	0.39	1.21
strict	hobby	low	10	6.40	0.52	0.47	0.03	1.40
jaded	insect	low	13	5.73	0.59	0.72	1.00	1.13
foul	lava	low	9	10.91	0.42	0.73	0.61	1.30
tidy	legend	low	10	12.50	0.40	0.43	-0.49	1.18
risky	luggage	low	12	4.84	0.45	0.52	0.63	1.07
devious	nest	low	11	13.86	0.67	0.67	0.00	1.35
cursive	pants	low	12	4.45	0.72	0.82	0.35	1.42
abstract	pastry	low	11	1.83	0.62	0.80	0.39	1.35
lethal	pebble	low	6	1.41	0.48	0.42	0.86	1.25
drastic	picnic	low	6	3.64	0.46	0.67	0.38	1.14
livid	poster	low	13	5.42	0.49	0.53	0.65	1.11
clumsy	realm	low	11	10.81	0.61	0.53	-1.13	1.21
obtuse	revenge	low	13	8.73	0.68	0.46	-1.31	1.06
speedy	robe	low	9	8.42	0.55	0.63	0.58	1.17
spongy	rumor	low	11	2.07	0.57	0.54	-0.92	1.08
earthy	sarcasm	low	13	3.42	0.58	0.67	-0.70	1.26
itchy	skeleton	low	14	4.57	0.56	0.62	1.11	1.19
tense	soda	low	10	4.06	0.68	0.61	0.63	1.20
wobbly	spree	low	11	1.48	0.31	0.53	-0.58	1.25
waxy	theater	low	14	16.77	0.47	0.54	0.77	1.24
festive	theft	low	12	7.44	0.68	0.60	0.03	1.30
slimy	trance	low	11	4.94	0.52	0.53	-0.45	1.27
clingy	umbrella	low	12	15.66	0.56	0.42	1.06	1.19
stoic	urge	low	9	6.88	0.60	0.46	-0.94	1.25

Appendix D

Correlations between Meaningfulness as a Continuous Variables and Other Relevant Continuous Variables (N=117)

	Meaningfulness	Concreteness	Adj SND	Noun SND	Letter Length	Mean OF
Pearson Correlation	1	.004	-.044	-.027	.142	.104
Significance		.969	.637	.772	.127	.265

Appendix E

Summary of Counts and Proportions for Different Interpretation Types for Novel Adjective-Noun Pairs in Experiment 5

Adjective-Noun	Adj.Concrete	N.Concrete	Adj.SND	N.SND	% Unknown	% Misc	Total Interpretations (#)	%Unique	%Slot-Filling	%Noun Elaboration	%Abstraction	%Adjective-Reversal
robust chill	abstract	abstract	sparse	sparse	0.03	0.05	55	0.38	0.73	0.16	0.11	0.00
sparse craze	abstract	abstract	sparse	sparse	0.17	0.05	47	0.47	0.57	0.21	0.21	0.00
gassy fantasy	abstract	abstract	sparse	dense	0.10	0.10	48	0.65	0.44	0.29	0.25	0.02
illegal boredom	abstract	abstract	sparse	dense	0.08	0.08	50	0.52	0.12	0.60	0.24	0.04
strict hobby	abstract	abstract	sparse	sparse	0.08	0.02	54	0.48	0.61	0.30	0.09	0.00
lyrical conspiracy	abstract	abstract	dense	sparse	0.08	0.07	51	0.63	0.31	0.24	0.41	0.04
earthy sarcasm	abstract	abstract	dense	dense	0.08	0.05	52	0.60	0.52	0.25	0.23	0.00
sassy devotion	abstract	abstract	dense	dense	0.08	0.05	52	0.56	0.52	0.35	0.10	0.04
stoic urge	abstract	abstract	dense	sparse	0.12	0.03	51	0.67	0.43	0.41	0.14	0.02
eerie tangent	abstract	abstract	dense	dense	0.13	0.02	51	0.73	0.29	0.20	0.51	0.00
clumsy realm	abstract	abstract	dense	sparse	0.13	0.03	50	0.86	0.26	0.42	0.30	0.02
corny fraud	abstract	abstract	dense	dense	0.08	0.07	51	0.45	0.43	0.41	0.16	0.00
obtuse revenge	abstract	abstract	dense	sparse	0.05	0.03	55	0.47	0.60	0.27	0.13	0.00
festive theft	abstract	abstract	dense	dense	0.05	0.02	56	0.39	0.61	0.21	0.18	0.00
hasty canvas	abstract	concrete	sparse	sparse	0.05	0.10	51	0.27	0.80	0.12	0.08	0.00
mystical stair	abstract	concrete	sparse	dense	0.07	0.15	47	0.32	0.83	0.04	0.13	0.00
foul lava	abstract	concrete	sparse	dense	0.12	0.07	49	0.49	0.67	0.14	0.18	0.00
risky luggage	abstract	concrete	sparse	sparse	0.07	0.05	52	0.37	0.71	0.23	0.06	0.00
drastic picnic	abstract	concrete	sparse	dense	0.07	0.08	51	0.41	0.76	0.24	0.00	0.00
chatty cider	abstract	concrete	sparse	dense	0.03	0.07	54	0.28	0.81	0.07	0.11	0.00
rowdy envelope	abstract	concrete	sparse	sparse	0.07	0.07	52	0.38	0.81	0.06	0.13	0.00
livid poster	abstract	concrete	sparse	sparse	0.07	0.05	53	0.28	0.87	0.02	0.09	0.02
graceful dessert	abstract	concrete	sparse	dense	0.03	0.05	55	0.51	0.75	0.22	0.04	0.00
speedy robe	abstract	concrete	dense	dense	0.13	0.05	49	0.41	0.59	0.10	0.31	0.00

clingy umbrella	abstract	concrete	dense	sparse	0.12	0.13	45	0.67	0.53	0.22	0.24	0.00
brutal denim	abstract	concrete	dense	dense	0.03	0.08	53	0.42	0.70	0.26	0.04	0.00
sleepy vaccine	abstract	concrete	dense	dense	0.03	0.02	57	0.28	0.88	0.09	0.02	0.02
jaded insect	abstract	concrete	dense	dense	0.03	0.08	52	0.60	0.67	0.23	0.12	0.00
vocal blizzard	abstract	concrete	dense	sparse	0.05	0.03	55	0.42	0.42	0.00	0.38	0.20
abstract pastry	abstract	concrete	dense	dense	0.12	0.05	50	0.58	0.58	0.34	0.06	0.02
devious nest	abstract	concrete	dense	dense	0.08	0.07	51	0.55	0.45	0.16	0.39	0.00
tense soda	abstract	concrete	dense	dense	0.05	0.12	50	0.40	0.76	0.12	0.10	0.02
wobbly spree	concrete	abstract	sparse	sparse	0.15	0.02	50	0.76	0.32	0.20	0.48	0.00
tidy legend	concrete	abstract	sparse	sparse	0.08	0.07	51	0.53	0.57	0.25	0.18	0.00
faded confession	concrete	abstract	sparse	dense	0.02	0.08	54	0.52	0.50	0.39	0.11	0.00
stale cognition	concrete	abstract	sparse	dense	0.18	0.05	46	0.50	0.46	0.20	0.35	0.00
slimy trance	concrete	abstract	sparse	sparse	0.12	0.12	46	0.78	0.43	0.09	0.46	0.02
bald format	concrete	abstract	sparse	sparse	0.08	0.07	51	0.33	0.67	0.10	0.10	0.08
spongy rumor	concrete	abstract	dense	dense	0.02	0.15	50	0.26	0.74	0.26	0.00	0.00
cloudy dilemma	concrete	abstract	dense	sparse	0.00	0.00	60	0.48	0.58	0.22	0.18	0.02
stocky legacy	concrete	abstract	dense	sparse	0.13	0.08	47	0.55	0.64	0.15	0.21	0.00
glossy bias	concrete	abstract	dense	sparse	0.12	0.07	49	0.59	0.27	0.45	0.29	0.00
starchy delusion	concrete	abstract	dense	sparse	0.13	0.07	48	0.71	0.33	0.15	0.52	0.02
wormy solo	concrete	abstract	dense	sparse	0.17	0.03	48	0.50	0.50	0.13	0.25	0.13
grilled accuracy	concrete	abstract	dense	sparse	0.08	0.05	52	0.44	0.27	0.15	0.46	0.12
hilly banner	concrete	concrete	sparse	sparse	0.17	0.07	46	0.48	0.72	0.20	0.02	0.07
steep crush	concrete	concrete	sparse	sparse	0.05	0.07	53	0.42	0.68	0.13	0.19	0.00
beefy fountain	concrete	concrete	sparse	sparse	0.08	0.05	52	0.37	0.73	0.10	0.15	0.02
starry cacti	concrete	concrete	sparse	dense	0.12	0.15	44	0.57	0.70	0.20	0.09	0.00
greasy raft	concrete	concrete	sparse	dense	0.07	0.08	51	0.37	0.69	0.12	0.16	0.02
waxy theater	concrete	concrete	sparse	dense	0.12	0.08	48	0.52	0.60	0.27	0.08	0.04
flabby barn	concrete	concrete	sparse	sparse	0.15	0.07	47	0.47	0.70	0.15	0.15	0.00
lethal pebble	concrete	concrete	sparse	sparse	0.07	0.05	53	0.38	0.68	0.19	0.13	0.00
frigid stump	concrete	concrete	sparse	sparse	0.20	0.00	48	0.60	0.58	0.13	0.29	0.00
solar mold	concrete	concrete	sparse	sparse	0.13	0.10	46	0.41	0.74	0.00	0.17	0.09

dense toad	concrete	concrete	sparse	sparse	0.13	0.03	50	0.36	0.76	0.06	0.18	0.00
sappy pyramid	concrete	concrete	sparse	sparse	0.30	0.08	37	0.68	0.49	0.24	0.27	0.00
itchy skeleton	concrete	concrete	dense	dense	0.17	0.05	47	0.49	0.26	0.13	0.38	0.23
dusty headache	concrete	concrete	dense	dense	0.03	0.03	56	0.54	0.57	0.30	0.13	0.00
ripe hike	concrete	concrete	dense	dense	0.08	0.03	53	0.53	0.62	0.26	0.11	0.00
rubbery orchid	concrete	concrete	dense	dense	0.10	0.05	51	0.37	0.82	0.10	0.08	0.00
cursive pants	concrete	concrete	dense	dense	0.18	0.07	45	0.47	0.76	0.18	0.04	0.02

VITA AUCTORIS

NAME:	Tara McAuley
PLACE OF BIRTH:	Windsor, ON
YEAR OF BIRTH:	1993
EDUCATION:	St. Joseph's High School, Windsor, ON, 2011
	University of Windsor, B.Sc., Windsor, ON, 2015
	University of Windsor, M.A., Windsor, ON, 2018