

Suppression and Enhancement in Metaphor Processing

Fartein Nordengen Tjessem

English Language and Linguistics
60

The Department of Literature, Area Studies and European Languages

Faculty of Humanities



Acknowledgments

Writing a master thesis completely from home, without fellow students to confide in, ask for help, or discuss problems with, has been challenging. I therefore wish to thank all my friends and family members who have provided me with someone to discuss my ideas and issues with and provided me with support when I have most needed it.

But most importantly, I would like to thank my supervisor, Dr. Camilo Rodriguez Ronderos. You have provided invaluable help in the areas of this thesis I struggled to grasp and have provided invaluable optimism and enthusiasm when my own dwindled in worries, frustration, and anxiousness.

Also, special thanks to my girlfriend, Rebecca. Not only for being as helpful as she has been, acting at times as best she could as a fellow student, giving me fruitful discussion on those topics and texts I either struggled to comprehend, or struggled to explain. But also, and most importantly, I must thank her for somehow dealing with the absolute insanity that is my anxiety levels before important deadlines.

I would also like to thank my best friend, Ruben. The joy he exudes when listening to anyone discuss the topics and subjects they are enthusiastic about is admirable, wonderful, and infectious.

I would also like to thank my puppy, Sonny. Although getting a puppy whilst writing a master thesis might have been the dumbest thing I have ever done, he has gotten me out of my own head, out of the apartment, and been an endless source of dopamine.

And finally, I would like to add one last silly little thanks to all the comedians and comedy shows that have provided me with some much-needed relief throughout the year, whilst still allowing me to pretend like I am consuming relevant pragmatic content.

Can I first of all say my memories of this are a bit sketchy. ... they're more like, just, fingerprints on an abandoned handrail.

Bob Mortimer

Abstract

Both Relevance Theory and the Class Inclusion theory of metaphor view metaphor processing as dependent on cognitive processes that create ad hoc categories. These ad hoc categories are created through lexical modulation of the metaphor vehicle, i.e., the word used in to convey something metaphorically. This modulation is hypothesised to use general cognitive mechanisms which suppresses contextually irrelevant features and enhances contextually relevant features of the vehicle. The main goals of this thesis are to explain and discuss these theories, as well as test this modulation through a replication of the experiment on suppression and enhancement by Rubio-Fernandez (2007). Because the replication is the most important part of the thesis, the thesis also includes a brief explanation and discussion of the replication crisis; an important discovery that many empirical experiments are not able to be replicated and achieve congruent results. Some of the important measures recommended to deal with the replication crisis include conducting replications of experiments with new material and new participants. The results of the experiment conducted as part of this thesis support the results found in Rubio-Fernandez (2007) and therefore also the hypothesis that contextually irrelevant features of metaphor vehicles are suppressed, and contextually relevant features are enhanced.

Contents

Acknowledgments	I
Abstract	II
1. Introduction	1
1.1 General Introduction	1
1.2 Aim of Study	4
1.3 Outline of Thesis	5
2. Literature Review	6
2.1 Psychological Investigations of Language and Cognition	6
2.2 Philosophical Theories on Literal and Figurative Language	8
2.2.1 Gricean Theories on Language and Communication	9
2.2.2 Searle's Explicit Predictions	13
2.3 Contemporary Theories on Metaphor	19
2.3.1 Metaphors as Comparison: The Analogical Approach	20
2.3.2 Lexical Modulation Theories of Metaphor	21
2.3.2.1 Relevance Theory	21
2.3.2.2 Class Inclusion Theory	32
2.4 Defining Key Experimental Concepts	41
2.4.1 Processing Factors	41
2.4.2 Concepts, Properties and Ad Hoc Categories	43
2.4.3 Suppression and Enhancement	45
2.4.4 Priming	57
2.5 The Replication Crisis	59
3. Experiment	64
3.1 Introduction	64
3.2 Method	64
3.2.1 Participants.	65
3.2.2 Design	65
3.2.3 Materials	67
3.2.3 Apparatus	71
3.2.4 Procedure.	71
3.3 Results	72
3.3.1 Results of Pilot Run	72
3.3.2 Results of Main Experiment	73
3.4 Discussion	76
4. General Discussion	79
5. Conclusion	81

References	82
<i>Appendix A: Primes and Features</i>	90
<i>Appendix B: Metaphoric contexts.....</i>	91

1. Introduction

1.1 General Introduction

The jigsaw puzzle that is *language* has many pieces. Some pieces have similar colours and patterns, and clearly go together, or at the very least, near each other. Some of these similarities make the puzzle easier; it becomes obvious which pieces go where. Other similarities make it more difficult; the pieces blend together. The similarities make it more difficult to figure out which pieces actually do join together. To complete the picture is a difficult task, and although we may have an inkling of what image the puzzle displays, actually placing that last piece that completes the puzzle is far away, if it is even possible. The puzzle piece this thesis attempts to help identify and position is the piece of metaphoric language comprehension.

In overly simplistic layman's terms, a *metaphor* is the description of something as something else that it is literally not. The type of metaphor focused on in this thesis is *nominal metaphors*. Nominal metaphors are metaphors where a noun is used as the metaphoric *vehicle*, meaning the word used in a metaphoric sense, such as:

1.1 Our boss is a *vulture*.

The noun *vulture* is used as a metaphoric vehicle to assert some features, properties, or behaviour of the metaphoric target *Our boss*. Presumably, the referent of the word *boss* is most likely not literally a scavenging bird of prey. However, the description of a human holding a certain amount of power over others as a scavenger may provide some information that most typical speakers would comprehend with relative ease. This is typically the case even if the metaphoric statement is not literally true, and even if the *exact* intended meaning of the statement might not be comprehended identically by each listener.

Metaphors are a common feature of human language and can be used without the awareness of listeners, and potentially even the speaker (Holyoak & Stamenković, 2018; Glucksberg, 2001). However, metaphors have also posed issues throughout the study of language and communication. Previously the common understanding of communication was that of the Shannon-Weaver model, a code model of communication (Noveck, 2018, p. 17).

The code model assumed that an utterance was a message transferred from the speaker, received, and then decoded by listeners. If the coded message is not deformed, the exact message intended to be communicated will be understood exactly as it is intended by the listeners. If the comprehension process consisted of a simple decoding of semantic meaning, the metaphoric use of words such as *vulture* should cause problems, seeing as the meaning cannot be the literal, lexical, semantic meaning of *vulture* and be applicable for a human. Although there are still significant disagreements of the nature of metaphors, the focus of metaphoric investigation has changed throughout time.

Metaphors, and the nature of metaphors, have been discussed since ancient Greece. Aristotle argued: ‘But the greatest thing by far is to have a command of metaphor.’ (Aristotle, 1999), due to the *rhetorical* power of metaphor. The view of metaphor in Ancient Greece was that metaphor is, in essence, the same as similes, i.e., they were viewed as implicit comparisons (Allott, 2010, p. 122). To utter (1.2)

1.2 *Frank is an avalanche*

and to utter (1.3):

1.3 *Frank is like an avalanche*

were then regarded as essentially identical. It should be noted however that Aristotle mainly discussed metaphor as a tool for rhetoric purposes. In modern scientific investigations, metaphors have been an interesting tool for the investigation of language, meaning, the mind, contextual effects, and language comprehension.

Metaphors, and figurative language in general, are interesting for several academic fields. Philosophers are interested in the relationship between language, the world, and meaning. This includes figurative language and its use and implications. Cognitive scientists on the other hand are interested in how language and communication are processed by the mind. The ways the brain comprehends, and processes figurative language is an essential aspect of communication, especially considering how ingrained figurative and metaphorical language use is in our daily communication (Lakoff & Johnson, 1980; Wilson, 2011). Finally, alongside philosophers and cognitive scientists, linguists naturally want to unravel the nature of all complexities of human language, including metaphoric meaning and how context affects meaning.

In language and communication there is a distinction between what is *said* and what is *meant* (Grice, 1975). Figurative language very clearly portrays this division. Although there are many types of figurative language, two of the most typical types of figurative language are metaphoric and ironic language use, who both portray this divide quite clearly. Metaphoric language use is as previously mentioned in simplistic reductive terms the description of something as something it is literally not. To state (1.4)

1.4 *Frank is an avalanche.*

when discussing one's disorganised, clumsy, impulsive, maybe unfortunate, or unlucky friend, indicates that we may not always speak literally, seeing as a human being cannot literally be a mass of snow falling down a mountainside. Therefore, there is a divide between the words we use, and what we mean.

Although metaphors very clearly display the disconnect between what is literally stated and what is actually meant, irony displays this disconnect maybe even more succinctly. Irony is (also too simplistically) understood as a speaker intending to convey the 'opposite' of what is literally stated. The fact that a simple statement such as

1.5 a) *That went well.*

can be used to mean the supposedly 'exact' opposite:

1.5 b) *That went terribly.*

means that language, and meaning, is flexible. Both irony and metaphors display an apparent disconnect between what a speaker literally stated and what he or she actually intended to convey.

When we examine language, it becomes clear that many utterances can have several meanings depending on context and intention. To use a common example of a metaphor, the utterance 'Frank is a shark' can mean that the presumably aforementioned *Frank* is a named member of the suborder Squali, under the Plagiostomi order of fish, observable in a local aquarium. However, it could also be used, for example, to mean that *Frank* is a cold, merciless, brutal, predatory human being. The same utterance gives wildly different meanings, based on the context it is uttered in. With the former meaning, the utterance is literal, the latter meaning however, is figurative, a metaphor. Therefore, it becomes clear that the study of semantics, the lexical meaning of words, is not sufficient to explain all meanings that speakers are able to express in human communication. If the meaning of the word *shark*

is lexically used to describe a member of the Plagiostomi family, how do listeners comprehend *shark* to mean anything other than that when used to describe a human?

1.2 Aim of Study

So how do listeners comprehend novel metaphors in communication? How does the mind process metaphors, and comprehend the non-literal meanings intended? Does the mind process metaphors differently than it does literal language, and if so, is it more costly for the mind to compute metaphorical language? Does the mind suppress irrelevant literal semantic features of metaphors whilst enhancing relevant figurative features? This thesis includes a review of some of the most prominent contemporary answers to these questions, alongside their empirical investigations. The paramount aim of this thesis is an investigation of metaphoric processing, testing the empirical predictions of two important cognitive theories, the pragmatic Relevance Theory and the psycholinguistic Class Inclusion theory of metaphor. Both theories predict that metaphors are processed through the suppression of irrelevant literal features, and the enhancement of relevant figurative features. To attempt to answer this research question, the thesis includes a replication of a previous experiment, Rubio-Fernandez (2007) that found supporting results for these theories but had an insufficient sample size for generalising the results.

Because the experiment conducted as part of this thesis is a replication, the thesis also includes a brief introduction and discussion of the ‘replication crisis’. The replication crisis is a scientific crisis in the academic community caused by the lack of replications of experiments, and the revelation that when experiments *are* replicated a significant number of replications do not lead to the same results as the original experiments (Ioannidis, 2005). The replication crisis has revealed fundamental issues with scientific research, and an objective need for improvements in the scientific community. Ioannidis (2005) is one of the most prominent articles on the replication crisis and it argues that a majority of research findings are false. The replication crisis likely affects all scientific fields (Ioannidis, 2005; Schooler, 2014), and linguistics is no exception (Grieve, 2021). Regardless of the replication crisis, experimental linguistics and pragmatics struggles to answer fundamental questions without there being several studies undermining the research and theories. Of course, this is how all research functions; new studies and new research leads to more information, and more nuanced answers. However, new research is often based on the evidence provided by older research, and the tendency for studies on metaphor comprehension to undermine each other’s

finding leads to uncertainty of the scientific validity of the research methods utilised in many experiments. The significant amount of disagreement prevalent in pragmatics, as in all experimental fields, coupled with the problematic revelations caused by the replication crisis strengthens the need and necessity for solid research and replications.

The research questions of this thesis can therefore be summarised as:

1. Does metaphor processing include the suppression of irrelevant literal features, and the enhancement of relevant figurative features?
2. Can the results of Rubio-Fernandez (2007) be replicated?
3. How can the empirical investigation of metaphors use the shortcomings revealed by the replication crisis to improve its approach to scientific research?
4. How do the different cognitive theories differ in their treatment and understanding of metaphors?

1.3 Outline of Thesis

This thesis consists of 4 main sections. The first section is a Literature Review. This section briefly presents how the scientific field of psychology reached a point where it was both willing, and able to, investigate and empirically test cognitive processes. Previously, cognitive processes were something psychology had been either unwilling to investigate or viewed itself as incapable of doing. The first section also includes the seminal philosophical theories that have been crucial to the development of contemporary pragmatic and psycholinguistic theories on metaphors and figurative language. The Literature Review also discusses these contemporary theories of metaphor. Furthermore, The Literature Review presents and discusses some key experimental concepts, relevant for the experiment that has been conducted as part of this thesis, which also includes a discussion of the original experiment, Rubio-Fernandez (2007). Finally, the Literature Review also contains a brief presentation and discussion of the recent replication crisis affecting experimental research. The third section presents the experiment conducted as part of this thesis. The fourth and fifth sections consist of a general discussion and the conclusion of the thesis.

2. Literature Review

Language and meaning have been pondered and studied for centuries, even millennia. As mentioned, metaphor alone has been discussed in depth over 2000 years ago by Aristotle. The philosophical, psychological, and linguistic discussion of language, meaning, and metaphors have developed and diverged greatly since then. To explain the contemporary ideas on metaphor *comprehension* it is near impossible not to discuss the essential works those ideas and theories sprung from. The first three sections in this literature review explain (1) how the field of psychology reached a point where it was both willing, and capable of, dealing with cognitive processes. And (2), it also explains some of the first theories from the philosophy of language, that psychologists examining language and cognition were interested in empirically testing. The next two sections highlight the contemporary theories that were created based on empirical investigations of language and communication theories, and how they furthered experimental work on metaphors. These two sections are divided into one theory that views metaphors as dependent on comparison processes, and the second into two theories that view metaphor processing as dependent on lexical modulation. Finally, the last section explains the replication crisis, and the proposed measures meant to curb the effects of the experimental shortcomings that have resulted in the replication crisis.

2.1 Psychological Investigations of Language and Cognition

The theoretical explorations of ‘meaning’ by philosophers have given us a vast amount of relevant knowledge to draw from for the understanding of metaphors. However, a crucial component of understanding metaphor *processing* is found in the work by experimental psychologists and psycholinguists working on language processing and comprehension. Rieber and Vetter (1980) provide an extensive overview of the historical roots of psycholinguistics, which include a vast number of philosophical theories, movements, and thinkers that have influenced how the modern field of psychology has reached its approach to the study of language and cognitive processing. In psychology, two important movements have been the essential approaches to experimental work, and the psychology of language: *Behaviourism*, and *Cognitivism*.

Behaviourism started with a revolutionary tear from the previous paradigm of Structuralism, caused by psychologist John R. Watson. In 1913 a scathing article written by

Watson was published in *Psychological Review*, called *Psychology as the Behaviorist Views it*. The article caused an upheaval, and a total change in the then modern approach to psychological research. The behaviourist approach was clear: it was an objective study of psychology, and more specifically, 'behaviour'. Behaviour was seen as the only directly observable aspect of the psychology of humans, and therefore the only one worth pursuing. Watson's behaviourism discarded all traces of introspection and all traces of mental philosophy, deemed as meaningless sources of subjective science. Behaviourism therefore focused on observable behaviour that could be explained by objective facts, utilising stimuli, and response. Watson's behaviourism was heavily influenced by the field of Animal Psychology. One famous example of the behaviourist approach was done by psychologist Ivan Pavlov, who observed and described dogs' response: salivation, caused by stimulus, the approach of food handlers, and many more stimuli (Schultz, 1975). According to Mandler (2002), Watson's behaviourism '*... avoided sophisticated investigation of human problem solving, memory, and language*' (p. 340). Mandler (2002) argues therefore that behaviourism's failures were caused by its lack of satisfactory description and investigation into human action and thought.

Although several researchers and experimentalists began to be dissatisfied with behaviourism's limited view of experimental possibilities, two crucial 'revolutionaries' were psychologist George Miller, with his article 'The Magical Number Seven, Plus or Minus Two' (1956) and linguist Noam Chomsky, who published two key works that challenged behaviourism: 'Syntactic Structures' (1957) and 'Review of B. F. Skinner's Verbal Behavior' (1959). Chomsky (1959) argued that behaviourism was not properly suited to deal with human language acquisition. In its pursuit of objectivity, behaviourism has dismissed mental states and cognitive processes since they are not possible to be directly observed and described. *Cognitivism*, the school of psychology that the Cognitive revolution caused, focused on the areas that behaviourism had neglected: cognitive processes, perception, thinking, memory, language and more. Therefore, although the Cognitive Revolution was not a quick revolutionary upheaval, it can be called revolutionary simply by being a total transformation of how the psychology of language was approached.

Cognitivism utilises empirical investigations to make inferences based on the computational input and output observed and tested in controlled conditions (Brown, 2007). During the period when cognitivism was becoming established as a serious psychological movement a new interdisciplinary field was emerging, joining psychology and linguistics into

psycholinguistics, in an effort to tackle the psychological processes part of language production and comprehension.

The academic interest in language and all its intricacies is vast, various, and unwavering. The focus and hard work of several academic fields has been an endless well of insights for modern pragmatists. The theoretical insights of philosophers, the knowledge of linguists, and the experiments and empiric research of psychologists, and much more has led us to an interdisciplinary field that has vast scientific sources to employ and utilise, as was proposed by the philosopher of language J. L. Austin (1979):

In the history of human inquiry, philosophy has the place of the initial central sun, seminal and tumultuous; from time to time it throws off some portion of itself to take station as a science, a planet, cool and well regulated, progressing steadily towards a distant final state ... Is it not possible that the next century may see the birth, through the joint efforts of philosophers, grammarians, and numerous other students of language, of a true and comprehensive science of language?

(p. 232)

A truly comprehensive science of language must be able to explain how figurative language is comprehended. The tear from behaviourism and the establishment of cognitivism and psycholinguistics lead to researchers and experimentalists being ready and willing to investigate theories on language, cognition, and communication. For cognitive researchers, one important and extensive source for theoretical frameworks to investigate was the Philosophy of Language.

2.2 Philosophical Theories on Literal and Figurative Language

Within Philosophy of Language, the ‘Ordinary Language approach’ established that a strictly logical, semantic approach to language, detached from contextual effects, was not sufficient for the understanding and investigation of the language utilised by real participants of everyday communication (Chapman, 2010). These insufficiencies become even more apparent when attempting to explain figurative language. One of the key Ordinary Language philosophers whose work inspired the development of the linguistic field of Pragmatics was the philosopher, Paul Grice.

2.2.1 Gricean Theories on Language and Communication

The seminal work *Logic and Conversation* by Grice was published in 1975. In it, Grice introduced several crucial concepts for the subsequent academic field of pragmatics, including his explanation for how listeners comprehend commonplace communication, which includes metaphoric utterances such as *Frank is an avalanche*. These crucial concepts include *implicatures*, the *Cooperative Principle* and the *Maxims of Conversation*. Together, these concepts explain how, in Grice's view, listeners understand communication, including metaphors, which was initially adopted and developed by subsequent pragmatic theories.

Implicature is a term coined by Grice (1975), as a noun for the meaning that is conveyed beyond what is literally stated in an utterance. Grice divided implicatures into two groups, *conventional* and *conversational* implicatures. *Conventional implicatures* are meanings caused by specific lexical words, such as *but*. An example of a conventional implicature would be *He is rich but nice*, where the word *but* conveys contrast; the person is nice *despite* being rich. *Conversational implicatures* are the relevant form of implicature for metaphoric language use, and are caused by what is said, how it is said, and the context of the utterance. Conversational implicatures are frequently used, in many different types of language, such as indirect answers and metaphors. An example of conversational implicature is the meaning intended by the indirect answer from Speaker Y:

Speaker X: *Do you want to join the football game later?*

Speaker Y: *I ran 5 kilometres this morning.*

In this conversation speaker Y does not literally answer the question posed by speaker X. However, participants in this conversation would likely understand that speaker Y *implicates* that he or she is too tired to play football, because he or she has already had a run that day. Therefore, the conversational implicature in Speaker Y's statement could for example be: *I would like to, but I am too tired to run any more today, and must therefore decline* or, in short: *No*. Metaphors are also dependent on implicatures, the intended meaning beyond what has actually been said. One common example is the metaphoric use of the word *princess*:

Speaker X: *Would you ask Mary to help clean up after dinner?*

Speaker Y: *Mary is a princess.*

The implicature here being that Speaker Y conveys that there is no use asking Mary to help, because Mary is spoiled and unhelpful, but not actually necessarily literally a female royal.

The meanings of implicatures are often semantically disconnected from the literal words used, so how can listeners reach the intended meaning, the implicatures of utterances?

According to Grice's (1975) theory of conversations, participants utilise the maxims of conversation, guided by the 'Cooperative Principle' to infer the implicature of Speaker Y's statement.

The Cooperative Principle is as Grice states:

Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.

(Grice, 1975, p. 45)

In short, the Cooperative Principle states that communicators should be cooperative in conversations, and we, as conversational participants, assume that other participants adhere to this principle. Grice's *Maxims of Conversation* are conversational principles that speakers *tend* to follow when communicating, i.e., the maxims convey the different ways in which we attempt to cooperate as communicators. The four maxims are the maxim of Quantity, Quality, Relation and Manner. The maxims are in Grice's words:

Quantity

- 1. Make your contribution as informative as is required (for the current purposes of the exchange).*
- 2. Do not make your contribution more informative than is required.*

Quality

Supermaxim: Try to make your contribution one that is true.

- 1. Do not say what you believe to be false.*
- 2. Do not say that for which you lack adequate evidence.*

Relation

- 1. Be relevant*

Manner

Supermaxim: Be perspicuous.

1. *Avoid obscurity of expression.*
2. *Avoid ambiguity.*
3. *Be brief (avoid unnecessary prolixity).*
4. *Be orderly.*

(Grice, 1975)

For figurative language, the relevant maxim is, according to Grice (1975), the maxim of Quality. As mentioned, the maxims are principles or rules that speakers *tend* to follow, but they may easily be broken. If they are covertly violated, a speaker is for example lying. If they are *overtly* violated, or *flouted*, the speaker breaks the rule in such a way that the listener is meant to understand that the maxim is *purposefully* being broken. In Grice's view, metaphors are 'floutings' of the maxim of Quality. As Grice puts it, metaphors:

... involve categorically falsity, so the contradictory of what the speaker has made as it to say will, strictly speaking, be a truism; so it cannot be THAT that such a speaker is trying to get across. The most likely supposition is that the speaker is attributing to his audience some feature or features in respect of which the audience resembles...

(Grice, 1975, p. 53)

Although Grice discussed metaphor rather briefly, his propositions are based on a 'literal-first' idea, meaning that the literal meaning of an utterance is initially derived, then understood as inaccurate. When the Maxim of Quality has been flouted, alternative interpretations must be derived. Because a failure in communication has occurred and been recognized, the listener begins to find other viable, non-literal meanings. Grice's argument for metaphor comprehension has been structured into what is now called the Standard Pragmatic Model:

- I. *Derive the literal reading.*
- II. *Recognize that the literal reading cannot be true.*
- III. *Search for a non-literal reading that makes sense.*

(Noveck, 2018, p. 69)

Essentially, a metaphor such as *Frank is an avalanche* will be initially processed to mean that *Frank is a mass of snow falling down a mountainside*. When this metaphor is heard, the listener will recognize that the human being *Frank* cannot literally be a mass of snow and will then search for valid non-literal interpretations of the word *avalanche*. Grice only briefly discussed metaphors, and it is not difficult to spot insufficiencies in his proposals on metaphor. A poetic metaphor that is commonly used to undermine Grice's thoughts on metaphors is the quote '*No man is an island*' by poet John Donne (1624). This metaphor is clearly a truism; no man is literally an island, yet as a metaphor it is 'easily' understood. More everyday language also displays this principle, when speakers use other negated metaphors, such as *She is no angel*. These examples include no 'categorical falsity', no flouting of the maxim of quality, and the Standard Pragmatic Model would be insufficient as an explanation for how the statement would be comprehended. There is also a problematic relationship between the maxim of quality and Grice's distinction of what a speaker 'said' and what a speaker 'made as if to say'. Grice claims that what is 'said' is the explicit content of an utterance, as opposed to what is *implicated*. According to Grice, when a speaker utilises figurative language, said speaker only 'makes as if to say', not literally saying. Therefore, since nothing has been 'said', then the maxim of quality has, technically, not been broken (Allott, 2010).

The Standard Pragmatic model has inspired a multitude of pragmatic studies. Although experimental work is seldom easy, the Standard Pragmatic model is relatively simple, and therefore provides clear starting points for experimental work. Noveck (2018) argues that there are two main ways of testing it. One way is that figurative language must take longer to process than literal language, as literal language does not require the processing of steps II and III. The other aspect of the model which has a 'clear' route to be tested is that processing input literally is obligatory, seeing as it is the first step. The Standard Pragmatic Model has been debunked several times over (Glucksberg, 2003; Noveck, 2018). Glucksberg (2003) argues that the use and processing of idioms such as *to kick the bucket* disprove the literal first priority that the Standard Pragmatic Model assumes. People do not need to

consider the literal meaning of idioms when they encounter them. This is further displayed by the existence of a street joke that exploits the listeners lack of literal first processing:

2.1 -*What did the old man say before he kicked the bucket?*

 -*Look how far I can kick this bucket.*

The intended humoristic ‘surprise’ of the punchline only functions because the figurative meaning is assumed to be not only the first meaning to be processed, but also the only one to be initially accessed.

Regardless of the shortcomings of Grice’s theories on metaphor, Grice is responsible for essential contributions to the scientific investigation of language, meaning and metaphors. Furthermore, Noveck (2018) argues that the falsification of the Standard Pragmatic Model does not equal falsification of Gricean theory. Noveck (2018) argues that the three levels of theoretical analysis provided by Marr (1982) display how Gricean theory and the Standard Pragmatic Model are not on the same level. The *computational level* is the most abstract level. This level makes explicit the input and output of a process, and is the level Grice operates on, as a philosopher and theoretician. According to Noveck (2018) the Standard Pragmatic Model ‘mimics’ Grice’s ideas, but on the *algorithmic level* (p. 75). The algorithmic level explains the process of getting from input to output, and how said processes are used. Therefore, Noveck (2018) argues that the falsification of the Standard Pragmatic Model does not undermine Gricean theory.

2.2.2 Searle’s Explicit Predictions

Paul Grice was notorious for his perfectionism and was reluctant to publish work (Chapman, 2010). As previously mentioned, Grice’s total exploration and explanation of metaphoric language was rather brief, but it is also responsible for the multitude of further academic explorations. In Grice’s own field of the philosophy of language, philosopher John Searle adopted and further developed Gricean theory. Searle further developed and made more explicit the theory’s implications for metaphoric language. Searle’s investigations focused on the problem of metaphor understanding, i.e., how can listeners comprehend something that is not literally communicated, or simply, how can metaphors be understood? (Searle, 1979)

Searle utilised philosopher J. L. Austin’s *speech act theory*, and argued that because metaphors convey additional meaning than the meaning lexically expressed by the words

uttered, they are ‘indirect speech acts’ (Searle, 1975). As Grice does, Searle distinguishes two different types of communicative meaning, *sentence meaning* and *the speaker’s utterance meaning* (Mácha, 2012). A speaker may therefore state ‘S is P’, which would be the ‘sentence meaning’, but intend to convey that ‘S is R’, which is then the ‘speaker’s utterance meaning’. An explanation of how this is possible must then explain the systemic and strategic abilities used to both convey, and understand, the intended but implicit meanings. Searle argues that literal utterances are those where the sentence meaning and the speaker’s utterance meaning coincides, and that metaphoric meaning will always be the speaker’s utterance meaning (Searle, 1979). Mácha (2012) argues that one crucial aspect of Searle’s work that distinguishes him from others is his commitment to define literal utterances, that therefore gives him a clearer starting point to define figurative meaning as opposed to literal, without muddling the definitions. Searle (1979) summarises the three relevant features of literal utterances for the understanding of metaphoric utterances as:

First, in literal utterance the speaker means what he says; that is, literal sentence meaning and speaker's utterance meaning are the same; second, in general the literal meaning of a sentence only determines a set of truth conditions relative to a set of background assumptions which are not part of the semantic content of the sentence; and third, the notion of similarity plays an essential role in any account of literal predication.

(Searle, 1979, p. 87)

Searle highlights how literal utterances are semantically under-informative, and that background assumptions are necessary for the comprehension of literal utterances, just as with metaphoric utterances. In addition, in the expansion of literal meaning, general terms must be understood through the similarities between properties expressed by that term, meaning that similarity between concepts are integral in many literal utterances. Searle also highlights the inadequacy of literal paraphrases of even simple metaphors and argues that some meaning intended by metaphors seem to have been either lost or inadequate when expressed literally. To use one of Searle’s (1979) examples:

2.2 Metaphoric: *Richard is a gorilla*

can be paraphrased to the literal

2.3 Literal: *Richard is fierce, nasty, and prone to violence*

The need to expand upon the single noun with several adjectives and an adjectival phrase shows the condensed information packed in the metaphoric vehicle expressed by a single noun, and Searle argues that still something seems lost in the transformation. The difference from 'S is R' to 'S is P' must be, and typically is, somehow understood by hearers. This understanding follows in Searle's view from a 'simple' principle:

The basic principle on which all metaphor works is that the utterance of an expression with its literal meaning and corresponding truth conditions can, in various ways that are specific to metaphor, call to mind another meaning and corresponding set of truth conditions.

(Searle, 1979, p. 89)

'Call to mind' is an expression that Searle himself admits to being imprecise, but he explicitly states that the truth conditions of the sentence's meaning, and the truth conditions of the speaker's utterance meaning are not the same. Finally, Searle makes explicit some of the pragmatic views taken for granted by Grice, such as figurative defectiveness, and the literal-first principle. Searle argues that hearers go through three steps when encountering a metaphor:

First, he must have some strategy for determining whether or not he has to seek a metaphorical interpretation of the utterance in the first place. Secondly, when he has decided to look for a metaphorical interpretation, he must have some set of strategies, or principles, for computing possible values of R, and third, he must have a set of strategies, or principles, for restricting the range of R's - for deciding which R's are likely to be the ones the speaker is asserting of S.

(Searle, 1979, p. 103)

He also explicitly states that *Where the utterance is defective if taken literally, look for an utterance meaning that differs from sentence meaning* (Searle, 1979, p. 103). Although Searle further developed Gricean theory and discussed metaphor in significantly greater detail than Grice, the contemporary view of metaphor has changed immensely since these first discussions on communicative comprehension.

As previously mentioned, psycholinguistic experimentalists wanted to test theories of language comprehension and processing. Testing the 'literal-first' aspect of Gricean theory, made explicit by Searle, could be done through numerous different experimental designs.

One of the earliest experiments testing the Standard Pragmatic Model by investigating metaphoric processing was Ortony et al. (1978). They proposed that it was not literalness that affected metaphoric processing speeds, but context. They conducted two experiments, using reading times of target sentences as a measure for processing speed. In the experiments participants were shown vignettes ending in a target sentence that would either be metaphorical or literal depending on the given context. One of their target sentences was 'Regardless of the danger, the troops marched on'. The relevant contexts shown before the target were either literal, describing several imminent dangers for soldiers, or metaphoric, describing children disregarding their babysitter's threats of punishment. In the experiment participants were shown either short or long versions of the contexts. Ortony et al. 's (1978) results showed that metaphoric utterances were understood slower than the literal utterances when preceded by short contexts. However, longer contexts lead to differences between literal and figurative utterances becoming insignificant. Their experiment immediately questioned the validity of the SPM and showed that it is at the very least insufficient for explaining what really occurs cognitively when we encounter metaphors. It should be noted that their experiment included both 27 practice items, and a five second interval (Ortony et al., 1978, p. 469) before the target metaphor was shown, thereby making the situation less natural, and more likely to prepare the participants for something out of the normal, potentially undermining their findings.

Glucksberg et al. (1982) gave participants short metaphoric sentences, scrambled metaphors, true literal sentences, and false literal sentences. The participants were to respond by identifying whether the sentences were true or false. The response times to indicate that metaphoric sentences were false were longer than literally false sentences, indicating that the participants could not help but process metaphors. Glucksberg et al. (1982) conclude that this means that the Standard Pragmatic Model is wrong in implicating that processing nonliteral language is a secondary process, only activated if a literal statement is found 'defective'.

Gildea and Glucksberg (1983) further investigated context's role in metaphor comprehension. They used contextual priming to activate either the figurative sense, the literal sense, or the general semantic field of a metaphor ground. Their results lead them to conclude that all types of contexts provided by them lead to 'immediate and automatic metaphor comprehension' (1983, p. 577). The automatic activation of metaphor comprehension again leads to problems for the Standard Pragmatic Model.

Pynte et al. (1996) utilised electroencephalography (EEG) to measure event-related potentials (ERP). They prepared metaphoric sentences as well as literal controls, and measured brain activity during the reading of said sentences. They provided either no context, or short relevant, or irrelevant, preceding contexts. The terminal words of the metaphor created larger N400 than the literal ones, which they argued meant that the literal meanings had been processed during comprehension. They also performed an experiment where ‘familiar’ metaphors, meaning metaphors more typically used in natural speech (an example being ‘X (human man) *is a pig*’, were given irrelevant contexts, and unfamiliar metaphors were given relevant contexts. Unfamiliar metaphors with relevant contexts created smaller N400s, or activation, than familiar metaphors with irrelevant contexts, exemplifying the importance of context for metaphor comprehension.

The Standard Pragmatic Model is generally considered insufficient and disproven (Noveck, 2018; Glucksberg, 2001). However, there are studies that give some credibility to it, although they clearly do not prove its validity completely, as the evidence against it is substantial.

Blasko and Connine (1993) utilised a cross-modal priming paradigm to investigate the processing time of metaphoric utterances. Their experiment was based on the psychological phenomenon of semantic priming. As will be discussed in greater detail later, semantic priming is a phenomenon that shows how related words, such as *doctor* and *nurse*, ‘prime’ each other, or make each other easier to process. For example, when the word *doctor* is heard or read, people process the word *nurse* more quickly than other semantically unrelated words (McNamara, 2005). Blasko and Connine (1993) utilised this phenomenon to investigate whether metaphoric language is activated as quickly as literal language.

Throughout the experiment, participants would listen to recorded utterances, such as *Aunt Mabel showed the boys how A DUSTY AND CROWDED ATTIC IS A PARADISE that lazy Saturday afternoon*. Whilst the participants were listening, a string of letters would appear on the computer screen in front of them, and they had to decide, as quickly and accurately as possible, if the string of letters was a real English word, or if it was nonsensical. The string of letters would appear either at 0 milliseconds after the metaphoric vehicle, in this case *paradise*, or it would appear 300 milliseconds after the vehicle had been said. The real words that appeared were either metaphorically related words, literally related, or control words. When participants heard the example *A DUSTY AND CROWDED ATTIC IS A PARADISE* the possible words appearing could be *memories*, *heaven*, or *joining*.

Their results showed that literal targets were more quickly reacted to than control words, at both delays. Metaphoric targets were also reacted to more quickly than control at both delays, but only when the metaphors were highly apt metaphors, meaning the metaphors that had been rated as good metaphors by an independent group of participants.

Blasko and Connine's (1993) results are congruent with the multitude of studies that shows that metaphor comprehension demands the same amount of time as the processing of literal language (Glucksberg, 2001). However, the study is inconsistent with research investigating the resolution of ambiguous words in context, such as Onifer and Swinney (1981). Onifer and Swinney (1981) display how only contextually appropriate meanings of ambiguous words are facilitated when there was a 150-millisecond delay in presentation of the target words. Blasko and Connine's (1993) results, however, show that the literal meanings were active at 300 milliseconds. These confounding results may be due to inadequate material. As Glucksberg (2001) highlights, in several of Blasko and Connine's (1993) contexts the literally related target words could be functional in the metaphoric contexts, and still make sense. He exemplifies this with the metaphor *his anger is a blizzard*. In the experiment the two targets for this context are *blinding* and *snowing*. *Blinding* functions as a metaphorically related word, but *snowing* will logically also be highly activated after the vehicle *blizzard*. An additional example would be the context *ritual is a prison*, with the targets *restricting* and *criminal*, again, both highly related words to the vehicle *prison*. As Glucksberg (2001) concludes, the experiment is adequate in displaying how metaphoric and literal meanings are processed at the same speeds but are inadequate in displaying the processing demanded by contextually inappropriate literal meanings.

Many of these early experiments rested on measuring the reading times of the actual metaphors, up against literal controls. And they were appropriate for addressing the insufficiencies of the Standard Pragmatic Model. However, as shown, the conclusions are not always complimentary. Although many portray the Standard Pragmatic Model as insufficient at an algorithmic level, they do not lead to a conclusion on metaphor processing. Unlike Grice (1975), Lakoff and Johnson (1980) working on conceptual metaphors, Sperber and Wilson (1996) working on Relevance Theory, and Glucksberg (2001) working on the Categorization approach to metaphor, all conclude that metaphors are ordinary language. However, metaphors being ordinary and metaphors requiring extra effort, are not necessarily incompatible.

2.3 Contemporary Theories on Metaphor

Although Philosophy provided the earliest extensive discussions of metaphor and its nature, experimentalists from psychology, linguistics, and pragmatics, have created extensive investigations and experiments attempting to reveal the nature, the features, and the processing, of metaphoric language. Holyoak and Stamenković (2018) provide a discussion and review of the key psycholinguistic theories dealing with metaphors.

In their review, Holyoak and Stamenković (2018) highlighted three major positions that attempt to deal with metaphors. They call these the ‘categorization position’, which will be further discussed here mainly through the work of Sam Glucksberg, the ‘analogy position’ discussed later through Gentner’s (Gentner & Clement, 1988) Structural Mapping theory of metaphor, and the ‘Conceptual mapping position’.

The conceptual mapping position is a theory on metaphors with a significant amount of literature. However, Holyoak and Stamenković (2018) are critical of the position as a relevant theory for the comprehension of novel metaphors. Unlike the analogical and categorization positions, the conceptual mapping position concerns mainly conventional metaphors. The theory argues that with conventional metaphors, a source conceptual domain is applied to a target conceptual domain. The source domain being generally more concrete, and the target domain being generally more abstract (Holyoak & Stamenković, 2018, p. 658). An example of this is the metaphoric statement *Time is money*.

The conceptual view argues that metaphors are common and part of ordinary language use because metaphors help explain conceptual mappings already existing in the brain (Wilson, 2011). Critically, Holyoak and Stamenković (2018) argues that the conceptual mapping theory is lacklustre as a psychological theory, due to many of the ‘abundant’ metaphors we encounter in daily communication, and mainly discussed by the positions, being so conventional that they are likely not registered as metaphoric at all by listeners.

Although Holyoak and Stamenković (2018) found evidence for, and against, all three metaphor positions in their review, they do argue that they have found some key aspects of metaphor comprehension that are well established. These key aspects include the rejection of the ‘literal first’ proposal that have been prevalent in earlier metaphor theories. They also include the proposal that metaphor comprehension is likely done through the same general

processes that thinking and language comprehension in general require. Additionally, they propose that through their review they have found more persuasive evidence for the category position. However, they also argue that the unfortunate focus on simple nominal metaphors may have given the analogy position a disadvantage, as the analogy position may be more apt to deal with more complex metaphors. Holyoak and Stamenković (2018) concludes based on their meta-analyses that more evidence is needed for conclusive theories on metaphor comprehension, especially studies and evidence based on complex and non-nominal metaphors. They also propose that it is likely that mechanisms based on both the analogical and categorization approaches should be integrated for a clearer understanding of metaphor comprehension.

2.3.1 Metaphors as Comparison: The Analogical Approach

One significant psycholinguistic theory dealing with metaphor comprehension is the analogical position, or the comparison view of metaphor. The comparison view, as mentioned earlier, argues that metaphors function as comparisons, as opposed to categorical assertions, as Glucksberg (2001) does. The earliest comparison views of metaphor seem to struggle with several features of metaphors but have been further developed through the work done by, among others, psychologist Dedre Gentner. Gentner developed the Structure Mapping Theory, which is applicable for analogies, similes, and metaphors (Gentner & Clement, 1988). The Structure Mapping model proposes that metaphors are comprehended through a comparison of similarities between the structures that each representation in the metaphor provides. Structure mapping consists of three stages. The first stage is an alignment of the representations used in the metaphor, i.e., the topic and the vehicle. After the structural alignment the features that are identical for both representations are matched. Additionally, relations between features are matched as well, which provides the opportunity for non-identical features to be matched, a problem for previous comparison view models. The second stage consists of the grouping of ‘structurally consistent connected clusters’, called ‘kernels’ (Gentner & Bowdle, 2008, p. 111). Furthermore, irrelevant features that cannot be structurally connected are dropped, which leaves only relevant features. Finally, in the third stage these kernels of relevant features are merged into one or more ‘structurally consistent global interpretations’ (Gentner & Bowdle, 2008, p. 111). During processing there are also inferences projected from vehicle to topic.

To use an example provided by Gentner and Wolff (1997), in the metaphor '*Tree trunks are (like) straws.*' the topic *tree trunks* and vehicle *straws* are structurally aligned and compared. After the structural alignment the features that are identical for both representations are matched. Additionally, relations between features are matched as well. This makes it possible for the structural alignment model to account for feature selection issues; non-identical features that otherwise would not logically be able to be matched become matched through their relations with features. In this example there are concepts linked between the representations that include that: both tree trunks and straws transport liquid, for both objects the liquid is typically water, and the method of transportation is suction. There are more elements that are cognitively linked between the subjects that would be irrelevant to link together for the comprehension of this specific metaphor (Gentner & Wolff, 1997). Metaphor comprehension therefore involves a complex comparison between several layers of conceptual knowledge, including the qualities of objects, the relations between objects, in addition to higher-order relations between relations (Holyoak & Stamenković, 2018).

One noteworthy development to the analogical position was the creation of the 'Career of Metaphor' theory. The theory combines the analogical and categorization positions and argues that novel metaphors are understood through analogical reasoning, whilst conventional metaphors are understood as a categorical statement. It therefore also argues that as metaphors get conventionalized (through their 'career' of use, if popular enough to become less novel) the processing method used by listeners changes (Bowdle & Gentner, 2005).

2.3.2 Lexical Modulation Theories of Metaphor

In the interdisciplinary field of Pragmatics, there are several theories that attempt to tackle metaphor comprehension and processing, from several different fields. Two important theories from different fields have fundamental similarities in how they view metaphor. Both the pragmatic Relevance Theory, and the psycholinguistic Categorization approach, view lexical modulation as the crucial mechanism for metaphor processing.

2.3.2.1 Relevance Theory

One of the most important pragmatic theories for metaphor comprehension, based on Gricean theory, is Relevance Theory. Unlike the computational theorising of philosophers of

language, Relevance Theory proposes a cognitive framework for communication, and makes empirical predictions for the cognitive processes involved in communication. The theory was created by Dan Sperber and Deirdre Wilson (1996) and has been further developed through both theoretical and experimental work. The theory synthesises Grice's maxims into one crucial notion of Relevance. Relevance Theory claims that 'Relevance' (*Relevance* being a technical term in this sense rather than the more commonplace non-technical term) is the crucial aspect of any human communication, where the balance between *cognitive effects* and *cognitive effort* is what provides listeners the ability to make sense of everyday communication. Listeners do so, because of the *Cognitive Principle of Relevance*.

The Cognitive Principle of Relevance is one of the crucial fundamental principles for Relevance Theory and is explained by Sperber and Wilson (2008) as: '*Human cognition tends to be geared to the maximisation of relevance*' (p. 7). There are two main aspects to explain an utterance's relevance, the cognitive effects achieved by the utterance, and the processing effort required to achieve these effects. Sperber and Wilson (2008) explains that:

- a. *The greater the cognitive effects achieved by processing an input, the greater its relevance.*
- b. *The smaller the processing effort required to achieve these effects, the greater the relevance.*

(p. 6)

Cognitive effects are vaguely described as something that '*... contributes positively to the fulfilment of cognitive functions or goals*' (Sperber & Wilson, 1996, p. 265). Meaning how much an input fulfils a goal, for example how well something answers a question. Cognitive effects are caused by three different types of utterances: (1) *New information provided by a contextual implication*, (2) *strengthening of an existing assumption*, and (3) *a contradiction and possible elimination of an existing assumption* (Gibbs & Tendahl, 2006, p. 397).

Processing effort is how much effort a listener needs to mentally expend to understand something as relevant. For example, if someone asks,

2.5 '*How is John and Louise doing?*'

Two possible responses that might be intended to convey approximately similar information could be:

2.6 '*Well, one of them isn't drinking alcohol for quite some months now.*'

2.7 'They are doing great; Louise is finally pregnant.'

2.6 might be a completely comprehensible, relevant answer but requires additional processing effort than saying 2.7. Because of the additional processing effort, the former answer is *less* relevant than the latter, but the latter utterance would lack some of the potential cognitive *effects* the speaker of the former utterance wants to create, in this case potentially the wish to create a 'dramatic' build-up to the joyous news. Some Relevance theorists argue that metaphors do lead to additional cognitive effort, with the payoff that the effort leads to greater cognitive effects (Gibbs & Tendahl, 2006).

Noveck et al. (2001) investigated whether or not cognitive efforts caused by metaphors lead to additional cognitive effects. They did two experiments focusing on metaphor comprehension and development in children. The experiments investigated whether metaphors require extra effort, and if that effort is beneficial. Using vignettes that included either a critical metaphoric sentence or a synonymic literal sentence and recording the responses to simple questions about the vignettes, as well as reading times for the critical sentences. Their results showed a slowdown for metaphoric sentences, however, in adults the correct responses for the simple questions regarding the vignettes were more correct when the critical sentence was metaphoric. The researchers concluded that metaphors cause extra effort, but that the effort caused a benefit for the adult participants.

These findings have however been questioned by other experimentalists. Gibbs and Tendahl (2006) argue that Noveck et al's (2001) use of imbalanced material undermines their results and conclusion. One example of their literal and metaphoric pairs is the literal utterance 'All children to the side of the pool' and the metaphoric 'All toads to the side of the pool'. Gibbs and Tendahl (2006) argue that the literal word 'children' only assigns a referent, whilst the metaphoric 'toads' both assign a referent and attributes properties to said referent. This may cause additional cognitive effects that interferes with the response times and undermines Noveck et al's (2001) conclusion. However, as previously discussed, this seems to be one of the crucial aspects of why we use metaphors; they convey more information than their supposed literal counterparts and this will therefore easily be a potentially confounding aspect of experiments dealing with metaphors.

Alongside the Cognitive Principle of Relevance is the *Communicative Principle of Relevance*, described as the central claim of Relevance Theory by Sperber and Wilson (2008). The Communicative Principle is that: *'Every act of inferential communication conveys a*

presumption of its own optimal relevance.' (Sperber & Wilson, 2008, p. 8). Ostensive-inferential communication is defined by Sperber and Wilson as such:

... The communicator produces a stimulus which makes it mutually manifest to communicator and audience that the communicator intends, by means of this stimulus, to make manifest or more manifest to the audience a set of assumptions {I}.

(Sperber & Wilson, 1996, p. 63)

That is, communicators make it mutually manifest to their audiences that we purposefully intend to convey some assumptions, and the communicative principle is that the way these assumptions are conveyed is optimally relevant for the purposes of the communicator.

Gibbs et al. (1991) conducted 3 different experiments, where participants were provided with metaphoric, literal, or anomalous comparisons, and told that the comparisons were either made by a 20th century poet, or randomly generated by a computer program. The experiments showed that the participants considered the metaphoric and literal statements to be more meaningful, and to contain more potential meanings when produced by a poet. The participants also had quicker reaction times when making judgments on meaning when they were produced by poets, and slower to reject the anomalous statements when supposedly produced by poets. Although rather intuitive, the study makes it clear that participants of communication are more willing to undergo additional cognitive efforts to understand statements when there is a perceived intention behind the statements, even if the literal words are the same. That is, the experiments provide additional support for theories that claim that utterances are heavily influenced by context. There is also an assumption that the metaphors bring some cognitive effects that would not occur with a literal alternative. However, in their experiment, Gibbs et al. (1991) claim to investigate metaphor comprehension, but seem to exclusively use 'metaphoric comparisons', i.e., similes. Their examples of material include the metaphoric comparisons 'A family album is like a museum' and 'A sermon is like a sleeping pill'. According to both researchers within the Relevance Theory framework, and the Categorization view of metaphor, metaphors and similes do not function the same way, and are not interchangeable (Carston & Wearing, 2011; Glucksberg, 2008; Glucksberg & Haught, 2006).

The cognitive and communicative principles of Relevance are some of the core concepts of Relevance Theory in general, which are clearly relevant for metaphor comprehension. Listeners assume that metaphors are relevant, and therefore worthwhile

processing. The Cognitive and the Communicative principles of communication can be used to create the '*Relevance-Theoretic comprehension heuristic*':

- a) *Follow a path of least effort in constructing an interpretation of the utterance (and in particular in resolving ambiguities and referential indeterminacies, enriching or adjusting the encoded meaning, supplying contextual assumptions, deriving implicatures, etc.).*
- b) *Stop when your expectations of relevance is satisfied (or abandoned).*

(Wilson & Carston, 2006, p. 408)

This Comprehension heuristic exemplifies how a listener processes an utterance. A listener would attempt to understand a speaker's meaning in a way that would satisfy the presumption of optimal relevance. In this attempt, the listener would have to enrich the decoded utterance at the explicit level, through disambiguation, reference assignment and other processes, and complement the utterance at the implicit level, by utilising contextual assumptions. The contextual assumptions alongside the enriched explicit utterance provide the listener with implications or cognitive effects that presumably makes the utterance relevant in a way expected by the listener. Simply put, the listener attempts to find the easiest explanation of an utterance that satisfies the listeners expectations of relevance (Wilson & Carston, 2006). Wilson and Carston (2006) claims that these processes crucially occur in parallel, not in sequence. They also claim that the comprehension heuristic is not exclusively used for comprehending figurative language, but also for processing literal utterances.

In contrast with Grice's 'literal first' approach to language comprehension, Sperber and Wilson (2008) claim that verbal comprehension '*involves no presumption of literalness and no default interpretation, and that metaphors are in no way exceptional*' (p. 5). The view that metaphors form a natural part of ordinary human language is not unique to Relevance Theory. Cognitive Linguists working on Conceptual Mapping Theory also view metaphors as an ordinary aspect of human communication. The Conceptual Mapping Theory argues that metaphors are common and ordinary because they reflect the metaphorical way our cognitive abilities function (Lakoff & Johnson 1980). Wilson (2011) highlights how Relevance Theory, just as cognitive linguistics, sees metaphors as ordinary, natural language, but for different reasons. Relevance Theory views metaphors as tools to convey complex thoughts that are often vague, but not metaphorical in nature (Wilson, 2011, p. 178).

Relevance Theory crucially claims that *all* verbal communication is inferential and context sensitive (Sperber & Wilson, 1996, p. 64). This claim is based on the idea that communication consists of both ‘coded’ communication as well as ‘ostensive-inferential’ communication, where the ostensive-inferential communication is argued to be the crucial aspect of communication. Sperber and Wilson (1996) argue that coded communication is *only* used to ‘strengthen’ the inferential communication (p. 64). Prior to Grice’s influential work, the ‘code-model of communication’, was the assumed model of communication (Noveck, 2018). A code model view of communication assumes that speakers have thoughts, or messages, that they encode and then transmit. This transmitted code is then received by whomever is able to receive it, who then decodes the message. If this process is done without issue or interference, the receivers of the code will then have decoded and duplicated the encoded message, and successfully understood the exact thought intended by the speaker (Wilson, 1998, p. 2).

As opposed to this code model, the inferential aspect of Relevance Theory is based on the idea that speakers are only capable of providing clues to their audience, and these clues help their audience infer the speakers’ approximate thoughts. If so, it follows that metaphor comprehension could be dependent on the same processes that literal comprehension requires, seeing as essentially all communication needs to be comprehended through inference, and the exact intended meaning is not assumed to ever be completely reached by the audience of communication. According to Wilson (1998, p. 4) Grice’s ‘main contribution’ to pragmatics was that he put forth the basis for an inferential model of communication, giving an alternative to the previously default view of communication, based on the code model of communication.

Relevance Theory, as a theory based on Grice’s theories, puts a significant weight on inference (yet does not completely dismiss the process of coding and decoding as an aspect of communication). Utterances are clues that provide hints of the speaker’s thoughts. As this indicates, utterances are seldom, if ever, able to fully convey the exact thoughts of the speaker. Utterances typically underdetermine the speaker’s thoughts, in many ways (Carston, 2015). Examples of how utterances may underdetermine speakers’ meanings include:

Underdetermined Language	Examples
1. Lexical or structural ambiguity	<i>Bank</i> – riverbank or financial institution
2. Referential Indeterminacies	<i>I, we, they, 'the workers'</i>
3. Unspecified quantifier domains	<i>'Everyone' loves linguistics.</i>
4. Incomplete expressions	<i>(The item you are looking for is) behind the door.</i>
5. Vague expressions	<i>She is 'young', He is 'tall'</i>
6. Implicit clausal connections	<i>They left – the professor talked for ages</i>
7. Approximations	<i>The classroom is 'empty'</i>
8. Figurative language	<i>His words are empty</i>
9. Illocutionary indeterminacies	<i>You are leaving</i> – this could be e.g. a bet, prediction or a promise.

Table 1 Examples of underdeterminacies of language.

(Carston, 2016, p. 2)

Underdeterminacies are abundant in everyday language, and communicators constantly create and encounter utterances containing ‘gaps’ that need to be filled in for the utterance to become ‘complete’. Therefore, according to Relevance Theory, metaphors are on a continuum of ‘loose use’ language and is processed in the same ways as many ‘literal’ utterances, through ‘Meaning Adjustment’. This continuum of loose use language includes hyperboles, category extensions and approximations (Carston, 2010). The same utterance may be used in many different ways, based on the context and manner it is used in, for example:

2.8 Speaker X: *The room is empty.*

Spoken by a member of a school's janitorial staff, for example, 2.13 can be considered a literal, but incomplete expression, such that it 'should be' *The room is empty (of chairs)*. Disregarding the need for chairs, the utterance could also be an approximation, there are a few tools, some paint buckets, but the room is approximately empty. Or the utterance may be hyperbolic, spoken by a lecturer expecting many more students than the few who showed up. Or it may be an ironic statement, spoken by a colleague of the pessimistic lecturer who expected few students, when the room appears full of students. The meaning of the word *empty* is dependent on the context, and listeners in their respective contexts will infer the relevant meaning through meaning adjustment.

Unlike Grice's view of communication, where communicators follow the Maxims of conversation either by what they say, or by what they mean, Relevance Theory does not propose that speakers are dependent on Maxims to create valid interpretations of utterances. Relevance Theory does not argue that speakers attempt to be truthful by following a maxim of quality, but that every act of communication carries presumption of relevance. These acts may fail to be relevant in a given context, but still carry the presumption of relevance. Through the Cognitive and Communicative principles of Relevance, communicators will attempt to comprehend the most relevant meaning intended. Therefore, metaphors do not break any rules of truthfulness, but carry presumptions of relevance that listeners will attempt to infer, just as they will attempt to infer the most relevant intended meanings conveyed in any form of communication (Noveck, 2018).

Relevance Theory categorises metaphor as one of many types of 'loose use' of language and claims that there is a continuity between loose use of what is commonly viewed as literal language, but is not necessarily 'strictly' literal, such as approximation, and figurative language, including creative metaphors (Sperber & Wilson, 1996, p. 235). Relevance Theory explains the processing of metaphors, and other loose use language, through meaning adjustment. According to Relevance Theory, loose use is one way for speakers to reduce cognitive effort when communicating yet maintain optimal relevance and create sufficient cognitive effects (Carston, 2017). Common examples of utterances that are viewed as literal language but require meaning adjustment to be strictly literal would be:

2.9 *I live thirty minutes from work.*

2.10 *I make 50 000 pounds a year.*

2.11 *The orchids are dead.*

The speakers of these utterances may live *approximately* thirty minutes from work, make *approximately* 50 000 pounds a year, and the orchids may be so close to death that the speaker will not attempt to salvage them, and wish to convey as such. All three utterances are common types of utterances, easy to understand, and sufficiently conveys the meaning the speaker wishes to convey. Uttering

2.12 *I live 28 minutes and 35 seconds from work.*

2.13 *I make 48950 pounds a year.*

2.14 *The orchids are so close to death I have given up on attempting to revive them.*

Would naturally provide the audience with more literal, more detailed and more ‘correct’ information, but also demand more cognitive effort for minuscule gains in cognitive effects, making (2.12), (2.13), and (2.14) less relevant for typical speakers’ communicative purposes, and their attempts at optimal relevance.

Some Relevance Theorists argue that a full explanation of metaphor comprehension must include a split between two modes of processing (Carston, 2010). Carston (2010) proposes that there are generally two types of metaphors, the ‘ordinary’ and the ‘literary’, each typically requiring different modes of processing. Carston (2010) uses the commonly used metaphor

2.15 *My lawyer is a shark*

As an example of an ‘ordinary’ metaphor. The ordinary metaphors are described as *conventional, conversational, spontaneous, single-word metaphors*, and having propositional content, meaning it is possible to agree or disagree with the statement. Carston (2010) argues that when ordinary metaphors are encountered, a quick, local, on-line meaning adjustment process is used for comprehending the metaphor. Literary metaphors are exemplified by Carston (2010) with the poem:

2.16 *The fog comes*

on little cat feet.

It sits looking

over harbor and city

on silent haunches

and then moves on.

(Sandburg, 1916)

The literary metaphors are described as ‘*carefully crafted, extended and developed, expressive of a feeling or sensation, highly imagistic*’ (Carston, 2010, p. 297). Carston (2010) argues that it would be problematic to contest the propositional content of such a literary metaphor, and it would be more apt to comment on the metaphor’s expression of a feeling or experience.

Carston (2010) also argues that the literary metaphors require a ‘*slower, more global appraisal of the literal meaning of the whole*’ (Carston, 2010, p. 297). When a speaker encounters extended literary metaphors as in (2.16), the ‘lingering of the literal’ becomes a more essential aspect of the metaphor comprehension. Due to the extensive priming of the continued metaphor use in the developed metaphors, quick on-line ad hoc concept creation becomes too demanding, and the second mode of metaphor processing gets activated, where the literal meaning becomes more important. Carston (2010) concludes that although there seem to be two modes of processing metaphors, and potentially two ‘types’ of metaphor, this does not break from Relevance Theory’s deflationary account of metaphor comprehension. The two modes of processing have ‘no difference in kind’, and both modes are always available, if a communicator wants to explore the implications of the literal. The cognitive processes required by metaphors are also crucially not exclusive to metaphors (Carston, 2010, p. 318).

Meaning adjustment leads to either a narrowing or broadening of the concept utilised in the utterance, or potentially both. Narrowings are when the meaning intended by a word is narrower than the typical lexical meaning of said word. Some examples of narrowings include:

2.17 *Now this is wine.*

2.18 *That guy can punch.*

2.19 *I have never had a drink.*

Both (2.17) and (2.18) contain a concept that is narrower than the lexical definition of the word, but with an included degree of quality, or severity. The *wine* in (2.17) may be meant as an especially high-quality wine, not ‘simply’ fermented grape juice, as per definition, or in a wine discussion it could be something more specific, high-quality wine stored in oak barrels,

as opposed to cheap wine oaked with oak ‘chips’. In (2.18) *punch* would most likely be considered a punch that is especially painful or damaging, not just any strike with a fist. Example (2.19), although now lexicalised, has been a narrowing of the general concept of any liquid meant for consumption to specifically an alcoholic drink. As shown in these examples when a concept is narrowed a degree of literalness is maintained (Sperber & Wilson, 2012), a high-quality wine is still a wine, a ‘good’ punch is by definition a punch, and an alcoholic drink is simply a subcategory of drink. Wilson & Carston (2007) argues that a degree of meaning adjustment happen in possibly all of communication, citing the nuanced usages of common verbs such as *open*, *cut* and *leave*, as argued by Searle (1980) and the variance of antonyms to the same lexical item, as highlighted by Murphy (1997). Searle (1980) displays how common verbs, such as *open*, has slightly varied meaning depending on context, we *open a book*, *open a road*, *open a store*, *open a bottle* and so on, although there is a vague common meaning in the act, it is clear that the literal exact meaning in each context varies greatly. Similarly, Murphy (1997) did an experiment where he asked participants to name antonyms of adjectives but used the adjective to modify different nouns. Depending on the noun, the adjectives' antonyms varied severely. An example would be the adjective *fresh* and its possible antonyms:

Adjective	Noun	Antonym
Fresh	Vegetables	Rotten
Fresh	Fish	Frozen
Fresh	Water	Salt

Table 2 Examples of varied antonyms of the same noun *Fresh*

The nouns in the study were not exclusively food items, but even by displaying exclusively food nouns, *fresh*'s antonyms change dramatically based on context.

The inverse of narrowing, broadening, does not maintain literalness. One example of broadening has already been mentioned in Table 1, the approximation, but more examples of broadenings include:

2.20 *Denmark is flat.*

2.21 *The earth is a sphere.*

2.22 *What a nice car you have there* (Told to a toddler in a cardboard box.)

Both (2.20) and (2.21) are approximations, Denmark is not literally flat, just as the earth is not a perfect sphere. Example (2.22) is a category extension, and has been used to include a cardboard box, maybe with a cardboard seat and a cardboard steering wheel to the concept *CAR*, so the box might share a few physical characteristics (in their vaguest of definitions) and is therefore included in a much more detailed concept.

Metaphors may contain *both* narrowing and broadening (Sperber & Wilson, 2008, p. 6). An example of this would be for the metaphorical utterance of the previously used example

2.23 *Frank is a shark*

An ad hoc concept *SHARK** is constructed, which is both narrower and broader in meaning than the lexical meaning of *shark*. It is narrower, because it does not necessarily include fish, but it is also broader, because it now includes vicious, aggressive humans. Crucially, Relevance Theory argues that these utterances are understood through the employment of the same cognitive process, *meaning adjustment*. Through meaning adjustment, an *ad hoc* concept is created, which contains the relevant features needed for the utterances to become sufficiently relevant for the purposes of the exchange (Carston, 2015). The creation of ad hoc categories is also crucial for how the psycholinguistic Class Inclusion theory explain the processing of metaphoric utterances.

2.3.2.2 Class Inclusion Theory

The Class Inclusion Theory of Metaphor is a prevalent psycholinguistic theory dealing with how metaphors function, and how they are processed. Although the theory has been developed by many experimental researchers, it has been especially shaped by the prolific experimentalist Sam Glucksberg and is discussed here mainly through the work of Glucksberg. Glucksberg has dedicated a substantial amount of his academic work to investigate the nature of metaphors and how they are processed. He has also based a lot of his work on disproving the implications of Gricean theory in terms of metaphoric language.

Although Glucksberg (2001) supports Grice's theory of the Cooperative Principle (p. 5), he vehemently denies all empiric implications of the Standard Pragmatic Model, as proposed by Grice (1975) and Searle (1979). As previously stated, the Standard Pragmatic Model consists of three steps:

1. *Derive the literal meaning of an utterance.*
2. *Test the derived literal meaning against the context of the utterance.*
3. *If the literal meaning makes sense, accept that meaning as the utterance meaning, that is, the speaker's intended meaning. If it does not make sense, then seek an alternative, nonliteral meaning that does make sense in the context.*

(Glucksberg, 2001, p. 10)

Metaphors are typically 'false' statements and as mentioned, are in Searle's view 'defective' (1979). 'Defective' utterances will in interpretations be transformed by hearers from categorical statements to comparison statements, i.e., similes. This comparison model views metaphors as implicit similes. Glucksberg (2001, pp. 10-11) argues that there are four main implications that follow from the Standard Pragmatic Model and the Comparison view of metaphor that may be investigated. These four are:

1. *... Literal meanings are unproblematic and context-free.*
2. *... Literal meanings have unconditional priority.*
3. *... Literal meanings are derived automatically, but nonliteral meanings are derived only optionally.*
4. *... [Metaphors] are implicitly transformed into true comparison statements and interpreted via a comparison process.*

Glucksberg (2001) claims that every single implication of the Standard Pragmatic Model is false, including the view that metaphors are comparison statements.

To dispute the first implication, *literal meanings are unproblematic and context-free* Glucksberg (2001) first wishes to establish what exactly constitutes *literal* meaning. In doing so Glucksberg (2001) differentiates between two processes of language comprehension, *linguistic decoding* and *linguistic interpretation*. Linguistic decoding is defined by Glucksberg as the linguistic operations utilised to comprehend an utterance, such as phonological, lexical and syntactic operations. Therefore, in Glucksberg's (2001) view, literal meanings can be described as ... *the products of a particular (one hopes, the "best") theory*

of semantics and syntax, a theory that does not pretend to describe or explain what people actually do when talking and listening (p. 11). Meaning that literal meanings ‘should’ be able to be comprehended based solely on the results of linguistic decoding.

In his explanation and definition of metaphoric language Glucksberg (2001) first attempts to define literal language. In a non-linguistic folk theory of literal language, literal language is viewed as the primary meaning of a word. These primary meanings are commonly viewed as context independent, that is, literal utterances will be unambiguously understood regardless of the context. Glucksberg exemplifies this with the utterance

2.24 *Dogs are animals.*

This utterance, with its context independent primary meaning would be interpreted as a category statement claiming that *dogs* are part of the category *animal*, as opposed to, for example, the category *vegetable*. A nonliteral interpretation of the same utterance would be context dependent, and based on a certain context, the utterance may be used to explain the behaviour of a dog, claiming that one cannot expect any less, because *dogs are animals*.

Glucksberg (2001) questions the context independence of literal words, due to the ambiguous nature of even the most basic terms, assumed to be literal by laypeople. These basic terms include *connectives*, such as the *implication* and *conjunction* connectives form formal logic; *if – then*, and *and*. Furthermore, they also include words from common word classes such as quantifiers (some), adjectives (good) and pronouns (he, she, they).

In formal logic, connectives are used to connect logical content. *If – then*, or the *implication* connective is symbolised by \rightarrow , and carries context independent meaning within formal logic. This implicative meaning is that two conditions are connected through the implication that *if p then q*, or, to use an example, *p* may equal *the dogs play well together* and *q* could be *we stay for an hour*. Two conditions connected with an implicative are considered true through several different options, the total amount of conditions is exemplified in 1-4:

1. *p and q*: The dogs play well together and we stay for an hour.
2. *not p and not q*: The dogs do not play well together and we do not stay for an hour.
3. *p and not q*: The dogs play well together and we do not stay for an hour.
4. *not p and q*: The dogs do not play well together and we stay for an hour.

In formal logic, $p \rightarrow q$ is only considered false when the conditions are 3: *p and not q*. Both 1 and 2 seem pragmatically logical, but 4 is considered true under formal logical, but causes

pragmatical issues. Statements such as *if you have drunk wine then you are older than 18* expresses a prerequisite condition for drinking wine, being above the age of 18, rather than expressing a logical relation, *if you have not drunk wine you are older than 18* seems pragmatically problematic, yet is considered a true statement for the formulation $p \rightarrow q$.

Similarly, the connective *and*, formalised as \wedge , is true when both p and q are true separately. However, in natural language use, *and* is used to convey both conjunction and implications, i.e., *Bring me the bottle and I will give you a glass*, conveying an implication, or an *if – then* condition. As shown, because even the most basic logical terms utilised in communication, typically viewed as both basic and literal, is clearly dependent on contextual effects, Glucksberg (2001) argues that lexical modulation occurs at almost every stage of communication, regardless of the figurativeness, or so-called literalness of the words used. Furthermore, Glucksberg (2001) highlights how commonly used ‘literal’ words such as *good* may carry several different senses, such as the previously mentioned *open* (Searle 1980) and *fresh* (Murphy 1997). To expand on it, Glucksberg uses examples from (Bierwisch, 1967):

2.25 *He got a good whipping for being late.*

2.26 *Harry Truman was a good president.*

2.27 *Hannibal Lecter was more than just a good villain.*

In these examples, the common adjective *good* seems to carry greatly varied meanings, such as *painful*, *honest*, *terrifying*.

Glucksberg (2001) argues that although folk theory of language has clear divisions between what is literal and what is non-literal language, literal meanings cannot be ‘*explicitly defined in formal linguistic-theory terms*’ (p. 14). He concludes that literal language, as well as metaphoric language, behaves as natural kind concepts, i.e., they have clear prototypical examples, but also unclear non-prototypical examples that are more difficult to include in the category of literal, or metaphoric, language. One such unclear example that Glucksberg (2001) uses is *glued to the TV set*. *Glued* does not carry the primary meaning of the verb *to glue*, yet intuitively seems literal in its usage. This displays the disconnect between how people view language, and how they use it. Glucksberg (2001) claims that:

... *people may use and understand metaphorical expressions without being aware that the expressions are metaphorical at all. This should certainly be the case if metaphor*

and literal understanding depend on the same linguistic, cognitive, and pragmatic principles.

(p. 15)

Glucksberg (2001) concludes that the vague nature of both literal and metaphoric language makes the distinction less clear than one might expect. As shown, *literal meanings* being *unproblematic* and *context-free* seems unsatisfactory at best. The vagueness of literal language also makes it possible that metaphoric and literal language is dependent on the same cognitive operations.

To dispute the second implication of the Standard Pragmatic Model, *literal meanings have unconditional priority*, Glucksberg (2001) argues that 1. Literal decoding is context sensitive, as we have seen, and 2. Word recognition is context sensitive. The sensitivity of word recognition has been proven through literature on disambiguation, which has shown that appropriate contexts will make disambiguation and meaning selection clear. Research such as Simpson and Krueger (1991) showed that contexts will have immediate effects on disambiguation, and that both meanings of ambiguous words were only activated after a delay, as opposed to an early selection stage with activation of both meanings immediately. Contextual information also activates word recognition before words have been fully heard (Tannenhaus et al., 1995).

The third implication, that figurative language is optional, is based on the idea that figurative meanings are only processed if an utterance has been found 'defective' as a literal statement. So literal meanings are automatically processed, whilst figurative meanings are optionally investigated once an utterance has been ruled out as literal. The 'optional' nature of metaphors has been investigated by Glucksberg, inspired by the psychological 'Stroop effect'. The Stroop effect is that when people are shown a word and asked to name the colour of the ink, or pixels, of the word, if the word is a colour term, such as *blue* whilst the colour of the ink is red, meaning the participant should say *red*, their reactions are delayed. This is because the participants are unable to not read the word. Had the participants solely looked at the colour of the ink, they would have responded more quickly, but participants cannot help but read and process the word, delaying their response in naming the colour of the ink. Glucksberg et al. (1982) asked participants to judge whether statements were literally true or false. These statements could be either literally true, such as *Some fish are trout*, some were literally false, *Some fish are eagles*, scrambled metaphors *Some jobs are snakes* and

metaphors *Some jobs are jails*. The results showed that participants required longer time to reject metaphors as literally false than they did when presented false statements and scrambled metaphors. Because of this delay, Glucksberg et al. (1983) conclude that people are unable to inhibit the consideration of metaphorical meaning, just as people are unable to inhibit reading words in front of them, meaning, metaphoric meanings are not optionally processed.

The last implication requires a more thorough explanation, as it separates Glucksberg's view on metaphor from other psychologists investigating metaphors. The fourth implication of the model as stated by Glucksberg (2001): *Metaphors are implicitly transformed into true comparison statements and interpreted via a comparison process*. Glucksberg argues that this last implication, as all the other implications, are false. This implication is in accordance with the previously discussed analogical position. In short, the comparison view of metaphor sees metaphors as implicit similes, so that literally false metaphors such as

2.28 *My lawyer is a shark*

Are transformed into the literally true simile:

2.29 *My lawyer is like a shark*

Which is then treated as any other literal comparison statement. Glucksberg argues that the comparison view is inadequate as an explanation of the function of metaphor. His rejection again requires a basic understanding of what constitutes a comparison, and how a metaphor would, or would not, function as a comparison.

Literal comparisons, such as *Orange wine is like red wine*, compare features of *a*, here *Orange wine* with *b*, *red wine*. In such a comparison, only relevant features are used for interpretation. Some essential features are presumed to be the same for both *a* and *b*, and therefore irrelevant for comparison. In our example these essential, irrelevant features for comparison might be that they are both liquids, both made for consumption, both made in similar regions. In this simple example, there is one crucial feature that might be used for comparison, and several more possible. The crucial feature is that both types of wine are made with the grape juice being in contact with the grape skin for a certain amount of time, as opposed to white wine where the grape juice is quickly extracted without a significant amount of skin contact. Other features could be that they are both alcoholic, both fermented, and

much more. There are also those relevant features that are not common for *a* and *b*, which are relevant as they set the two components apart. Orange and red wine are obviously set apart by their defining colours, which is due to one more feature that sets them apart, the type of grapes they are made from. Orange wine is made from green grapes, and red wine is made from blue grapes. In the process of comparison, the relevant features in common and the relevant features not in common are weighted and compared against each other, which provides a source for directionality in comparison. Often, a comparison statement of *a is like b* will differ from the statement *b is like a*; *Red wine is like orange wine*, seems to convey something else than the original example. More weight is assigned to the subject of the comparison, *a*, than the predicate, *b*. For many, red wine is more prototypical than orange wine, and the more prototypical a component is, the more salient it is. Variants of a prototype will always be more similar to the prototype, than the prototype will be of its variants. Glucksberg (2001) argues that these features of literal comparisons are inadequate as explanations of metaphorical comparisons.

With metaphorical comparisons, there will often be no relevant distinct features, which removes directionality from the equation. The common features should be applicable in both directions. Using Glucksberg's (2001) example of a metaphoric comparison *My job is like a jail*, shows clearly that the metaphor *is* directional, *a jail is like my job* does not work. Glucksberg argues that metaphors are not only asymmetrical, but they are also non-reversible. Those metaphors that do seem reversible do however apply different meanings in the reversed form, and therefore become new, different metaphors. A common example in the metaphor literature *My surgeon is a butcher*, implies features of gross incompetence or carelessness to the surgeon, whilst the reversed metaphor *My butcher is a surgeon* conveys the opposite, applying high skill and carefulness to *my butcher*.

The non-reversible nature of metaphors causes problems for the applicability of models of literal comparison statements. Additionally, these models do not provide any way to judge levels of metaphoricity. Although literalness may sometimes be vague, communicators are generally able to judge if a comparison is literal or metaphorical, as well as judging levels of metaphoricity between metaphoric statements. Glucksberg exemplifies this with the two statements *John's face was like a beet*, and *John's face was red like a beet*, both metaphorical in nature, but the former 'more' metaphoric than the latter. Ortony (1979) explains metaphoricity with a salience imbalance. A metaphoric comparison requires a matching feature that has low salience for the topic but a high salience for the vehicle. In the

previous example, *My job is a jail*, ‘my job’ would then have a low salient feature of ‘constraining’, for example, which is a highly salient feature of ‘jail’. However, Glucksberg (2001) argues that this might be adequate as an explanation of how metaphors function, but the explanation does not manage to separate literal and metaphoric statements. This imbalance in information is essential for all informative comparisons (Glucksberg, 2001). Additionally, if the comprehension of metaphors includes matching properties between topics and vehicles, metaphors that utilise vehicles conveying completely ‘irrelevant’ information for the metaphor topic are only able to be matched for features *after* the metaphor itself is understood. This comparison view is therefore inadequate as an explanation of how listeners understand metaphors.

Glucksberg (2001) does argue that Gentner’s (Gentner & Clement 1988) structure mapping model is capable of accommodating many of the previously discussed problems for comparison views, such as feature selection, the attribution of new features, as well as the non-reversible nature of metaphoric comparisons, but falls short for explaining two problems, how people identify metaphors as metaphors, and how similes can be transformed into metaphors and vice versa.

Glucksberg (2001) highlights one key feature that separates metaphoric comparisons from literal comparisons. Metaphoric comparisons, i.e., similes, can be transformed into class inclusion statements, whilst literal comparisons cannot. Similes such as:

2.30 *My lawyer is like a shark*

Is easily transformed into

2.31 *My lawyer is a shark*

Whilst our earlier literal comparison becomes problematic:

2.32 *Orange wine is red wine*

Glucksberg (2001) argues that literal comparison *always* involves components at the same level of abstraction. Literal comparisons may be understood by joining the two components into a common category, in ‘Orange wine is like red wine’, the components could be in the category ‘fermented grape juice’. Literal comparisons may also be used to introduce new information, such as *cider is like wine*. In this example, the new information may be that the alcoholic beverage cider is more alike with wine than beer, a type of alcoholic drink it may be more closely associated with. However, both are in the category ‘fermented fruit juice’; the

comparison may be used to imply this relevant information. Crucially, Glucksberg (2001) argues that metaphors are understood in exactly the same way, only through the use of abstract categories. In (2.31), both *lawyer* and *shark* are part of the same abstract category of ‘predatory animal’.

Essential for Glucksberg’s (2001) view of metaphor are two concepts, the concept of *dual reference*, and that of *novel metaphorical categories*. Novel metaphorical categories are as they seem; new categories used for the creation and comprehension of metaphors. Words may be described as part of an indefinite number of categories, based on what the word is used for. ‘Shark’ may be part of basic literal categories such as ‘seafood’, ‘predatory fish’, and ‘group of typically endangered animals’. Furthermore, ‘shark’ may be used as a representation of more abstract categories, in our example, ‘predatory animal’, this abstract category may contain only properties that do not exclude human lawyers from inclusion in the group, such as ‘vicious’, ‘cold-hearted’, and ‘highly-driven’.

The dual reference of words is essentially that any word may be used as representations of several different categories. This dual reference may be explicitly expressed. In Glucksberg’s (2001) example:

2.33 *Cambodia is Vietnam’s Vietnam.*

The first *Vietnam* refers to the concrete South-East Asian country, whilst the second *Vietnam* refers to the abstract category of ‘horrendous military failures’, lexicalised due to the United States’ failure in the American – Vietnam war. This example succinctly shows the dual reference possibilities of words. Interestingly, because of this dual reference, Glucksberg (2003) argues that metaphors are in some sense meant literally. Metaphors do not employ a figurative use of a literal concept, but a literal use of an abstract category; separate but related to the dual literal category. This explains in short Glucksberg’s (2001) view of how metaphors function, but it does not explain how listeners comprehend metaphors. For the comprehension metaphor, Glucksberg proposes what he calls an *interactive property attribution model*.

The interactive property attribution model claims that constituents of a metaphors provide dimensions and salient features for attribution. This model is also applicable for the comprehension of nominal compounds (Estes & Glucksberg, 2000). When we encounter nominal compounds, i.e., noun-noun compounds such as *glass bottle*, the head *bottle* provides the dimension of which the modifier provides salient features to attribute. *Bottle* may provide the dimensions of *MATERIAL*, *COST*, *SHAPE*, whilst the modifier provides features, such as

'glass', for the dimension *MATERIAL*. Similarly, in nominal metaphors salient features of the metaphor vehicle are attributed to the metaphor topic. Depending on the context the word *shark* may provide salient features such as 'tasty' 'quick-swimmer' and 'aggressive'. In (2.31) only 'aggressive' would be a salient feature for the dimension provided by this specific metaphor topic, *TEMPERAMENT*. Furthermore, because of the novel metaphoric category *shark*, which is, as mentioned, an abstract category named by its literal referent, features such as 'quick-swimmer' would not be salient, as these are not properties of the abstract category 'shark'. Literal features of the fish 'shark' should therefore not be valid, nor considered during metaphor comprehension, as metaphors such as (2.31) does not include any word referring to a literal marine animal.

2.4 Defining Key Experimental Concepts

This section will further elaborate and discuss some of the key concepts relevant for contemporary theories on metaphor, and crucially, the experiment conducted as part of this thesis. These key concepts include metaphor's processing factors, ad hoc categories, the cognitive mechanisms of suppression and enhancement, as well as priming. These concepts are also all linked together. Firstly, the relative ease of the processing metaphors is dependent on several different processing factors: the conventionality, the familiarity, and the aptness of the metaphor. Secondly, Relevance Theory and the Categorization approach to metaphors both propose similar cognitive processes for the comprehension of metaphoric utterances. They both argue that lexical modulation occurs through the creation of an Ad Hoc category. The Ad Hoc category is constructed through the enhancement of relevant properties of the category, and the suppression of irrelevant properties, i.e., metaphor processing is dependent on the cognitive processes of suppression and enhancement. One of the most common ways of testing suppression and enhancement is through priming tasks, which attempt to manipulate the activation of different relevant or irrelevant concepts to the experimental material. Before discussing the experiment conducted as part of this thesis these crucial aspects of the experiment will first be explained and discussed.

2.4.1 Processing Factors

How difficult a metaphor is to understand is based on many factors. Three key factors that affect the level of difficulty are *conventionality*, *familiarity*, and *aptness* (Ronderos et al.

(under review)). Though separate factors, they are often confounded (Jones and Estes, 2006). They differ in what level they affect the metaphor property, at either word or sentence level. Conventionality refers to the frequency of which a metaphoric vehicle is used as such, as opposed to its literal meaning (Bowdle & Gentner, 2005). Conventionality is therefore a word-level property, seeing as it refers to a word's frequency of metaphoric use. Familiarity is a sentence level property and refers to the frequency of the metaphor as a whole, not just the metaphoric vehicle. Lastly, aptness is a sentence level property and refers to the degree in which the figurative features of the metaphor vehicle are relevant for the metaphoric topic (Jones & Estes, 2006). These factors seemingly all affect the processing effort required for comprehending metaphors, in different ways, and at different degrees. They also potentially represent different manners in which metaphors are processed, depending on the type of metaphor. The 'career of metaphor' hypothesis argues that a metaphor's level of conventionality affects how the metaphor is processed. The hypothesis claims that metaphors most often require analogical reasoning, where the topics and vehicle are scanned and compared for relational similarities. After the scan, inference is used from vehicle to topic. Crucially, the Career of Metaphor claims that as a metaphor becomes more frequently utilised, and encountered by communicators more often, the mode of processing will with increasing likelihood change from analogical reasoning to category statement. The increased use of a metaphor vehicle will eventually, potentially, result in a new polysemous word. When the vehicle has become polysemous, the process required changes again, and the process becomes a class inclusion, rather than analogous reasoning. Therefore, the hypothesis argues that conventionality, or the frequency of use of the metaphor vehicle, is the crucial factor and component that determines the mode of processing for metaphor comprehension (Bowdle & Gentner, 2005).

The Quality of metaphor hypothesis argues that *aptness*, not *conventionality*, is the crucial factor that affects metaphor comprehension (Glucksberg, 2008). If a metaphor is especially apt, meaning the vehicle describes highly relevant features of the metaphor topic, the metaphor as a whole is quickly understood through the creation of an ad hoc category, and the metaphor is understood as a category statement (Ronderos, et al. (under review)). If the vehicle does not create a quick, categorical comprehension, analogical reasoning is activated, and the topic and vehicle are scanned for similarities.

Unfortunately, these factors are often confounded (Jones & Estes, 2006), making previous research and experiments potentially problematic, if these factors are not accounted

for. According to Glucksberg (2008, p. 76), Bowdle and Gentner (2005) varied the metaphors they used to investigate conventionality not exclusively based on conventionality, but also in aptness, making their results skewed. Jones and Estes (2006) highlight an unfortunate trend in metaphor research where researchers control for one condition but do not control for the others (p. 22), making the results muddled.

2.4.2 Concepts, Properties and Ad Hoc Categories

Ad Hoc concepts are an integral component of two intriguing theories that attempt to explain metaphor comprehension: the previously discussed Relevance Theory and the Categorization approach to metaphor. The idea of ad hoc concepts is crucial for Relevance Theory's understanding of metaphor, although it was not a part of the initial theory put forth by Sperber and Wilson (1996). According to Gibbs and Tendahl (2006), Pilkington (2000) was first in expanding the ideas of Relevance Theory by asserting that poetic metaphors are understood through the creation of *ad hoc categories*, based on the work of Carston (1996). *Ad Hoc 'categories'*, the inspiration, or basis, for ad hoc concepts, were first explained by Barsalou (1982). Psychologist Lawrence W. Barsalou (1982) performed two experiments on context dependent and context independent properties of concepts. In his second experiment Barsalou (1982) utilised two types of categories to highlight the distinction between the two properties: *common* and *ad hoc* categories. His examples for common categories are '*birds*', '*furniture*', and '*vegetables*'. Unlike the relatively concrete common categories, *ad hoc* categories are broader, vaguer, and less typical categories; Barsalou (1982) exemplifies *ad hoc* categories with the categories '*things that have a smell*', '*things that float*', and '*things that can be thrown*' (p. 88).

Ad hoc concepts are indicated typically with upper case letters, and always with an asterisk, as such: AVALANCHE*, unlike proper concepts, indicated only by upper case letters: AVALANCHE. The ad hoc concept is, according to Relevance Theory, a part of the explicature (Carston, 2016, p. 9), meaning an aspect of the proposition communicated, that is both intended and not implicated (Allott, 2010). Relevance Theory describes the ad hoc concept as part of the explicature, and part of the creation of the implicature, as such:

2.33 Explicature: FRANK_x IS AN AVALANCHE*

Implicature: FRANK_x IS CHAOTIC

FRANK_x IS IMPULSIVE

FRANK_x CAUSES DEVESTATION

FRANK_x IS UNRELIABLE

One aspect of ad hoc concepts that are both intriguing and frustrating is the ‘emergent property issue’, as discussed by Wilson and Carston (2006), and highlighted by others (Glucksberg, 2008). The ‘Emergent Property Issue’ is simply that when speakers utilise a metaphor, the vehicle of the metaphor is typically used to propose properties of the topic that is not typically, semantically, connected to the lexical meaning of the vehicle. A common metaphor used to exemplify this is the metaphor:

2.34 *Bill is a bulldozer*

In this utterance, *Bill* is the *topic* of the metaphor, and *bulldozer* the *vehicle* of the metaphor. (2.34) may be used to implicate that *Bill is stubborn*, or that *Bill is unrelenting, insensitive*, and much more. All of these meanings may be implied by (2.34), even though the lexical item *bulldozer* does not include *stubborn, unrelenting, insensitive*, in its dictionary definition. (It should be stated at this point that *bulldozer* has become lexicalised as a metaphor and the Oxford English Dictionary does include the figurative sense of a ‘bulldozer’ as a ‘forceful’ person (‘Bulldozer’, n.d.), however, it still exemplifies the gap between metaphors and their emergent properties). As in (2.34), if ‘*Frank is an avalanche*’ is used to imply that ‘Frank’ is *impulsive*, or *unreliable*, there is a disconnect between the properties applied to Frank, and the properties that the vehicle contains lexically. Metaphor comprehension must then include an explanation for this gap between the *emergent properties*, and the literal features that we assume the metaphor vehicle logically would be providing. Opposed to associative approaches, in which *bulldozer* would activate associate features, Wilson and Carston (2006) argue for a fully inferential approach. Wilson and Carston (2006) argue that the loosening of lexical definitions that takes place in metaphors also appear in other *loose-use* language, as well as in literal language itself.

2.4.3 Suppression and Enhancement

As previously stated, both Relevance Theory and the Categorization view of metaphor argues that metaphor comprehension depends on cognitive processes that elevates relevant meanings, whilst suppressing irrelevant meanings, and by so doing, creates an ad hoc category. These processes can be understood through Gernsbacher's (1991) psychological framework called the 'Structure Building Framework'.

Gernsbacher (1991) argues that both the production and processing of language are based on 'general cognitive processes and mechanisms', that are likely not exclusive to language. Gernsbacher (1991) proposes the 'Structure Building Framework' to explain the processes involved in language comprehension. The structure building framework claims that when listeners process and comprehend language, our cognitive processors build structures of mental representations. These structures can represent any meaningful unit, including both phrases and paragraphs (Gernsbacher & Faust, 1991). The structures are built through the employment of several different processes. First, the foundation of the structures is constructed. Secondly, the mental structures are developed by 'mapping' relevant information onto the structure. If new information is less coherent or relevant, substructures may potentially be built.

The components these structures and substructures are made of are called 'Memory nodes'. The memory nodes are activated by stimulus. Additional coherent information, or coherent stimuli, will either activate the established memory nodes, or related ones, which will then be added to the structure. Less coherent or incoherent stimulus will be less likely to activate either the already established memory nodes, or related memory nodes. However, these stimuli might create the foundation for new substructures.

Activated memory nodes transmit processing signals which either increase or decrease the activation of other nodes. These two cognitive mechanisms are named enhancement and suppression, respectively. Gernsbacher (1991) proposed that these processes likely occur because the information represented by the nodes are either relevant or irrelevant for further structure building.

Gernsbacher (1991) argues that suppression and enhancement are used in several different linguistic and nonlinguistic phenomena, among them what she calls 'fine-tuning' the meaning of words (p. 238). When hearing ambiguous, and even unambiguous words, several meanings are immediately activated, and subsequently either enhanced or suppressed.

However, an important question is how meanings are selected, and how irrelevant information is processed. After periods as short as 200ms, typically only one meaning will be consciously available for the processor (Gernsbacher, 1991, p. 239), meaning that irrelevant information has somehow been removed, suppressed, or simply decayed. Gernsbacher (1991) argues that contextually irrelevant features do not just decay, they are actively suppressed.

Gernsbacher et al. (2001) conducted three experiments on suppression and enhancement in metaphor comprehension. In the three experiments, metaphorical prime sentences were contrasted with a baseline prime sentence. The baseline prime sentences were either literally meaningful, nonsensical, or unrelated:

Metaphoric sentence: *That defense lawyer is a shark.*

Literal meaningful sentence: *That hammerhead is a shark.*

Nonsensical sentence: *His English notebook is a shark.*

Unrelated sentence: *That new student is a clown.*

The participants were to verify a target property sentence after reading the primes. Their results showed a slowdown in verifications when property statements relevant for basic, literal meaning were shown after metaphor primes versus the literal primes. On the other hand, verifications were quicker when the target sentence was relevant to the superordinate category and following a metaphoric priming sentence. They concluded that this shows an active suppression of irrelevant properties of the metaphor and enhancement of relevant properties. However, as Rubio-Fernandez (2007) highlights, the literal control sentences such as ‘That hammerhead is a shark’ ends in the same word as the metaphor vehicle. It is therefore likely that the control sentences will still activate the relevant metaphoric properties to some extent, and therefore undermine the strength of the results.

Glucksberg et al. (2001) did two experiments investigating metaphor property suppression and activation. The first experiment consisted of showing a metaphor vehicle, followed by a property that was either relevant or irrelevant for the metaphor such as:

Vehicle: *Shark*, Properties: *Tenacious/Swimmer*

They also used irrelevant metaphor vehicles as primes, to include scrambled versions. The participants were to indicate if a word was a real English word or not. The results of the first experiment showed a zero difference between the primed responses and the baseline

responses. They conclude that a simple lexical priming is not enough to enhance or suppress properties of metaphors. In the second experiment participants read sentences on a computer screen and had to indicate if each sentence was sensible or not. The sentences were either a metaphoric prime followed by either a metaphor relevant or irrelevant probe, or a literal prime followed by either a metaphor relevant or irrelevant probe. The results of the second experiment did show faster responses when the property sentences were metaphor relevant, showing suppression of metaphor irrelevant properties.

Based partially on the findings of these experiments, Rubio-Fernandez (2007) investigated whether the mind suppresses irrelevant meanings of metaphors when creating an ad hoc concept, whilst enhancing relevant meanings, using a Relevance Theoretical lens. Rubio-Fernandez' study was a cross-modal lexical priming study, using metaphor vehicles as primes, and measuring response times for target words. Participants heard 20 short stories, ranging from one to three sentences, where each ended in a metaphor, such as:

After six months without going to the barber, John was a lion.

After hearing the short story, a string of letters would appear on the computer screen, and the participants were to indicate if the string of letters was a real English word or a non-word. The relevant words, or targets, would either be relevant for the literal interpretation of the metaphoric sentence, relevant for the figurative meaning of the metaphor, or a scrambled target from a different short story. For the metaphor *John was a lion* the literal target, or the *superordinate* was *animal*, and the word relevant for the figurative meaning, the *distinct feature*, was *mane*. The targets appeared either 0ms after the short story, 400ms after, or 1000ms after. At 0ms and 400ms, the results were relatively similar. Furthermore, at 400ms the superordinates had slightly higher activation than distinct features. At 1000ms however, the superordinate was lower in activation than the distinctive features. Rubio-Fernandez (2007) argues that based on the findings of Rubio et al. (2003), where neutral contexts lead to superordinates maintaining activation at 1000ms, the literal meanings of the metaphoric words are being actively suppressed, as opposed to there being passive decay of meaning. Rubio-Fernandez (2007)' design is well formed, with a relatively simple design for the participants. The results are intriguing, and the design provides a solid foundation for replications.

Rubio-Fernandez (2007) did an empirical investigation of metaphor comprehension, based on the theory of ad hoc concept creation for metaphor comprehension. Rubio- Fernandez

(2007) bases her design on concept prototype studies such as Barsalou (1982). As previously mentioned, the work of psychologist Lawrence Barsalou has been crucial for the development of ad hoc concepts. Barsalou (1982) argues for concepts containing two different types of properties. These two properties are *context-independent* properties that are always activated when a word occurs, and *context-dependent* properties, which only occur in specific contexts. Barsalou's (1982) examples of context-independent properties are the words 'gills' as a contextually independent property of 'fish', and 'edible' for the word 'apple'. These properties are always activated when their relevant words are heard. Context-dependent properties are only activated when they are relevant for the context provided. As Barsalou (1982) highlights, the property 'fits in a suitcase' may very well be activated by the word 'flashlight', but most likely only if the context demands or implies it. Barsalou (1982) finds supporting evidence for his theory on concepts through two conducted experiments, and concludes:

Rather, the meaning of a word also contains weakly associated and inferable properties that are inactive in irrelevant contexts and active in relevant contexts. Given the existence of CI properties, the meaning of a word is not completely determined by context.

(p.89)

In his own brief discussion of metaphor, Barsalou (1982) mentions that 'good' metaphors are those where a relevant metaphoric property is context dependent for the topic of the metaphor, but context independent for the vehicle. It is also worth noting that similarly to how some Relevance Theorists argue that potentially all communication requires meaning adjustments (Carston, 2015), Barsalou (1982) claims that the mechanisms of context dependent and independent properties may 'be central to sentence comprehension' (p. 91).

Rubio-Fernandez (2007) investigates if metaphor comprehension is done through the enhancement of relevant properties of the metaphor, and the active suppression of irrelevant properties. The use of enhancement and suppression of properties leads to meaning adjustment creates a unique concept created for that unique utterance, the ad hoc concept. Rubio-Fernandez (2007) utilises the theories of metaphor comprehension put forth by Relevance Theory (Sperber & Wilson, 1996), Glucksberg and Keysar (1990), as well as philosopher Francois Recanati (2004) to create a thorough background for her experiment.

Recanati (2004) similarly argues for the contextual adjustment of a reference for the understanding of a concept, either through dropping or adding 'conditions of application' of

the reference. All three of these important Pragmatic theories argue for a cognitive process including enhancement or elevation of relevant metaphor properties, and a suppression of irrelevant literal properties of the metaphor. Recanati (2004) is unique in Rubio-Fernandez' (2007) discussion on theories dealing with metaphors, as Recanati (2004) argues that the primary process for metaphor processing is associative, and that inferential processes are secondary. Rubio-Fernandez (2007) argues however that although using the metaphor *John was a lion* to describe *John* having grown his hair long, i.e., he has acquired a 'mane' may seem associative. If so, it would follow that contexts that use the same vehicle but a different premise, such that John's hair makes him a lion because he is 'prideful' would needlessly also associate 'mane' even if it is not relevant based on context.

Rubio-Fernandez (2007) utilised a cross-modal lexical priming paradigm, as used by previous lexical ambiguity studies. Cross-modal simply means that the modes involved are not the same, the participants have an auditory stimulus: the metaphors, followed by a visual stimulant: a target word. The study had 60 participants, all undergraduate students. In the study, 20 proper nouns were chosen to be used metaphorically in vignettes, or short stories. After each vignette was played, a string of letters would appear on the computer screen, and the participants were to decide whether the letters were a proper word in English or not. The word was either a related superordinate of the noun, a related distinct feature of the noun, a scrambled target, or a made-up English-like word. To figure out the most appropriate superordinate and distinct feature of the proper nouns two questionnaires were made. These were based on prototype literature, and the most frequent results were chosen as the superordinates and the distinct features for the study. Some examples of these are:

Proper noun	Superordinate	Distinct feature
Cactus	Plant	Spike
Lion	Animal	Mane
Slippers	Shoe	Comfortable
Skyscraper	Building	Tall
Lullaby	Song	Sleep

Table 3 Examples of Primes and Targets from Rubio-Fernandez (2007).

An example of the vignettes is:

John doesn't like physical contact. Even his girlfriend finds it difficult to come close to him. John is a cactus.

All the vignettes ranged from one to three sentences, where the last sentence was always a nominal metaphor, with the form of 'X is a Y'. Meaning that the nominal metaphors ended in the metaphor vehicle, Y, and began with the metaphor topic, X. The contexts were created to be metaphor enhancing, making the metaphors as easily understood as possible, but still being limited to a few sentences. The vignettes were split into two equal groups. One group had superordinates and distinct features, meaning relevant targets, whilst the other group had scrambled targets from different vignettes. Two material lists were made where each group of contexts had relevant targets in one list, and irrelevant targets in the other list, and vice versa. To randomise further, another 20 metaphor contexts were made that had English-like non-words as targets. These non-word contexts were added alongside the critical contexts. To effectively randomise the study, and avoid participants seeing the same context several times, each participant was randomly assigned a Target Type, List, and Inter Stimulus Interval (ISI). The ISI was either 0ms, 400ms or 1000ms after the word-recognition point of the last word in the vignette.

The experiment was preceded by two practice trials, the first being a lexical decision task, the other consisting of sentential contexts and targets for another lexical decision. In the experiment, the participants were told that the understanding of the auditory vignettes and the response to the visual targets, were two separate, equally important, tasks. After the experiment there was a short memory test, which the participants had been informed about, to increase likelihood of attentive participants. The study utilised four-way analyses of variance (ANOVAs) using participants as the random variable.

Table 4 displays the results of the study. It shows mean reaction times of both related and unrelated superordinates and distinctive properties, at each ISIs. It also displays the facilitation, meaning the mean reaction time of the unrelated condition subtracted from the mean reaction time of the related condition. The result of the study showed minor differences between related superordinates and related distinct properties at the 0ms and the 400ms offset. However, when the superordinates were displayed at 1000ms the responses were significantly

slower than the distinct property responses, with almost 200ms in difference. Similarly, there is a 140ms difference between the related superordinate and the unrelated distinctive property.

		0ms	400ms	1000ms
Superordinate	Related	883	644	791
	Unrelated	919	698	799
	Facilitation	-36	-55	-8
Distinctive Properties	Related	740	658	598
	Unrelated	782	710	651
	Facilitation	-42	-52	-53

Table 4 Mean reaction times in milliseconds and the facilitation of reaction times from Rubio-Fernandez (2007).

The activation curves of the results can be seen in Figure 1, which are based on the calculated facilitation of the means.

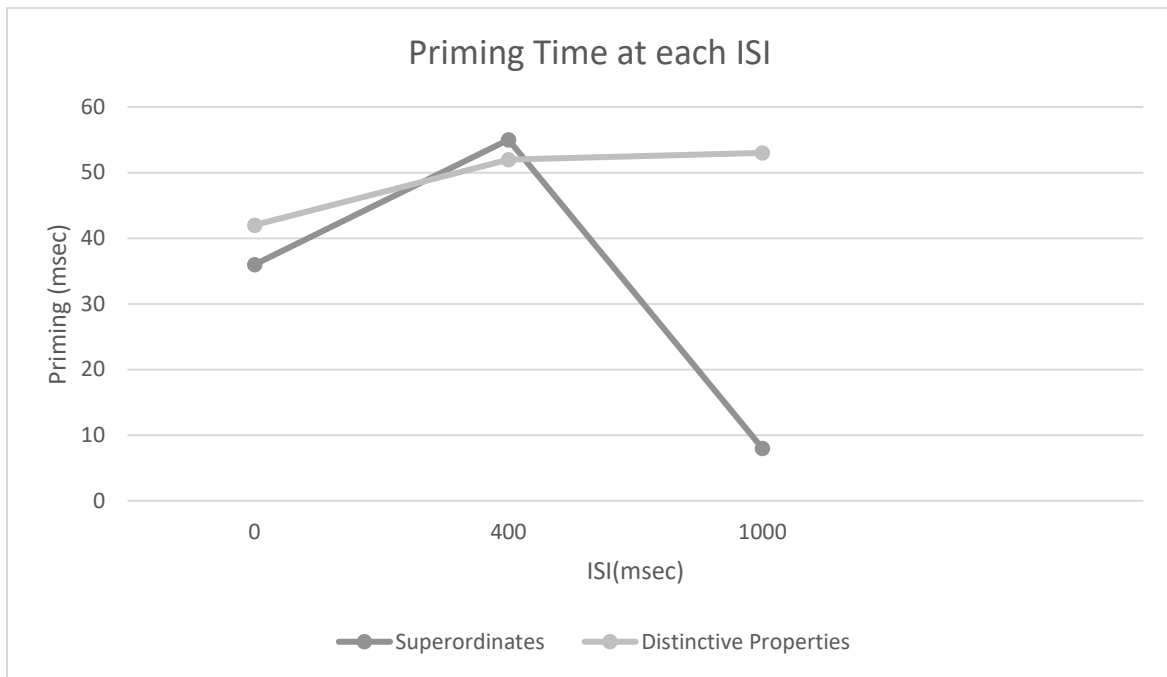


Figure 1 Activation curves of superordinates and distinctive properties in metaphoric contexts from Rubio-Fernandez (2007).

Rubio-Fernandez (2007) provides a new way of investigating metaphor processing and is therefore a welcomed addition to the experimental scene. Rubio-Fernandez (2007) utilised an older study design used for lexical ambiguity to investigate metaphor comprehension in a new way. By using cross-modal lexical priming studies to investigate metaphor- relevant or irrelevant properties the experiment is similar to the first experiment in Glucksberg et al. (2001) but it uses short vignettes to create more context for the properties. Although earlier metaphor-based experiments commonly measured reading times of the critical sentences, measuring property reading times seems like a natural way to investigate metaphor comprehension from a new angle. The study is a welcome addition to the discussion of metaphoric processing and addresses some issues with previous studies. Rubio-Fernandez (2007) highlights how Glucksberg et al. (2001) do not have zero level of activation between critical and control lines in some of their experiments, such as ‘That hammerhead is a shark’ and ‘Sharks are tenacious’. The word *shark* becomes activated in both lines. In Rubio-Fernandez (2007), the material was made such that each participant only saw crucial lines once, and only one ISI. Since ISI was a crucial aspect of the experiment, problems with pause times should be non-existent.

As with potentially every experimental study, there are issues with Rubio-Fernandez (2007). The biggest issue is addressed by Rubio-Fernandez (2007) herself, namely the small scale of the experiment, the experiment was performed with only 60 participants. Because of the small scale, the results cannot be generalised to other metaphors. I would argue that one other key issue is a number of potentially problematic metaphor vehicles and targets.

The metaphors investigated are claimed to be novel, yet three of the vehicles are defined with the relevant figurative meaning in the Oxford English Dictionary, meaning they are highly frequent metaphor vehicles. These three metaphor vehicles are *lion*, *steel*, and *sapling* ('Lion', n.d.; 'Steel', n.d.; 'Sapling', n.d.). As previously discussed, the frequency, or conventionality, of a metaphor may affect the processing required to comprehend it. The Career of Metaphor theory proposes that conventional metaphors are understood as category assertions, whilst novel metaphors demand a more complex analogical process to be understood (Bowdle & Gentner, 2005). Holyoak and Stamenković (2018) argue that if a metaphor becomes conventional enough, it may not even be recognised or processed as a metaphor at all. Furthermore, if we assume that these vehicles are recognised as metaphorical, and that these conventional vehicles are processed in the same way as other, less conventional metaphors, they may be conventional, i.e., frequently used, because they have been popularised because they are highly apt metaphoric vehicles. If these conventional metaphors are more apt than other vehicles in the experiment, then the Quality of Metaphor hypothesis would also argue that they get processed more quickly than their less apt counterparts. And to that point, I would argue that these conventional metaphor vehicles likely are processed more quickly than other vehicles in the experiment, in part because there are other vehicles that are simply generally infrequent words: *minnow*.

Three other metaphors contain compound or phrasal vehicles: *comfy chair*, *old slippers*, and *game of rugby*, which makes the point of priming difficult, and the lexical complexity potentially a lot higher than other vehicles. Two vehicles were unlike all the others by being highly physically associative metaphors, *banana* used for a boy in a long yellow jacket, and *dalmatian*, used for a boy that had spilled milkshake on himself.

Finally, I would argue that one metaphor is both unfamiliar, which is not necessarily problematic, but also inapt, and I would argue that the metaphor is therefore difficult to grasp: '*When Maria first came to England, she was very surprised that pubs closed at 11 pm. In Spain, closing time is breakfast*'. Glucksberg (2001) highlights how metaphor topics vary in their relevant attributional dimensions, and vehicles vary in the number and variety of

properties they may provide to a topic. Glucksberg further argues that although some vehicles are relatively unambiguous in their attributive qualities, such as *shark*, other vehicles can be quite ambiguous, because they ‘do not uniquely exemplify an attributive category...’ (p. 55). This highlights the possibility that vehicles do not need to simply exemplify one relevant feature for metaphoric meanings. Glucksberg goes on to argue that the properties of metaphor vehicles that are attributable to their topics are based on two requirements. These two requirements are as he states: ‘(a) the higher-order category (or categories) that the vehicle may exemplify, and (b) whether the prototypical properties of that category characterize the topic in a meaningful way’ (Glucksberg, 2001, p. 55). I would argue that the vehicle *breakfast* exemplifies strongly either other irrelevant features such as *meal*, *light meal*, or *early*, a feature of the breakfast meal that is potentially conflicting with Rubio-Fernandez’ (2007) distinct feature: *morning*. If *breakfast* activates the property of *early* alongside *morning* it provides potentially conflicting features for the understanding of this metaphor. If a pub’s closing time is *early*, this would imply early in the evening. Additionally, the feature *morning* seems to be potentially hyperbolic as a feature of Spanish pub’s closing time, making the feature less clearly available. Of course, the contextual sentences should specifically facilitate the activation of *morning*, however, the metaphoric use of the vehicle *breakfast* seems needlessly ambiguous, and less than optimal as a ‘meaningful’ property attributional vehicle to the topic of *closing time*. Furthermore, Jones and Estes’ (2006) argued requirements for metaphor aptness similarly potentially undermine the quality of the metaphor ‘*closing time is breakfast*’. Jones and Estes argue that aptness is how well a vehicle’s figurative meaning portrays a crucial feature of the metaphoric topic. The two requirements for metaphor aptness are: a salient vehicle feature for metaphor attribution, and the relevance of said salient feature. Even if we accept ‘morning’ as the most salient feature for the metaphor topic, there is still a question of relevance. I would argue that ‘morning’ is less typically relevant as a feature for a pub’s closing time than say ‘afternoon’, ‘evening’ or even ‘night’, potentially decreasing the metaphor’s aptness. However, it should be repeated that the metaphors in Rubio-Fernandez (2007) are preceded by facilitating contexts, which may potentially make these concerns invalid. To conclude, Rubio-Fernandez (2007) may include 10 potentially problematic vehicles out of a total of 20, potentially causing uncertainty for the study’s results.

Other issues, although minor ones, include not using exclusively right-handed people (at least it is not indicated that she did), seeing as the button for ‘yes’ (the word is a real word in English) was always on the right side of the keyboard, potentially giving quicker results for

right-handed people and slower for left-handed. Although the population of left-handed people is small, the small sample size of the investigation makes it even more important to know if there are confounding variables that may undermine the results. Another issue is the somewhat varied quality of the metaphoric sentences. Rubio-Fernandez (2007) does mention that the superordinate and distinct properties are based on a questionnaire. However, some primes are used in a less obvious manner than others. The metaphoric sentence:

2.35 *Her latest boyfriend is a Mercedes*

is a metaphor based on the premise of a woman being materialistic. However, the utterance may also be interpreted to be metonymic. Because of the woman's materialist attitude, the boyfriend may be metonymically referred to as his 'most important' feature, his expensive Mercedes. *Mercedes* may also invoke a form of a somewhat scalar implicature nature. *Mercedes* is supposed to invoke the property of 'expensive'. One could therefore naturally not use any car for the property of *expensive*. However, another vignette ends in the metaphor:

2.36 *The river was champagne*

This metaphor is meant to enhance the property of 'bubble', due to the river being a raging rafting river. If *Mercedes* is read, followed by the word 'expensive', *champagne* could just as easily have the assumed property of 'expensive' or 'luxurious'. This is however a minor issue, especially if the findings can be generalised for all metaphors. As mentioned, however, the small size of the study makes it impossible to generalise for other metaphors.

Another issue of the material is the difference in vignettes, that should not be necessary. The vignettes range from just one sentence, to three. As previously discussed, Ortony et al. (1978) found that when using short metaphoric contexts, metaphoric utterances were processed slower than literal utterances. However, when using longer accommodating contexts, the difference in processing speeds were removed and the metaphoric and literal utterances were processed at equal speeds. Although the differences between the shorter and longer contexts were bigger in Ortony et al. (1978) than in Rubio-Fernandez (2007), the point still stands. Some novel metaphors might warrant additional context to be properly activated but having more equal vignettes might strengthen the results.

Some vignettes also end in metaphors significantly different in complexity to others. Most metaphors have a very basic form of 'X is a Y', such as '*John was a lion*' or '*John was a cheetah*'. One vignette, exclusively, includes an adjective before the prime, '*John was a*

pair of old slippers'. The adjective might activate other properties of the prime than the chosen property '*comfortable*'. Similarly, one vignette has a metaphor that consists of a full noun phrase: '*Sometimes life can be a game of rugby*'. These differences do not necessarily need to affect the results negatively but do needlessly increase the differences between the metaphoric contexts, *potentially* causing problematic results for these specific vignettes. In their critique of Relevance Theory, Gibbs and Tendahl (2006) argue that in natural communication, the experimental phrase 'other things being equal' is seldom applicable (p. 386), and this seems to be one instance of aspects of experimental material becoming 'less equal'.

Rubio-Fernandez (2007) argues that the study's results mean that the literal connotations of the prime are being actively suppressed. Although this conclusion could solely be based on the difference of activation of superordinates and distinct properties at 1000ms, Rubio-Fernandez (2007) also mentions that a previous experiment using neutral contexts have shown superordinates to remain active at 1000ms. This previous study was done by Rubio et al. (2003). Rubio et al. (2003) also concludes that superordinates, weak features and strong features all become primed regardless of context type. Because neutral contexts cause superordinates to remain active at 1000ms, Rubio-Fernandez (2007) concludes that the literal, irrelevant, meaning of the word is actively being suppressed, supporting the Relevance Theoretic view of metaphor comprehension. As briefly mentioned earlier, at 1000ms the mean reaction time of *unrelated* distinct properties is still 141ms quicker than the reaction time of related superordinates, at 651ms. This is not addressed by Rubio-Fernandez (2007).

The study's results support the lexical modulation view of metaphor in a new way. The study's main issue being its size makes the need to replicate the experiment even greater. Similarly, a crucial aspect of Rubio's conclusion hinges on a different study, also by Rubio et al. (2003). Both Rubio-Fernandez (2007) and Rubio et al. (2003) should be replicated for a stronger, more conclusive understanding of metaphor comprehension, and feature suppression.

The extensive list of metaphor experiments has shown a wide spectrum of results. These previously mentioned metaphors are some of the many experiments that show that metaphors may be processed and understood as quickly as, or quicker than, literal language. However, this extensive list also includes experiments that show a slowdown in processing when listeners encounter metaphors, which argue that metaphors demand additional

processing effort than literal language. Additionally, there have been plenty of experiments that have shown enhancement of relevant metaphoric properties, and suppression of metaphorically irrelevant literal properties. The priming of relevant features of metaphoric vehicles has been used in several different fruitful ways. But it is worth noting that several well established important experimental works dealing with metaphors have been argued to be less than optimal. These flaws might be caused by problematic delays used in experiments, or by problematic material causing either interference or activation of additional processing requirements, potentially affecting the results of delicate response times. Additionally, as previously discussed, there has also been an unfortunate tendency to confound the processing factors that affect the quality of metaphors, and therefore also their processing.

2.4.4 Priming

In psychology, priming is the effect that occurs when one lexical item ‘activates’ different related words (McNamara, 2005). This activation means that listeners more easily, or more quickly, comprehend the related words activated by the prime word (Hutchison et al., 2013). A ‘lexical decision task’ is called the typical method for investigating semantic priming by McNamara (2005). One of the simplest lexical decision task designs would be showing participants a word, a *prime*, then showing them a second word, a *target*, which is either a real word or a non-word. The task participants are asked to do is to decide if the second word displayed, the target, is a real word or not. The results will then show if participants answer more quickly if the real words are semantically related to the prime (Neely, 1991). Experimental priming tasks are one potential way of testing the cognitive mechanisms of suppression and enhancement.

One of the most influential articles that investigated semantic priming was done by psychologists David E. Meyer and Roger W. Schaneveldt (1971), where they presented participants with word pairs, where some pairs were semantically related, such as doctor-nurse, bread-butter, and others were unrelated, such as bread-doctor and butter-nurse. Their results showed that participants responded on average 85 milliseconds quicker when the pair was related, as opposed to unrelated.

The investigation of semantic priming has been substantial the last few decades, and the effect has been used to investigate several cognitive processes, such as word recognition, sentence comprehension and knowledge representation (McNamara, 2005). Meyer and

Schaneveldt's (1971) experiment displayed a priming effect based on strictly visual targets. Cross-modal priming, meaning priming caused by different modes of stimuli such as auditorily presented words' effect on visual targets, has also been substantially investigated.

Swinney (1979) presented participants with aural sentences. The sentences, such as '*...The man was not surprised when he found several bugs in the corner of his room.*' contained either ambiguous homonyms, in this case *bugs* or an unambiguous alternative, *insect*. After the targets were heard, a string of letters was visually presented to the participants, which was either related, inappropriate, or unrelated, in this case *ant*, *spy* and *sew*. *Spy* was either ambiguously related (for bugs) or unambiguously inappropriate (for insects) depending on the context presented to the participant. The investigation revealed a facilitation of the inappropriate related meanings of the homonyms, meaning the ambiguous word *bugs* caused a priming effect on the word *spy*, even if the context was biased towards the *insect* sense of the word *bugs*, and not contextually ambiguous. However, when the presentation of the visual targets was delayed, facilitation only occurred for the appropriate targets (Swinney, 1979). Gernsbacher and Faust (1991) argue that before the delayed presentation the cognitive process of suppression had occurred, and the irrelevant meanings were at that point suppressed and less activated than the relevant meanings.

There are pitfalls when investigating semantic priming, and experimental issues that need to be avoided. McNamara (2005) presents methodological topics that present issues which must be considered when designing semantic priming experiments. Among them are the *counterbalancing of materials*, *Baselines*, and lastly *Sensitivity or bias* (McNamara, 2005).

Counterbalancing of materials when it comes to semantic priming experiments simply means that the items used in the experiments need to be observed in both the related and unrelated conditions, and it is not necessarily intuitive how this should be done. McNamara (2005) mentions that one potential solution for this could be done by presenting each subject with the items in all conditions, but argues that this should be avoided due to the repetition of items potentially causing problematic results, especially when the number of conditions, and therefore repetitions, increases. The standard solution then becomes alternating conditions across participants, ensuring that each participant observes an item only once, but that every item is observed in all conditions across the total of participants.

It is also important to counterbalance material in terms of confounding factors. McNamara (2005) asserts the need for carefully designed material, and includes several factors that may affect semantic priming, including associative strength, semantic relatedness, type of semantic relatedness, and word frequency (p. 54). Additionally, potential problems may be caused by reaction times being quicker for words without common synonyms such as *milk* (Pecher, 2001).

Due to the relative nature of semantic priming, the targets must be tested against neutral, or baseline, materials that are neither related or unrelated to the prime. There is a distinction between whether related primes are facilitating reaction times, or unrelated primes are inhibiting reaction times. A neutral baseline helps this distinction become clear when comparing the different conditions. An outdated way to include a baseline condition is to use words such as *blank* or *neutral* (McNamara, 2005, p. 59). However, this may confound the results as the baseline targets are repeated whilst the related and unrelated targets are not. McNamara (2005) concludes that the best current solution for a baseline condition is using ‘orthographically regular, pronounceable non-words’ (p. 60) and that these baseline targets should not be repeated. Finally, McNamara (2005) highlights how performance in priming experiments is also influenced by *sensitivity* and *bias*, two terms from *signal detection theory*. Sensitivity is essentially how easy it is to detect the presence of a stimulus, whilst bias is the likelihood of one response from another.

McNamara (2005) argues that if there are uncertainties about an experiment's results, there is one crucial solution available: *If an experimenter is concerned about the generality of his or her findings, then by far the best strategy is to replicate the experiment with new subjects and new materials* (McNamara, 2005, p. 58).

2. 5 The Replication Crisis

In recent years the ‘*replication crisis*’ has become an important aspect of all empirical scientific research. The replication crisis is a term used to describe a discovery in the scientific community that when experiments are replicated, the replication results will often either disprove, or be unable to prove the original experimental results. One of the most famous articles dealing with the crisis is Ioannidis (2005), who critically argued that ‘Most published research findings are false’. Sönning and Werner (2021) highlights four ‘focal problems’ as the main issues that must be tackled in scientific research:

1. *A lack of transparency in methodology and data analysis.*
2. *The non-reproducibility of scholarly work, as, for example, original data and analysis procedures are not accessible.*
3. *Reluctance to undertake replication studies as purportedly “unoriginal” (and unprestigious) despite their potential to put previous findings in perspective.*
4. *Concerns about high rates of false-positive findings in the published scientific literature.*

(p. 1182)

The crisis is in part intensified by researchers who solely focus on attaining statistical significance, and by scientific journals who disregard research that has not obtained statistical significance (Green, 2021). According to Green (2021) this has led to a trend in scientific journals to publish both research that erroneously rejects the null hypothesis and publish research that fails to reject the null hypothesis when it should have. Although the replication crisis has received a tremendous amount of focus in recent years, the critique of scientific ‘dependency’ on achieving statistical significance is not new. During the 1960’s there were several researchers who argued for the improvement of experimental and statistical work, among them Jum Nunnally. Nunnally (1960) states:

Even so, the emphasis on the null-hypothesis models is unfortunate. As is well recognized, the mere rejection of a null hypothesis provides only meager information. For example, to say that a correlation is “significantly” different from zero provides almost no information about the relationship. Some would argue that finding “significance”; is only the first step, but how many psychologists ever go beyond this first step?

(p. 643)

The replication crisis has been especially prevalent and discussed in scientific fields such as medicine, biology, genetics and psychology. However, the replicability issues affect all scientific fields to a certain extent (Schooler, 2014). Psychology has especially struggled with discoveries of fragile experimental results. Schooler (2014) argues however that the high focus on dealing with the replication crisis within psychology is only to be celebrated; psychology has in recent years proven its willingness to tackle its shortcomings empirically.

In psychology the replication crisis supposedly began in force in 2011, caused by the aftermath of a published research article in *Journal of Personality and Social Psychology* dealing with precognition. Because of the presumed unscientific nature of precognition, a vast amount of replication research was done, but was not accepted by the journal, as it only published ‘original’ work (Green, 2021). The twofold blow to scientific integrity by the journal, who had provided a platform to problematic scientific research whilst denying those who wished to disprove it undermined the scientific reputation of experimental psychology (Green 2021). Although the replications were denied for being unoriginal, one response was eventually accepted by the journal: Wagenmakers et al. (2011). The paper highlights the problematic nature of relying on null hypothesis significance testing, as had been endemic to the field of Psychology (Wagenmakers et al., 2011). Wagenmakers et al. (2011) displays how the precognition article may pass the popularised scientific convention of significance: $p < .05$ but becomes immediately weakened by utilising Bayesian analysis.

Subsequently, the ‘Reproducibility Project’ was created, bringing together 270 researchers dedicated to replicate 100 established psychological experiments. Out of the 100 original studies, ninety-seven percent had significant results. However, from the project’s replications only thirty-six percent had achieved statistical significance. Additionally, the mean effect size of the replication effects was half that of the original effects (Open Science Collaboration, 2015). After the publication of the project’s devastating results many researchers critiqued the project, questioning the motives of those who performed the replications. Additionally, many highlighted the inadequacy of trusting *one* scientific replication whilst disregarding the original study, seeing as either one could be the problematic one (Green, 2021). This underlines the problematic trends within scientific publications, since it will always be possible that multiple failed replications have either not been attempted to be published by the researchers, or not been accepted for publication by journals.

Clearly, psychology has replication issues, but is generally committed as a field to empirically deal with those issues. Linguistics, as almost every other field, is also affected by the replication crisis. Within linguistics, Grieve (2021) argues that replication failures are likely due to ‘inherent issues’ with experimental methods when studying something as socially complex as language (p. 1343). He goes on to argue that linguistic experimental research will frequently fail to replicate, regardless of research practices, because: *language is an inextricably social phenomenon, making it impossible for linguists to fully control social*

context across independent replications (Grieve, 2021, p. 1344). The claim by Grieve (2021) is in part supported by Roettger (2021a). Roettger (2021a) highlights how it is impossible to determine if failed replication attempts fail due to a false discovery in the replicated study, or if the sample or context is different enough to cause a failed replication.

There are clearly several issues with experimental practices and within scientific journals. Multiple researchers have highlighted the many issues that reduce the quality of scientific research, and with these many issues, all the procedural changes that the scientific community should utilise to improve. Roettger (2021a) claims that linguistic researchers would minimise the chances of failed replication by taking these contextual failings highlighted by Grieve (2021) into account. He further argues that three key strategies should be utilised to minimise failed replication due to context sensitivity. These three strategies are: 1. explicitly describe any potential boundary conditions for generalisations, 2. provide detailed description of the sampled population and any contextual factor that may have affected the results, and lastly 3. utilise more conservative statistical models to their data (Roettger, 2021a). Vasisth et al. (2018) highlights four proposals they argue would be a significant improvement to the general scientific procedures in experimental work:

Researchers should (i) move their focus away from statistical significance and attend instead to increasing the precision of their estimates (e.g., by increasing sample size, or improving the quality of measurements, or designing stronger manipulations); (ii) carry out direct (not just conceptual) replications in order to demonstrate the existence of an effect; (iii) pre-register their designs and planned analyses and deposit them in venues like osf.io and aspredicted.org; and (iv) release their data and code upon publication. Journals can encourage these practices by favoring pre-registered analyses, introducing a short-article type featuring direct replications, and mandating open data and code release upon publication.

(p. 167)

Although all four of the proposals would cause improvements for the integrity of scientific research, one of Vasisth et al. (2018)'s proposals is repeatedly mentioned as a relatively simple procedure that will greatly improve scientific empirical work: preregistrations (Schooler, 2014; Roettger, 2021b). Preregistration increases the incentive for transparency, replications, as well as reduce biases and flexibilities in experiments that may lead to researchers adjusting their hypothesis after having gathered data and frame the adjusted

hypothesis as the initial hypothesis (Roettger, 2021b). Preregistrations can be either a simple or a highly detailed, time-stamped document that includes the researcher's intent for data collection as well as how the researchers intend to analyse said data (Roettger, 2021b). Roettger (2021b) argues that exploratory observations that are framed as if they were predicted prior to data collection may cause overconfidence in results. Preregistration would then reduce the chances of publication bias, meaning that only confirming studies are published, as well increase the transparency in what was predicted and what was explored impromptu during the research process.

The replication crisis has made it abundantly clear that scientific research is all too rarely replicated. Additionally, and even more damning, when they are replicated the results will more likely than not be conflicting with the original studies (Ioannidis, 2005). These facts jeopardise the trustworthiness of all scientific research. Within psycholinguistics and pragmatics, figurative language represents one of the fundamental ways that contextual use affects the semantic meaning intended by speakers, among many others. The replication of solid experiments dealing with figurative language improves the statistical probability of definite answers for the fundamental questions of an entire interdisciplinary field and should be pursued. The experiment replicated in this thesis, Rubio Fernandez (2007), is a well-designed experiment, but unfortunately consisted of a low number of participants. Replicating the experiment will itself be important for the solidification of intriguing findings, especially when the replication, as it is in this thesis, includes a substantially larger number of participants.

3. Experiment

3.1 Introduction

Metaphoric language is not unique and relies on similar cognitive processes as literal language. Cognitive theories from both linguistics and psychology propose that language comprehension and processing require lexical modulation. Relevance Theory proposes that metaphor processing requires ‘meaning adjustment’, by creating an ad hoc concept, either through narrowing, broadening, or both, of the concept being adjusted (Carston, 2015; Noveck et al., 2001). The Categorization approach from psycholinguistics argues that metaphor processing happens through the creation of an ad hoc category by suppressing and enhancing features of the metaphor vehicle (Gernsbacher et al., 2001; Glucksberg et al., 2001). The results of experiments such as Rubio-Fernandez (2007), Gernsbacher et al., (2001), and Glucksberg et al., (2001) support these theories, indicating that the features of metaphor vehicles are suppressed and enhanced depending on contextual relevance. However, the replication crisis has revealed the need for replications to ensure the quality and validity of studies. Rubio-Fernandez (2007) was one of the studies that has found supporting evidence for this theory, but had too few participants to generalise her results, and as previously stated, McNamara (2005) claims that: *If an experimenter is concerned about the generality of his or her findings, then by far the best strategy is to replicate the experiment with new subjects and new materials* (p. 58). In addition to replicating Rubio-Fernandez (2007), before the main experimental run, a preregistration was created on the website *osf.io*. This preregistration explained the main research question, the experiment, and the intended statistical analyses of the results. Preregistrations are one way of increasing experimental transparency and scientific integrity.

3.2 Method

The experiment done in this paper is a close replication of Rubio-Fernandez (2007), with minor adjustments. Most importantly, the material is new. Additionally, one condition has been removed, the 400ms Inter-stimulus Interval (ISI), i.e., the delay between stimuli. Although Rubio-Fernandez (2007) wanted to test when metaphor priming was active with three different interval conditions, this experiment exclusively tests the relevant results of

Rubio-Fernandez (2007), namely the primed activity recorded at 0ms and 1000ms. The 0ms ISI condition also had to be slightly changed due to limitations of the program used to code the experiment, but the change was insignificant. The change was done because the program did not recognise 0 milliseconds as a valid number, and the ISI had to be changed from 0 to 1 millisecond.

3.2.1 Participants.

The experiment reported here had an initial pilot run performed with 50 participants. The results were promising, and funding was granted by the University of Oslo for an experiment with a substantially larger sample size. The final, main experiment had a total of 197 participants.

The participants were found through the website *prolific.co*, where research participants can be selected based on requirements, and the participants are compensated fairly. The requirements for participating were that the participants had to be native British English speakers, right-handed and neurotypical. Additionally, participants were between the ages of 18-35. Information on gender was not collected.

When analysing the data all participants with an accuracy lower than 90% were excluded from the analysis. This exclusion was calculated based on responses to both the critical target words (which were always real words of English) and the filler targets, which were nonsensical words. This left 164 participants with acceptable accuracy after the exclusion. The results are based on the response times of these 164 participants. Additionally, all response times quicker than 250 ms and slower than 2500 ms were removed. The number of required participants was decided based on a power analysis by simulations performed using the R package SimR based on existing pilot data. We conducted 1000 simulations to determine the minimum number of participants needed to detect the predicted three-way interaction, with a predicted effect size smaller than the one found in the pilot study. The experiment was done online, and each attempt took circa 10 minutes.

3.2.2 Design

This study is a quantitative experimental design, meaning one or more variables are systematically and deliberately manipulated by the researcher (Rasinger, 2013, s. 41). The

study design was a within-subject design, meaning all participants were exposed to all variables and conditions. The experiment consisted of three variables, with two conditions each, for a total of eight conditions (2x2x2). The variables are Relatedness, Target type and ISI. The ‘relatedness’ variable consisted of two conditions, unrelated and related targets. The relatedness was based on whether or not the displayed word was either literally or figuratively related to the metaphor vehicle. All items functioned both as related targets, and as unrelated controls for other contexts. The ‘target type’ variable’s conditions were if the feature was literal or figurative, primed by the vehicle. Finally, the inter-stimulus interval, the ‘ISI’, variable consisted of the two conditions; 1ms and 1000ms, i.e., the time delay that was used between the aurally presented metaphor vehicle and the visually displayed string of letters. The participants were assigned to one of eight lists, and were exposed to target words in all conditions, i.e., repeated measures. The eight lists were created so that each metaphor was seen by participants with every possible condition. Meaning, the metaphors, for example *This store is a jungle*, was seen with two related and two unrelated targets, these related and unrelated targets were both figurative and literal, and all targets were shown at both 1ms and 1000ms, for a total of eight lists, as exemplified in Table 6, displaying each target shown for the metaphor ‘*This store is a jungle*’:

***I was done shopping an hour ago, and have been trying to get out since.
This store is a jungle.***

Related Literal Feature	Forest 1ms	Forest 1000ms
Related Figurative Feature	Confusing 1ms	Confusing 1000ms
Unrelated Literal Feature	Pill 1ms	Pill 1000ms
Unrelated Figurative Feature	Aid 1ms	Aid 1000ms

Table 5 All possible targets shown for the metaphor ‘This store is a jungle’.

The lists were designed such that all participants saw an equal number of targets with each condition.

3.2.3 Materials

As mentioned, the crucial difference between this experiment and Rubio-Fernandez' (2007) is the material. The material is new and intended to be an improvement of Rubio-Fernandez' (2007) material.

The metaphors in this study were all chosen from Jones & Estes' 2006 study on metaphor conventionality and aptness. All metaphors were nominal, i.e., *X is a Y*. In Jones & Estes' (2006) study they performed three experiments. In the experiments they manipulated the conventionality and aptness of metaphors individually, to investigate which affects metaphor comprehension the most. They conclude their investigation stating that aptness 'predicted the preference for metaphors over similes ..., the speed and ease of metaphor comprehension..., and the category membership of metaphorical terms ...' (p. 18). Crucially for the experiment presented in this paper, Jones & Estes (2006) made participants rate 100 pairs of metaphors, where each metaphor was displayed with either a high apt or a low apt metaphor topic. Meaning, a total of 200 metaphoric sentences were rated on a scale of one to seven, where one means the sentence is not an apt metaphor at all, whilst seven means the sentence is a highly apt metaphor. They also controlled for conventionality by having one group of participants create properties associated with the vehicle concepts of the metaphors, then having another group of participants rate the conventionality of the total. A minor excerpt of their metaphors with rated conventionality and aptness can be seen in Table 7.

Conventionalit y rating	High apt metaphor	Aptness rating	Low apt metaphor	Aptness rating
1.79	A business is a living organism	3.76	A kitchen is a living organism	2.85
2.64	Some stomachs are barrels	4.70	Some bladders are barrels	3.59
2.97	Research is mountain climbing	4.27	Grading is mountain climbing	2.90

Table 6 Excerpt from Jones and Estes 2006 appendix of metaphoric sentences and their ratings.

24 metaphors were chosen from Jones & Estes' (2006) study. Jones & Estes (2006) argue that the key processing factor for metaphors are their aptness. Therefore, aptness was the key factor for choosing metaphors, and each metaphor chosen for this experiment had an aptness rating above 4.

In an attempt to increase the likelihood of 'all other things being equal' and therefore reduce potentially confounding factors in the experiment, some metaphors were avoided due to their 'lexical complexity'. Based on the potential chance that priming may be affected by compound or phrasal vehicles, those metaphors that ended in a compound or phrasal vehicle were discarded as possible options. Meaning, metaphor vehicles that included adjectives, adverbs or compound words were avoided as best as possible.

The chosen metaphors were chosen based on having a high aptness score, and a low conventionality score, i.e., the metaphors should be as good and as novel as possible. Due to the aforementioned restrictions, i.e., the metaphors must be highly apt single noun vehicles, some of the chosen metaphors had a less than optimal conventionality rating. However, only one metaphor was above a conventionality rating of 5, at 5.03. The mean and median of the conventionality ratings were 4.00 and 4.11, respectively. Therefore, from the three examples shown in Table 7, only the second high apt metaphor was chosen, as the first metaphor did not have an aptness rating of 4, and the third example consisted of a compound word, i.e.,

mountain climbing. There were some minor grammatical changes made to some of the metaphors chosen, but no differences were made to the lexical aspect of the topic of the metaphors, and no changes were made to the metaphor vehicles at all (beyond minor grammatical inflections). An example of a change made to a metaphor can be seen with the chosen metaphor from Table 2, as shown in (3.2).

3.1 Some stomachs are barrels

3.2 His stomach is a barrel

This change was made to avoid any potential cognitive effects caused by the enrichment of the scalar term *some* and to avoid the general vagueness of the term, as well as making contexts easier to produce by making the metaphors less vague.

For each metaphor two related features were created. One feature was created based on some literal aspect of the word used as a metaphor vehicle, the other feature based on the figurative aspect of the metaphor vehicle. These features were chosen based on word frequency to minimise the differences in frequency of use between the literal and figurative features, hopefully reducing any serious differences in ease of activation. Meaning, a highly frequent, common word such as ‘forest’ was not paired with a highly infrequent, uncommon word, for example ‘discombobulating’. As seen in Table 8, the features created for the metaphor ‘this store is a jungle’ was ‘forest’, related to the literal meaning of the metaphor, and ‘confusing’, relevant for the intended figurative meaning of the metaphor. There was also an attempt to minimise differences in syllables in the target words. An excerpt of the created metaphoric contexts, the metaphors and their features can be seen in Table 7.

Critical Contexts	Literal Feature	Figurative Feature
1. I was done shopping an hour ago and have been trying to get out since. This store is a jungle.	Forest	Confusing
2. Throughout her rehabilitation Hannah has gotten all the support she has needed. Her husband is a gem.	Stone	Precious
3. I cannot help talking to the mail carrier every morning, his smile is a magnet.	Metal	Attractive

Table 7 Extract of metaphoric contexts with their literal and figurative features.

After features had been created and checked for word frequency, the calculated average of the frequencies was used to pair the metaphors to create scrambled targets, with as close frequency as possible to the original, related targets. Some pairs were changed to avoid scrambled targets that could potentially be coincidentally primed by the irrelevant metaphor, e.g., the metaphor vehicle *butterfly* was originally paired with the vehicle *flower*. However, *flower* had the critical target *delicate*, which would potentially be primed by the scrambled vehicle *butterfly*.

Each metaphor was given a short context, meant to strengthen the metaphoric meaning, as opposed to the literal meaning. In addition to these critical contexts, 24 additional contexts were made as filler contexts. As mentioned, Eight Lists were created, to ensure that each item was seen by participants with every condition. Therefore, every item was seen with the literal feature at both ISI's, the figurative feature at both ISI's, as well as with the scrambled literal and figurative features at both ISI's. Every list contained the same amount of literal, figurative and scrambled targets, as well as the same amount of both ISI's. Each participant was assigned randomly to one list.

The contexts were recorded by a female speaker with a native Estuary English accent. The recordings were edited to remove any unwanted silence before and after the contexts. Before being published to *prolific.co* for experimental runs, the experiment had two minor pilot runs. The first pilot run revealed problems for some participants due to programming errors. The issues were fixed, and the experiment had a second, successful minor pilot run before being published.

3.2.3 Apparatus

Due to the digital nature of the experiment, the participants were able to take the experiment at their own behest, where and when they themselves would prefer. The participants were told that headsets were needed for the experiment, to reduce noise and interference. The participants were asked to keep their index fingers on the two relevant buttons throughout the experiment, namely their left index finger on the ‘F’ button, and their right index finger on the ‘J’ button. The ‘F’ button indicated that the string of letters was not a real English word, whilst the ‘J’ button was used to indicate that the string was a real English word. The visual targets appeared in the middle of the screen in capital letters and stayed on the screen until a decision was made. After the participant had made a decision there was a 1000ms delay before the next audio item began playing.

3.2.4 Procedure.

The participants were told they were participating in an experiment on language comprehension. They were told that they would be listening to short ‘stories’ consisting of a few sentences, and that the story would end in a metaphor. After each metaphor a word would appear, and they were to decide whether the word was a real English word or not, by pressing the aforementioned computer keys. The participants were told to do this as accurately and as quickly as possible. The participants were told that comprehending the short stories was equally as important as the lexical decision task. To increase the likelihood that the participants had their full attention on the metaphors they were told that there would *potentially* be a short memory test after the experiment, however, there were none. To assist the written instructions there was one example context before the main experiment, to display what exactly they would be doing. In the computer program the experiment was coded in,

PCIBex, it was possible to completely randomise what order the items were displayed in, so the participants who were assigned to the same group did not see the items in the same order.

3.3 Results

3.3.1 Results of Pilot Run

The main results of the pilot run are shown in Table 8 and Figure 2. The priming results, i.e., the facilitations of all conditions are displayed in Table 8, which also includes the Standard Error and the Confidence Intervals. The priming results are also displayed in Figure 2.

Feature	ISI	PrimeTime	SE	CI
Figurative	1	17.34	58.67	121.36
Figurative	1000	51.53	52.18	107.94
Literal	1	30.81	79.80	168.37
Literal	1000	-99.93	64.41	134.36

Table 8 The priming results of the pilot run, including the Standard Errors and Confidence Intervals.

Figure 2 displays the priming results at both ISIs, including Standard Error bars, displaying the standard deviation of means within the group of participants.

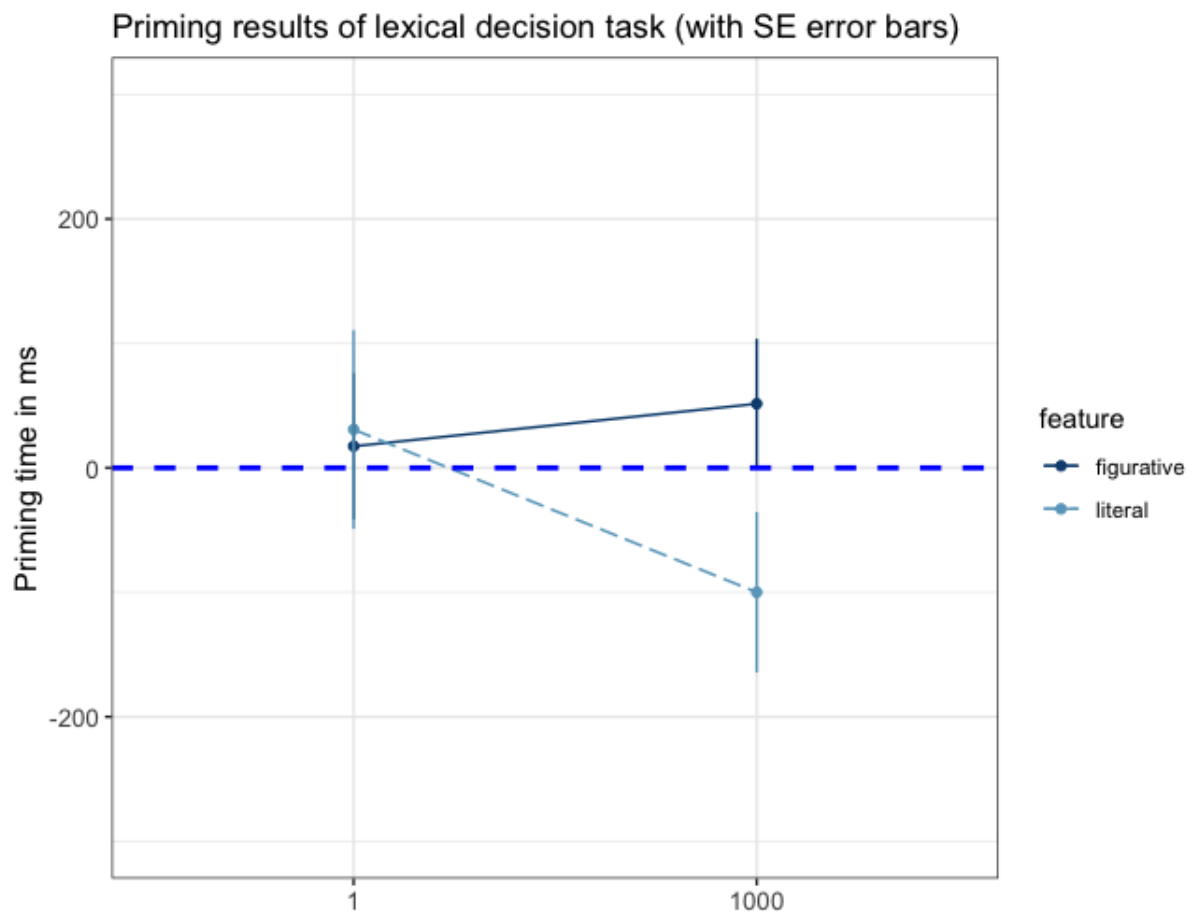


Figure 2 Prime time results of literal and figurative features at both ISIs of the pilot run.

3.3.2 Results of Main Experiment

The means of the reaction times of all conditions of the main experiment is displayed in Table 9. In addition, the table displays the calculated facilitations, meaning the differences between the related and unrelated conditions of the experiment at both ISI's.

		ISI	
Target	Relatedness	1	1000
Literal	Related	1036.91	942.63
	Unrelated	1057.67	907.85
	Facilitation	-20.76	34.78
Figurative	Related	1107.31	929.56
	Unrelated	1059.20	939.83
	Facilitation	48.11	-10.27

Table 9 Mean reaction times (in milliseconds) for each condition of the experiment, and the facilitation of each condition.

The Priming results of the main experiment is shown in Table 10, alongside the Standard Error and the Confidence Interval. The priming results, or prime times, are calculated by calculating the means of each item, then subtracting the mean of the item as a related target from the mean of the item as an unrelated target. This therefore shows the difference in activation of target types based on whether the items are related or unrelated targets.

Feature	ISI	PrimeTime	SE	CI
Figurative	1	-48.11	31.35	64.84
Figurative	1000	10.27	31.91	66.01
Literal	1	22.67	34.51	72.81
Literal	1000	-38.19	35.01	73.02

Table 10 Prime Time numbers, including Standard Errors and Confidence Intervals

The results from Table 10 can be seen plotted in Figure 4. This includes the prime times and Standard Error bars. The Standard error bars are the priming results minus and plus the Standard Error, showing the variability of gathered data.

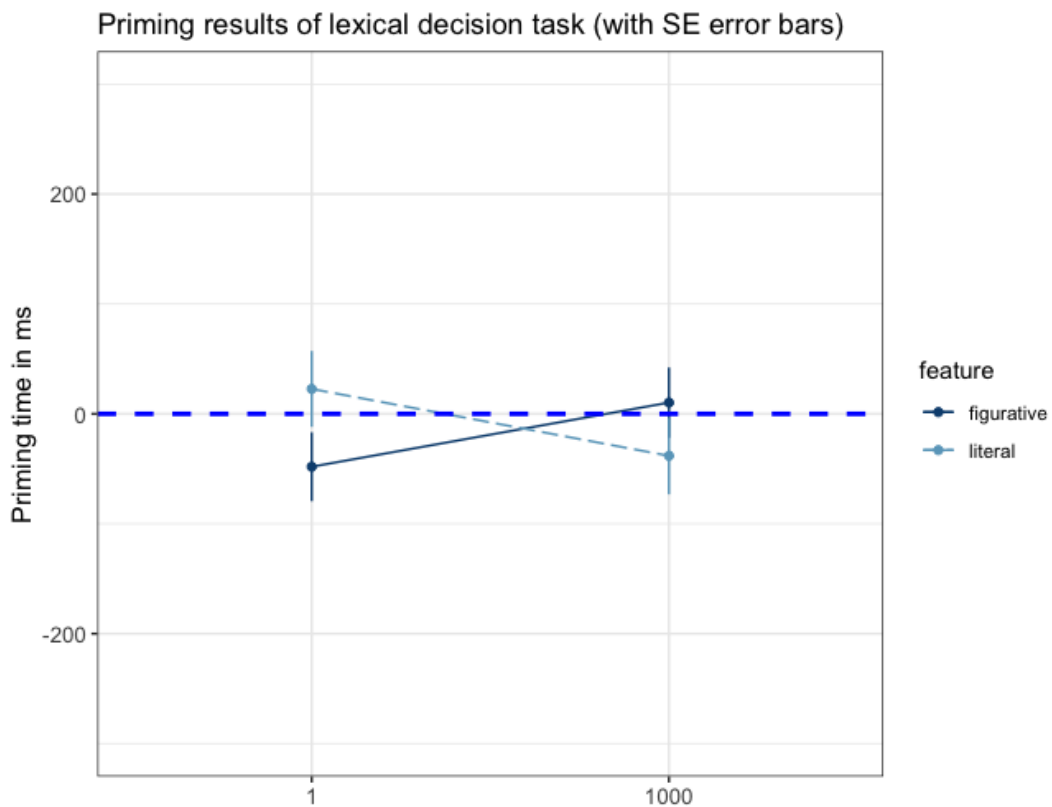


Figure 4 Prime time results of literal and figurative features at both ISIs of the final experiment.

3.4 Discussion

The results of both the pilot run and the main experimental run supports the previous experimental findings that indicate metaphoric processing to include a suppression of irrelevant literal features, and an enhancement of the relevant metaphoric features (Rubio-Fernandez, 2007; Gernsbacher et al., 2001; Glucksberg et al., 2001). Figure 4 shows that at 1000ms the figurative features have an increased activation, whilst the literal features have a decreased activation. The results of both runs of the experiment show an increased activation of the figurative features at 1000ms, and a decrease in activation of the literal features. Although the main experiment shows a smaller difference in activation than the pilot run, it also has smaller Standard Error bars, making the activation results less varied.

Based only on the means of the condition (Table 9), when the targets were shown immediately (1ms ISI) there is a slight preference for the related literal condition, compared to all other conditions, 1037ms versus 1058ms, 1107ms, 1059ms, respectively. The related figurative condition also stands out as noticeably slower than the others at 1107ms. However, after a 1000ms delay, the related figurative condition has a quicker mean response time than both the figurative unrelated and the literal related conditions, with a mean of 929ms, versus 940ms and 943ms, respectively. Interestingly, without any further calculations than the mean, the literal unrelated condition also has a substantially quicker response time, at 908ms. Comparing solely the raw means, this pattern is similar but not identical to Rubio-Fernandez (2007). Rubio-Fernandez' (2007) means, shown in Table 5, shows quicker reaction times for distinctive properties (figuratively related targets) than superordinates (literally related targets) at 0ms, for both the related and unrelated conditions, 740 ms and 782ms versus 883ms and 919ms, respectively. Unlike the results from this paper, Rubio-Fernandez (2007) reports slower reaction times for *both* the related and unrelated literal conditions, than the related and unrelated figurative conditions at 1000ms, 791ms and 799ms versus 598ms and 651ms, respectively. However, when calculating the facilitation of the means, the similar pattern of results becomes clearer, as seen in Figure 1 and Figure 4.

As Table 10 and Figure 4 show, the prime time results reveal a much higher activation of figurative features at 1000ms than the literal features, 10.27ms versus -38.19ms, respectively. We can additionally see an inverse result at 1ms, the literal features have a substantially higher activation immediately after the vehicle than the figurative features have, 22.67ms versus -48.11ms, respectively.

Interestingly, the reaction times recorded in this study, presented in Table 9, is overall noticeably slower than the reaction times recorded by Rubio-Fernandez (2007), presented in Table 4. The highest mean in Rubio-Fernandez (2007), 919ms, is only slightly higher than the lowest mean calculated in this study at 908ms. This may be partially due to when the targets were presented. Rubio-Fernandez (2007) presented the targets 0ms, 400ms, 1000ms after a chosen 'acoustic signal': 'For each of the 20 nouns in our materials a point was selected where the prime would be unequivocally recognized. Targets were presented visually at the end of the acoustic signal 0, 400 or 1000ms after the word-recognition point selected for each prime' (p. 355). In other words, the targets were presented before the entire metaphor vehicle was heard, based on words being comprehended before the entirety of the word has been heard. In the experiment conducted in this thesis, the targets were simply shown after the vehicles had been heard in their entirety. This difference in when the ISI's begin may explain the overall differences in means between this study and Rubio-Fernandez' (2007). However, one can hypothesise a bit further: Rubio-Fernandez (2007) does not specify exactly when this acoustic signal was. However, *if* the point chosen was typically about 100ms after the point when lexical and semantic information starts being retrieved when a word is heard, 200ms after the onset of the word (Hauk et al., 2012), this could explain why the overall mean reaction times collected by Rubio-Fernandez was around 300ms quicker than the mean reaction times collected in this experiment. Finally, the targets in Rubio-Fernandez (2007) were chosen based on questionnaires and may therefore be more connected to the vehicles than the features chosen in this study. This may also be part of why the features in this study had overall slower reaction times than the targets in Rubio-Fernandez (2007).

Unfortunately, the result is not as powerful as hoped, due to the high number of participants that were removed because of low accuracy. Out of the 197 participants that did the experiment, only 164 participants had above 90% accuracy. This high inaccuracy is somewhat expected due to the unfortunate nature of online experiments, taken when and where the participants chose. It is unlikely that all participants will be as diligent and focused at home, at a café, or at the library, as they would have been in an experimental setting. A simple improvement that could be made for future replications could be using an experimental setting, for increased likelihood of attentive participants.

Furthermore, due to time limitations, there are possible improvements to be made to the material. Although the metaphors were tested and rated by Jones and Estes' (2006) study, some of the features matched with the vehicles could be improved. By using literal and

figurative *features* of the vehicle, the spectrum of possible valid targets is broad. In contrast, Rubio-Fernandez (2007) used superordinates and polled distinct features of the vehicle, which limits the possible options. The literal and figurative features used in the experiment were intended to be contextually activated figurative features, and contextually independent activated literal targets. Some of the figurative features were however likely more literally connected to the vehicle than they optimally should. The metaphor ‘Her husband is a gem’ had the literal feature ‘stone’ and the figurative feature ‘precious’. ‘Precious’ is however part of the common description of ‘gems’, i.e., ‘Gems’ are ‘precious stones’. However, the literal feature ‘stones’ should still be suppressed as an irrelevant feature. ‘Precious’ will just likely also be activated as a literal aspect of the concept, at least immediately. Similarly, the figurative feature ‘aid’ is likely activated as a literal aspect of ‘medicine’, from the metaphor ‘music is medicine’, but again, the literal feature ‘pill’ should still be more suppressed than ‘aid’ as contextually irrelevant. Using questionnaires to find and test the most appropriate literal and figurative features, as Rubio-Fernandez (2007) did, is one way this experiment could have been improved, if the time and resources for more extensive questionnaire and experimental work had been available.

Although there were fewer acceptably accurate results than hoped, the priming results still display an initial preference for the literal features, but a substantial difference between the literal and the figurative features after a 1000ms delay, in favour of the figurative.

4. General Discussion

The presented experiment supports previous studies and theoretical predictions that metaphor processing involves suppressing contextually irrelevant literal features of the metaphor vehicle and an enhancement of contextually relevant figurative features. Specifically, both Relevance Theory and the Categorization Approach to metaphors predict that metaphor processing is done through lexical modulation, creating Ad Hoc categories by suppression and enhancement (Carston, 2002; Glucksberg, 2001). The study has shown that Rubio-Fernandez (2007) can be replicated with new material and new participants and produce similar, supporting results.

Throughout the thesis, several highly cited studies have been presented and discussed, that have provided valuable insights and empirical results for the investigation of figurative language. It is however an unfortunate reality of complex cognitive experiments that many of these studies are routinely, and somewhat justifiably, critiqued. And many studies do have potential room for improvement. As previously discussed, some could improve through better experimental design, such as Ortony et al., (1978), but more typically, many could improve through better material, such as Blasko & Connine, (1993), Noveck et al., (2001) Glucksberg et al., (2001), and Rubio-Fernandez, (2007). Replicating these experiments with improved designs or material would be a substantial improvement for the pragmatic and psycholinguistic theories on language comprehension.

McNamara (2005) argues that one of the potential reasons for why semantic priming has been such an attractive scientific method is that participants of semantic priming experiments may not be aware of the processes occurring (p. 5). This may potentially reduce the ‘Observer’s Paradox’, a term from sociolinguistics that explains how observed participants may be unconsciously influenced in an experimental environment, *by* the observation of an experimenter. This may cause a negative effect on the results of the observation (Labov, 1972). Interestingly, the experiment done in this paper was given to a friend for feedback, said friend then provided the feedback that maybe participants should be told that the words appearing on the screen would not be related to the aurally presented sentences, even though a substantial portion of the words he saw were related (disregarding non-words).

Finally, this experiment replicated the results of Rubio-Fernandez (2007), which used nominal metaphors. Nominal metaphors are metaphors where nouns are used as the

metaphoric vehicle, such as ‘X is a Y’. These types of metaphors have been the main source for experimental investigations in psycholinguistic investigations of metaphor. *Verbal*, or *predicate*, metaphors are when verbs are used metaphorically. Currently, verbal metaphors are underrepresented in the psycholinguistic investigation of metaphors (Holyoak & Stamenković, 2018; Ronderos et al. (Under Review)). A simple example of a verbal metaphor is the utterance:

3.3 That dog *flew* past us.

when used for example, to describe a dog running very quickly past the speaker. Although replicating previous experiments on nominal metaphors is an important endeavour, there is an unfortunate lack of investigations on verbal metaphors that must be rectified.

5. Conclusion

This study investigated and explained some of the most important cognitive theories attempting to explain metaphors and how they are processed. The main goal of the thesis was to test one theoretical prediction of metaphor processing put forth by several different cognitive theories: that metaphor processing involves the suppression of contextually irrelevant literal features, and the enhancement of contextually relevant figurative features. To do so, the study successfully replicated the experiment conducted by Rubio-Fernandez (2007). Finally, the thesis also discussed the measures that experimentalists should adopt, to tackle the replication crisis.

The experimental findings support the results of Rubio-Fernandez (2007), even though the sample size ended up smaller than desired and anticipated. Additionally, the limited scope of the study has limited the statistical analyses performed, and the results have been analysed and discussed at the most basic level. However, it still provides substantial support to the lexical modulation perspective on metaphoric processing by successfully replicating the findings of Rubio-Fernandez (2007).

Based on the discussion on the replication crisis, several measures that have been used in this study should be adopted for further research. Firstly, replications in general are an important tool to ensure the validity and quality of research. Additionally, using preregistrations, new material, new participants, a higher number of participants if possible, and better statistical methods will bolster future research.

Finally, within pragmatic and psycholinguistic research, using the scientific measures presented, metaphors should be further investigated focusing on more complex metaphors, and non-nominal metaphors, such as verbal metaphors.

References

- Allott, N. (2010). *Key Terms in Pragmatics*. Continuum.
- Aristotle. *The Poetics of Aristotle*. Project Gutenberg, Nov 1, 1999.
- Austin, J. L. (1956). A plea for excuses: The Presidential Address. *Proceedings of the Aristotelian Society*, 57, 1–30.
- Austin, J. L. (1979). Ifs and Cans. In *Philosophical Papers*. (eds. Austin, J. L., Warnock, G. J., & Urmson, J. O.) Oxford University Press.
- Barsalou. (1982). Context-independent and Context-dependent Information in Concepts. *Memory & Cognition*, 10(1), 82–93.
- Blasko, & Connine, C. M. (1993). Effects of Familiarity and Aptness on Metaphor Processing. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 19(2), 295–308.
- Bierwisch, M. (1967). Some semantic universals of German adjectivals. *Foundations of Language*, 3, 1–36.
- Bowdle, & Gentner, D. (2005). The Career of Metaphor. *Psychological Review*, 112(1), 193–216.
- Brown, C. (2007). *Cognitive psychology*. SAGE Publications Ltd.
- Bulldozer. (n.d.) In *Oxford English Dictionary*. Retrieved from <https://www-oed-com.ezproxy.uio.no/view/Entry/24527?rskey=GYAfyJ&result=1&isAdvanced=false#eid>
- Carston, R. (1996). Enrichment and Loosening: Complementary Processes in Deriving the Proposition Expressed. *UCL Working Papers in Linguistics*, 8, 205 – 232.
- Carston, R. (2010). Lexical Pragmatics, Ad Hoc Concepts and Metaphor: a Relevance Theory Perspective. *Italian Journal of Linguistics*, 22(1), 153-180.
- Carston, R. (2015). Contextual Adjustment of Meaning. In N. Riemer (Ed.) *The Routledge Handbook of Semantics* (pp. 195-210). Routledge.
- Carston, R. (2017). Relevance Theory and Metaphor. In E. Semino & Z Demnjen (Eds.) *The Routledge Handbook of Metaphor and Language* (pp. 42-55) Routledge.

- Carston, & Wearing, C. (2011). Metaphor, hyperbole and simile: A pragmatic approach. *Language and Cognition*, 3(2), 283–312.
<https://doi.org/10.1515/langcog.2011.010>
- Chapman, S. (2010). Paul Grice and the Philosophy of Ordinary Language. In Klaus Petrus (Ed.) *Meaning and Analysis: New essays on Grice* (pp. 31-46). Palgrave Macmillan.
- Chomsky, N. (1957). *Syntactic Structures*. Mouton Publishers.
- Chomsky, N. (1959). A Review of B. F. Skinner's Verbal Behavior. *Language*, 35(1), 26-58.
- Donne, J. (1959). *Devotions: Upon Emergent Occasions, Together with Death's Duel*. The University of Michigan Press.
- Estes, Z. & Glucksberg, S. (2000). Interactive Property Attribution in Concept Combination. *Memory & Cognition*, 28(1), 28–34.
- Frege, G. (1956). The Thought: A Logical Inquiry. *Mind*, 65(259), 289–311.
- Gentner, D., & Bowdle, B. (2008). Metaphor as structure-mapping. In R. W. Gibbs, Jr. (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 109–128). Cambridge University Press.
- Gentner, & Clement, C. (1988). Evidence for Relational Selectivity in the Interpretation of Analogy and Metaphor. *Psychology of Learning and Motivation*, 22, 307– 358.
- Gentner, & Wolff, P. (1997). Alignment in the Processing of Metaphor. *Journal of Memory and Language*, 37(3), 331–355.
- Gernsbacher, M. A. (1991). Cognitive Processes and Mechanisms in Language Comprehension: The Structure Building Framework. *The Psychology of Learning and Motivation*, 27, 217-263.
- Gernsbacher, M. A. & M. Faust (1991). The role of suppression in sentence comprehension. In G. B. Simpson (ed.), *Understanding Word and Sentence*. Elsevier Science Publishers.
- Gernsbacher, Keysar, B., Robertson, R. R., & Werner, N. K. (2001). The Role of Suppression and Enhancement in Understanding Metaphors. *Journal of Memory and Language*, 45(3), 433–450.

- Gibbs, Raymond W., Jr, Kushner, J. M., & Mills, W. Rob, III. (1991). Authorial Intentions and Metaphor Comprehension. *Journal of Psycholinguistic Research*, 20(1), 11–30.
- Gibbs Jr, R. & Tendahl, M. (2006). Cognitive Effort and Effects in Metaphor Comprehension: Relevance Theory and Psycholinguistics. *Mind & Language*, 21(3), 379–403.
- Gildea, P., & Glucksberg, S. (1983). On understanding metaphor: the role of context. *Journal of Verbal Learning and Verbal Behavior*, 22(5), 577–590.
- Glucksberg, S. (2001). *Understanding Figurative Language*. Oxford University Press.
- Glucksberg. (2003). The psycholinguistics of metaphor. *Trends in Cognitive Sciences*, 7(2), 92–96.
- Glucksberg. (2008). How Metaphors Create Categories – Quickly. In *The Cambridge Handbook of Metaphor and Thought* (pp. 67–83). Cambridge University Press.
- 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.
- Glucksberg, S & Haught, C. (2006). On the Relation Between Metaphor and Simile: When Comparison Fails. *Mind & Language*, 21(3), 360–378.
- Glucksberg, S., & Keysar, B. (1990). Understanding Metaphorical Comparisons. *Psychological Review*, 97(1), 3–18.
- Glucksberg, S. Newsome, M. R & Goldvarg, Y. (2001). Inhibition of the Literal: Filtering Metaphor-Irrelevant Information During Metaphor Comprehension. *Metaphor and Symbol*, 16(3-4), 277-298.
- Green, C.D. (2021). Perhaps Psychology’s Replication Crisis is a Theoretical Crisis that is Only Masquerading as a Statistical One. *International Review of Theoretical Psychologies*.
- Grice, P. (1975) *Logic and Conversation* In P. Cole and J. Morgan (eds), *Syntax & Semantics* 3: *Speech Acts* (pp. 41–58). Academic Press
- Grieve, J. (2021). Observation, experimentation, and replication in linguistics. *Linguistics*, 59 (5), 1343–1356. <https://doi.org/10.1515/ling-2021-0094>

- Hauk, O., Coutout, C., Holden, A., & Chen, Y. (2012). The time-course of single-word reading: evidence from fast behavioral and brain responses. *NeuroImage*, *60*(2), 1462-1477. <https://doi.org/10.1016/j.neuroimage.2012.01.061>
- Holyoak, K. J., & Stamenković, D. (2018). Metaphor Comprehension: A Critical Review of Theories and Evidence. *Psychological Bulletin*, *144*(6), 641–671.
- Hutchison, Balota, D. A., Neely, J. H., Cortese, M. J., Cohen-Shikora, E. R., Tse, C.-S., Yap, M. J., Bengson, J. J., Niemeyer, D., & Buchanan, E. (2013). The Semantic Priming Project. *Behavior Research Methods*, *45*(4), 1099–1114.
- Ioannidis. (2005). Why Most Published Research Findings are False. *PLoS Medicine*, *2*(8), e124–e124.
- Jones, L., & Estes, Z. (2006). Roosters, Robins, and Alarm Clocks: Aptness and Conventionality in Metaphor Comprehension. *Journal of Memory and Language*, *55*(1), 18–32.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*, XIII - 242. University of Chicago Press.
- Lion. (n.d.) In *Oxford English Dictionary*. Retrieved from <https://www-oed-com.ezproxy.uio.no/view/Entry/108800?redirectedFrom=lion#eid>
- Macha, J. (2012). Searle on Metaphor. *Organon F*, *19*, 186-197.
- Mandler. (2002). Origins of the cognitive (r)evolution. *Journal of the History of the Behavioral Sciences*, *38*(4), 339–353.
- Marr, D. (1982). *Vision*. Freeman.
- McNamara, T. P. (2005). *Semantic Priming: Perspectives from Memory and Word Recognition*. Psychology Press.
- Meyer, & Schaneveldt, R. W. (1971). Facilitation in Recognizing Pairs of Words: Evidence of a Dependence Between Retrieval Operations. *Journal of Experimental Psychology*, *90*(2), 227–234.
- Miller, G. A. (1956) The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97.

- Murphy, G. L. (1997). Polysemy and the Creation of Novel Word Meanings, in T. B. Ward, S. Smith, and J. Vaid (eds.), *Creative Thought: An Investigation of Conceptual Structures and Processes*, (pp. 235–265). American Psychological Association.
- Noveck, I. A. (2018). *Experimental Pragmatics: The Making of a Cognitive Science*. Cambridge University Press.
- Noveck, I., Bianco, M. & Castry, A. (2001). The Cost and Benefit of Metaphor. *Metaphor and Symbol*, 16(1&2), 109-121.
- Nunnally, J. (1960). The Place of Statistics in Psychology. *Educational and Psychological Measurement*, 20(4), 641–650.
- Onifer, & Swinney, D. A. (1981). Accessing Lexical Ambiguities During Sentence Comprehension: Effects of Frequency of Meaning and Contextual Bias. *Memory & Cognition*, 9(3), 225–236.
- Open Science Collaboration. (2015). Estimating the Reproducibility of Psychological Science. *Science*, 349(6251).
- Ortony, A. (1979). Beyond literal similarity. *Psychological Review*, 86(3), 161–180.
- Ortony, A., Schallert, D. L., Reynolds, R. E., & Antos, S. J. (1978). Interpreting Metaphors and Idioms: Some Effects of Context on Comprehension. *Journal of Verbal Learning and Verbal Behavior*, 17(4), 465–477.
- Pilkington. (2000). *Poetic effects: a relevance theory perspective*. J. Benjamins.
- Pynte, J., Besson, M., Robichon, F.-H., & Poli, J. (1996). The Time-Course of Metaphor Comprehension: An Event-Related Potential Study. *Brain and Language*, 55(3), 293–316.
- Rasinger, S. M. (2013). *Quantitative Research in Linguistics* (2nd ed.). Bloomsbury.
- Recanati, F. (2004). *Literal Meaning*. Cambridge University Press.
- Rieber, R. W. & Vetter, H. (1980). Theoretical and Historical Roots of Psycholinguistic Research*. In Rieber, R. W (Ed.) *Psychology of Language and Thought: Essays on the Theory and History of Psycholinguistics* (pp. 3-49). Plenum Press.
- Roettger. (2021a). Context Sensitivity and Failed Replications in Linguistics – a Reply to Grieve. *Linguistics*, 59(5), 1357-1358.

- Roettger. (2021b). Preregistration in Experimental Linguistics: Applications, Challenges, and Limitations. *Linguistics*, 59(5), 1227–1249.
- Ronderos, C.R, Guerra, E., Knoeferle, P (under review). “Mediating Factors of metaphor comprehension: the effects of familiarity, aptness and conventionality on processing verbal metaphors.”
- Rubio-Fernandez, P. (2007). Suppression in Metaphor Interpretation: Differences between M Meaning Selection and Meaning Construction. *Journal of Semantics (Nijmegen)*, 24(4), 345–371.
- Rubio-Fernandez, P., R. Breheny & M. Wei Lee (2003). Context Independent Information in Concepts: An Investigation of the Notion of “Core Features”. In F. Schmalhofer, R. M. Young, and G. Katz (eds.), *Proceedings of EuroCogSci 03: The European Cognitive Science Conference 2003*. Lawrence Erlbaum.
- Russell, B. (2005). On Denoting. *Mind*, 114(456), 873–887.
- Sandburg, C. (1916). ‘Fog’. In *Chicago Poems*. Henry Holt.
- Sapling. (n.d.) In *Oxford English Dictionary*. Retrieved from <https://www-oed-com.ezproxy.uio.no/view/Entry/170835?redirectedFrom=sapling#eid>
- Schooler, J. W. (2014). Metascience Could Rescue the 'Replication Crisis. *Nature (London)*, 515(7525), 9–9. <https://doi.org/10.1038/515009a>
- Schultz, D. P. (1975) *A History of Modern Psychology*. (2nd edition) Academic Press
- Searle, J. (1979). Metaphor. In: A. Ortony (Ed.) *Metaphor and Thought* (pp. 92-123). New York: Cambridge University Press.
- Searle, J. (1980). The Background of Meaning. In: Searle, J., Keifer, F. & Bierwisch, M. (eds.) *Speech Act Theory and Pragmatics*, (pp. 221-232). Kluwer Academic Publishers.
- Simpson, & Krueger, M. A. (1991). Selective Access of Homograph Meanings in Sentence Context. *Journal of Memory and Language*, 30(6), 627–643.
- Sönning, L. & Werner, V. (2021). The Replication Crisis, Scientific Revolutions, and Linguistics. *Linguistics*, 59(5), 1179-1206.

- Sperber, D. & Wilson, D. (1996). *Relevance: Communication and Cognition*. (2nd edition) Blackwell.
- Sperber, D. & Wilson, D. (2008). A Deflationary Account of Metaphors. In *The Cambridge Handbook of Metaphor and Thought* (pp. 84–106). Cambridge University Press.
- Sperber, D. & Wilson, D. (2012). A Deflationary Account of Metaphor. *Meaning and Relevance*. Cambridge University Press.
- Steel. (n.d.) In *Oxford English Dictionary*. Retrieved from <https://www-oed-com.ezproxy.uio.no/view/Entry/189540?rskey=F5o7xw&result=1&isAdvanced=false#eid>
- Swinney. (1979). Lexical Access During Sentence Comprehension: (Re)consideration of Context Effects. *Journal of Verbal Learning and Verbal Behavior*, 18(6), 645–659.
- Tannenhaus, Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of Visual and Linguistic Information in Spoken Language Comprehension. *Science (American Association for the Advancement of Science)*, 268(5217), 1632–1634.
- Tendahl, M. (2009). *A Hybrid Theory of Metaphor: Relevance Theory and Cognitive Linguistics*. Palgrave Macmillan.
- Vasishth, S., Merten, D., Jäger, L. A. & Gelman, A. (2018b). The Statistical Significance Filter Leads to Overoptimistic Expectations of Replicability. *Journal of Memory and Language*, 103, 151–175.
- Wagenmakers, E.-J., Wetzels, R., Borsboom, D., & van der Maas, H. (2011). Why Psychologists Must Change the Way They Analyze Their Data: The Case of Psi: Comment on Bem (2011). *Journal of Personality and Social Psychology*, 100, 426–432.
- Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review*, 20(2), 158–177.
- Wilson, D. (1998). Linguistic Structure and Inferential Communication. In B. Caron (Ed.) *Proceedings of the 16th International Congress of Linguists*. Pergamon.
- Wilson, D. (2011). Parallels and Differences in the Treatment of Metaphor in Relevance Theory and Cognitive Linguistics. *Intercultural Pragmatics*, 8(2), 177–196.

Wilson, D. & Carston, R. (2006). Metaphor, Relevance and the 'Emergent Property' Issue. *Mind & Language*, 21(3), 404–433.

Wilson, D. & Carston, R. (2007). A Unitary Approach to Lexical Pragmatics: Relevance, Inference and Ad Hoc Concepts*. In Noel Burton-Roberts (ed.) *Pragmatics*. Palgrave-Macmillan.

Winter, B. (2020). *Statistics for Linguists: An Introduction Using R*. Routledge.

APPENDICES: EXPERIMENTAL MATERIALS

Appendix A: Primes and Features

Primes	Literal Features	Figurative Features
Jungle	Forest	Confusing
Razors	Tool	Cuts
Lantern	Lamp	Inform
Gem	Stone	Precious
Joke	Funny	Easy
Veil	Cloth	Disguise
Dagger	Blade	Sharp
Butterflies	Insect	Hover
Storm	Weather	Destroy
Medicine	Pill	Aid
Rail	Bars	Thin
Arena	Stadium	Combat
Siren	Alarm	Annoying
Anchor	Weight	Safety
Thunder	Lightning	Loud
Magnet	Metal	Attractive
Shrimp	Animal	Tiny
Barrel	Wood	Huge
Portrait	Image	Description
Flower	Leaf	Fragile
Umbrellas	Plastic	Shelter
Encyclopaedia	Collection	Knowledge
Zoo	Cage	Chaos
Ice	Solid	Anger

Appendix B: Metaphoric contexts

I was done shopping an hour ago and have been trying to get out since. This store is a jungle.

Jimmy does comedy roast battles as a part time job; his insults are razors.

Studying history at university caused Matthew to switch political parties. Education is a lantern.

Throughout her rehabilitation Hannah has gotten all the support she has needed. Her husband is a gem.

I have yet to hand in an assignment and my professor still gave me an A; my computer skills course is a joke.

Ever since my stepdaughter became a teenager, I can't understand a word she is saying. Her constant sarcasm is a veil.

James has had a stomach-ache ever since he learnt of the affair. A lie is a dagger.

That was the best ballet I have ever seen, the dancers made it look as if their feet never touched the ground. Those dancers were butterflies.

Sophie has been looking tired the last few weeks, her divorce was a storm.

After they introduced live concerts at the nursing home the patients have been much more cognisant. Music is medicine.

You can see how unhealthy the fashion industry is, that fashion model is a rail.

I haven't seen anyone struggle as much as John did when Sarah and Liza told him to choose one of them; his mind was an arena.

I cannot listen to the new radio host; her voice is a siren.

Without Ryan's help I am not sure how this year would have gone; a best friend is an anchor.

I could not hear a word you said during the football match, the cheering crowd was thunder.

I cannot help talking to the mail carrier every morning, his smile is a magnet.

Not everyone in my family does body building, my young cousin is a shrimp.

Steve actually managed to be kicked out of an all you can eat buffet once; his stomach is a barrel.

Thanks to modern advances in technology, detective work has never been easier. A criminal's fingerprint is a portrait.

Every happy couple I know works hard to maintain their relationships. Love is a flower.

Since its raining, I would much rather go for a walk in the forest; trees are umbrellas.

My teacher lets us write essays on any topic we'd like, and we can ask him about anything.
He is an encyclopaedia.

Because of the budget cuts Frank's kindergarten has doubled the pupils and gone down to one teacher per class. Frank's kindergarten class is a zoo.

Ever since the teacher told my mom I have been skipping school she hasn't stopped staring at me. Her unflinching gaze is ice.