Assessing the Meaning of Metaphors in Real-Time: A Cross-Modal Investigation

Iola Patalas

A Thesis

In the Department

of

Psychology

Presented in Partial Fulfilment of the Requirements

For the Degree of

Master of Arts (Psychology) at

Concordia University

Montréal, Québec, Canada

July 2018

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Iola Patalas

Entitled: Assessing the Meaning of Metaphors in Real-Time: A Cross-Modal Investigation

and submitted in partial fulfilment of the requirements for the degree of

M.A., Psychology

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

______________________________ Chair

Dr. Andreas Arvanitogiannis

______________________________ Examiner

Dr. Ben Eppinger

______________________________ Examiner

Dr. Norman Segalowitz

______________________________ Supervisor

Dr. Roberto G. de Almeida

Approved by: ____

Dr. Andrew Chapman, Graduate Program Director

________________, 2018 ____

Dr. André Roy, Dean of Faculty of Arts and Sciences

ABSTRACT

Assessing the Meaning of Metaphors in Real-Time: A Cross-Modal Investigation

Iola Patalas, M.A.

Concordia University, 2018

Natural language is replete with figurative expressions such as *my lawyer is a shark*, and listeners are expected to intuitively understand the *intended*, rather than the *literal* meaning of such expressions. But what are the cognitive resources involved in attaining meaning for such sentences? According to proponents of the *pragmatic* model of metaphor comprehension, metaphors are first interpreted literally, and then, upon realizing they cannot be true, listeners search for *implicatures* that could convey the speaker's intended meaning (Searle, 1979; Grice, 1989). In contrast, direct-access models of metaphor processing have posited that metaphors can be understood *directly*, circumventing higher-order cognitive processes (e.g. Glucksberg & Keysar, 1990). The present thesis investigated these theories using a cross-modal lexical decision paradigm (Swinney, 1979) with a novel brief masked target presentation at two probe points, in order to assess the moment-by-moment *on-line* processes involved in metaphor comprehension. We predicted that, following the *pragmatic* model, literally related target words would yield greater priming effects at the vehicle (e.g., *shark*) recognition point (a), compared to figurative targets, in both metaphor and simile conditions. At the later probe point (b), 500 ms after the vehicle's recognition point, we expected that figurative targets would yield greater priming effects during metaphor comprehension as literal meanings were discarded. Results obtained from a preliminary sample demonstrated priming of related target words across conditions, but no significant differences between conditions. We discuss how these results may best be interpreted as supporting the dual-processing account of metaphor interpretation put forth by

Carston (2010), which suggests that metaphors are held in the mind as literally true even as fast *ad-hoc* concepts are simultaneously created to interpret intended speaker meaning.

Acknowledgments

I would like to take this opportunity to thank Dr. Roberto de Almeida, whose guidance and mentorship has allowed me to carry out this research and develop as a cognitive scientist. His suggestions and expertise were critical particularly to the development of this experiment and writing of the present thesis. My thanks goes out to my lab manager Caitlyn Antal, whose expertise and support in the lab couldn't have been done without. Finally, I would like to thank all the volunteers who made this research possible, particularly Kristen Charles and Isabelle Oppong, who helped consolidate the data from this experiment, Alexa Falcone for offering her assistance whenever it was needed, and Nathan Boulianne who helped with the initial data collection.

Contribution of Authors

This study has been possible thanks to the contributions of two authors. The study was conceived of by Dr. Roberto de Almeida and Iola Patalas. Iola Patalas was responsible for stimuli creation and data collection, performed all analyses and wrote the present thesis. All work was conducted under the supervision and guidance of Dr. Roberto de Almeida, who contributed to the experimental design and revisions to the manuscript.

Table of Contents

List of Figures

List of Tables

Assessing the Meaning of Metaphors in Real-Time: A Cross-Modal Investigation

Are lawyers sharks? On the surface, the answer to this question is self-evidently 'no' - lawyers (human beings belonging to a certain profession) cannot, logically, be large carnivorous fish. However, a reader with a certain experience of lawyers might be tempted to answer yes to this question, because most readers will understand the question was not meant literally. Natural language is replete with figurative language, and listeners are expected to intuitively understand the *intended*, rather than the *literal* meaning of such expressions – but what are the cognitive resources involved in attaining meaning for such sentences? How and why people interpret nonliteral utterances as meaningful is a question essential to the understanding of language, and a subject of debate among cognitive scientists.

Metaphor Versus Simile

Nominal metaphors in the form *X is Y* (e.g. *lawyers are sharks*) are traditionally thought to be an alternate form of the simile (e.g. *lawyers are like sharks*), which involve direct comparisons in the form *X is like Y* – a view dating back to Aristotle (trans. 1926). Both these forms imply a relationship between a topic (*lawyers*) and a vehicle (*sharks*) that depends on listener interpretation to be fully understood. In both cases, a listener's interpretation of the comparison likely relies on identifying features of the vehicle that could plausibly be applied to the topic – for instance, sharks are often thought to be aggressive and predatory, features which could also be attributed to certain lawyers.

The key difference between these two forms of expression is a word such as *like* in simile, which renders the simile literally comprehensible – after all, it is always possible for one thing to be *like* another in some way. In contrast, a nominal metaphor does not explicitly invite a comparison between topic and vehicle but rather suggests that the vehicle comprises a superordinate category encompassing the topic, or predicates something about the topic, as in the sentence *lawyers are people*. The listener's assumption that a sentence with the structure *X is Y* is a categorical statement denoting a superordinate category falls apart upon hearing a seemingly anomalous vehicle like *sharks*, but this generally does not make a metaphor incomprehensible, even though it is literally impossible.

The purpose of the present thesis is to investigate the nature of the representations computed during the real-time processing of metaphors and similes. Specifically, we aimed to compare these two types of expression in real time to elucidate differences in processing that might occur due to their fundamentally metaphorical and literal nature, respectively, and by doing so to shed light on the cognitive mechanisms involved in reaching an understanding of the intended meaning of such constructions.

Pragmatic Theory of Metaphor Comprehension

Various theories have arisen to account for how listeners derive meaning from metaphors of the *X is Y* form. Early theories suggested that metaphors are first interpreted literally, and then, upon realizing they cannot be literally true, listeners will search for alternate meanings (Grice, 1975; Searle, 1979). According to Searle (1979), literal utterances denote a set of truth conditions that are applicable only in a certain context that is agreed upon implicitly by the speaker and listener of a sentence. Thus, a sentence like “Sally is tall” is true only insofar as the referent for *tallness* is other women, and not, e.g., giraffes – but as long as the speakers share the same referent, no additional information is needed to interpret a literal statement (Searle, 1979). In contrast, Searle (1979) posited that in order to search for an appropriate utterance meaning for a metaphor, a listener must first fully process a sentence and find it “defective” relative to these truth conditions (p. 103) and then search for salient features which the topic and vehicle might have in common. Here Searle (1979) makes a distinction between s*entence meaning* (that is, the literal denotative meaning where *sharks* really means SHARKS), and *utterance meaning* (what the speaker might have intended to convey by the phrase) (p. 84). He proposes a three-step process involved in communicating the speaker's utterance meaning to a listener that involves first fully processing the literal, semantic meaning of the sentence, then using surrounding context to determine its truth value, and finally, upon finding some truth condition lacking, searching for alternative, figurative meanings (Searle, 1979). This theory of metaphor interpretation is commonly referred to as the *pragmatic* model.

Another pragmatic theory which deviates slightly from Searle's (1978) model was proposed by Davidson (1978). Davidson's (1978) model proposed that, rather than being understood strictly by their *propositional* content, metaphors invite a reader to *imagine* them as true: “A picture is not worth a thousand words, or any other number. Words are the wrong currency to exchange for a picture” (Davidson, 1978, p. 263). However, for the purposes of our comparison, we propose there is no significant distinction between envisioning a lawyer as a shark and interpreting that the lawyer is literally a shark in a semantic sense – both processes involve interpreting the referent *shark* literally prior to accessing the pragmatic system for alternate potential meanings. Indeed, Davidson (1978) describes this process thus: “Absurdity or contradiction in a metaphorical sentence guarantees we won’t believe it and invites us, under proper circumstances, to take the sentence metaphorically” (p. 42).

Before we move on, it is relevant to make note of the difference between *semantic* and *pragmatic* processes generally. A distinction between these two processes in the linguistic context is made by de Almeida and Lepore (2018) in their discussion on linguistic modularity. According to their theory, propositional content as encoded by the linguistic cognitive architecture is merely symbolic, allowing for fast computations to be made in the linguistic system (de Almeida & Lepore, 2018). This *semantic* output is necessarily separate from the *enriched* understanding of a sentence which arises after drawing upon other cognitive systems such as memory, world knowledge, etc. - what we refer to here as the *pragmatic* system. Thus, the surface structure of a sentence, referred to by de Almeida & Lepore (2018) as a "shallow" representation (p. 115), is merely the output of linguistic computations performed on symbolic aspects of language, and this representation does not automatically draw upon deeper or more nuanced potential sentence meanings. According to this theory, there is no way for the linguistic system to *automatically* (or *directly*) access all the connotations of *shark* as these deeper understandings rely on the *pragmatic* system.

Direct-Access Accounts of Metaphor Comprehension

Most modern studies on metaphor have rejected the pragmatic model, suggesting instead that metaphors are immediately comprehensible by the linguistic system and do not involve additional cognitive resources to process compared to literal statements (e.g. Glucksberg & Keysar, 1990; Gibbs, 1994; Wolff & Gentner, 2000). This can be explained by a mechanism where metaphors are taken as comparisons between categories (e.g. Glucksberg & Keysar, 1990; Glucksberg, 2003), or processed via mapping common word properties, or *constituent features*, which are stored in the linguistic system as lexical properties of individual words (Wolff & Gentner, 2000). The comparison theory proposed by Glucksberg and Keysar (1990) postulates that in a phrase such as *my job is a jail*, a *jail* could be interpreted as shorthand for a superordinate category denoting an unpleasant or involuntary situation to which a *job* could also plausibly belong. Thus, according to this theory, the vehicle in a metaphor takes on the function of a taxonomic category which logically includes the topic (Glucksberg & Keysar, 1990). Wolff and Gentner (2000) rejected the idea that metaphor processing is *directional* (i.e. a superordinate/subordinate relationship between vehicle and topic), instead proposing that metaphors are processed via alignment of "salient properties" common to topic and vehicle, both figurative and literal. Along the same lines, the graded salience theory developed by Giora (2003) claims that a listener chooses between a number of initially available interpretations of a sentence based on properties like familiarity and aptness.

Dual Processing Model of Metaphor Comprehension

What about situations that involve mixed metaphors, or long passages involving multiple layers of metaphor? Building on her interpretation of Davidson's (1978) model of metaphor as evoking *images*, Carston (2010) argues that there are two simultaneous processes driving metaphor comprehension in real-time: (1) A process of ad-hoc interpretation where a listener treats metaphors as propositional content and adjusts the understanding of each metaphor using *pragmatic* processes to understand individual metaphorical meanings; (2) a process wherein a listener simultaneously keeps the literal semantic meaning of metaphorical phrases in mind to facilitate ongoing understanding of speaker meaning. An example is given by Carston (2010) to demonstrate why this would intuitively be true: when hearing a construction such as (1a), many listeners would be confused by the concept of *watering* a *spark*, even though individually the two halves of the mixed metaphor are clearly understandable (Carston, 2010; p. 305).

(1) a. If you find a student with a spark of imagination, water it. (Tirrell, 1989)

In this case, no cognitive dissonance would be experienced by a listener if each metaphor were interpreted separately using pragmatic processes after rejection of literal, semantic representations. Thus Carston (2010) concludes that literal interpretations likely persist after they should be rejected. Again, this conclusion is not necessarily at odds with the *pragmatic* model as it is possible that the surface meaning is the only meaning initially accessed, while both surface and contextual meanings remain simultaneously accessible after pragmatic processes kick in.

Evidence Against the Direct-Access Model

According to Glucksberg (2003), “there is a consensus in the field that literal meaning does not have unconditional priority” over figurative meanings (p. 92) and thus figurative meanings of an utterance can be accessed directly by the linguistic system, although the consensus does not extend to the mechanism by which meaning is attained. However, this view is most often based on the fact that tests involving comprehension are usually *offline*, i.e., require conscious judgment, and are thus not informative regarding what happens as sentences containing metaphors unfold in real time. For instance, Glucksberg, Gildea, and Bookin (1982) asked participants to read literal and metaphorical sentences and to judge whether they were literally true or literally false. Based on a finding that it took longer for participants to judge statements as false if they had a common metaphorical interpretation (e.g. *jobs are jails*), the authors concluded that a metaphorical meaning is immediately available along with a literal meaning and thus interfered with subjects' classification of metaphorical sentences as literally false (Glucksberg et al., 1982). This result has been widely replicated since, using similar offline judgment tasks about various metaphor configurations (e.g. Keysar, 1989; Wolff & Gentner, 2000). It would be premature, however, to conclude that these differences in reading time are necessarily caused by automatic metaphor processing by the linguistic system, as an offline judgment task does not provide any evidence about what is actually happening as readers process sentences in real time. The results obtained by offline studies such as Glucksberg and colleagues' (1982) could be equally compatible with the hypothesis that pragmatic processes interfere with literality judgments after the sentence has been fully processed, but before participants register a response.

Eye-tracking studies measuring reading times are also frequently invoked as evidence for the automatic processing of metaphors (e.g. Ortony, Schallert, Reynolds & Antos, 1978; Inhoff, Lima & Carroll, 1984). Inhoff and colleagues (1984) manipulated the length of contexts preceding short metaphorical statements and found that a short metaphor-biasing context such as (2a) was processed more slowly than a literal-biasing context such as (2b). When longer context was given, no significant difference in reading time was found between metaphor-biasing and literal-biasing contexts (2c-d) (Inhoff et al. 1984).

(2) a. At a meeting of the women's club, the hens clucked noisily.

b. In the back of the barn, the hens clucked noisily.

c. At a meeting of the women's club the youngest member requested the floor and brought up the issue of supporting the equal rights amendment. The importance of the issue outweighed her discomfort in speaking before the roup. They reacted as she expected. The hens clucked noisily.

d. In the back of the barn, the farmer's youngest child gathered pebbles and skipped them deftly across a puddle by the chicken coop. He knew that he was supposed to be feeding the animals but he kept on flicking at the birds. The hens clucked noisily.

While Inhoff and colleagues (1984) interpreted this result to mean that sufficient biasing context can enable direct processing of a metaphor in the linguistic system, alternate explanations for this finding have been proposed as equally plausible. De Almeida, Manouilidou, Roncero and Riven (2010) argued that these results could be consistent with the pragmatic model of metaphor processing if a sufficiently informative biasing context speeds up the search for alternate (figurative) meanings once the literal meaning of a sentence is rejected, compared to a short or uninformative biasing context. Context could lead to insignificant differences in overall reading time by making pragmatic cognitive processes much faster, not necessarily by circumventing them entirely. Studies directly comparing reading times for simile and metaphor have found that metaphors take longer to read relative to similes (e.g. Janus & Bever, 1985). In a more recent eye-tracking study comparing nominal metaphors such as *knowledge is a river* to their corresponding similes (*knowledge is like a river*), Ashby, Roncero, de Almeida and Agauas (2018) found longer reading times for the vehicle word (e.g. *river*) in metaphors compared to similes. Participants in this study also regressed from the vehicle region more when reading metaphors than when reading similes (Ashby et al., 2018), suggesting that readers had more initial difficulty processing metaphorical, compared to literal, statements immediately upon encountering the vehicle word.

It should be noted that metaphors are not all processed equally easily. Reversing the topic and vehicle in nominal metaphors – for example, changing *some jobs are jails* to *some jails are jobs* - can impede full comprehension by readers, although this transformation does not significantly alter the time it takes to make literality judgments (Wolff & Gentner, 2000). More importantly, metaphor processing can be facilitated by properties such as aptness and familiarity. Aptness, which describes the extent to which a vehicle's properties encompass salient features of the topic, has been found to mediate a preference for metaphor over simile (Chiappe, Kennedy & Chiappe, 2003; Roncero, de Almeida, Martin & de Caro, 2016), and ease metaphor processing (Bowdle & Gentner, 2005). Highly familiar metaphors, but not unfamiliar metaphors, primed figurative meanings in a cross-modal lexical priming study (Blasko & Connine, 1993) and were read faster than literal sentences in an eye-tracking study (Columbus, Sheikh, Cote-Lecaldare, Hauser, Baum & Titone, 2015). Furthermore, Columbus et al. (2015) found a relationship between executive control and comprehension of unfamiliar metaphors, concluding that interpretation of novel metaphors relies upon pragmatic processes to come up with possible figurative meanings. This difference could be due to a process wherein highly conventional metaphors become lexicalized, like common idioms, and can be easily retrieved from memory. Searle (1979) refers to these as 'dead metaphors' (p. 110) – explaining that in some cases, the meaning of a figurative phrase can eventually become so common that it effectively functions as a literal expression denoting the intended meaning. However, Ashby et al. (2018) found that moderately familiar, highly apt metaphors were still processed with more difficulty than similes containing the same topic-vehicle pair.

Studies investigating on-line metaphor processing by measuring event-related potentials (ERP) have demonstrated that figurative targets elicited larger N400 amplitudes than literal targets (e.g. Pynte, Besson, Robichon & Poli, 1996; Lai, Curran, & Menn, 2009), which suggests that figurative language is more difficult to process. This could be due to the detection of an incongruence between literal and intended speaker meaning. However, ERP is flawed due to the low ecological validity of timed serial target presentation, compared to self-paced reading. Furthermore, relatively few studies have investigated metaphor in auditory speech contexts, with most studies on metaphor restricted to reading contexts. In light of this mixed evidence, and dearth of studies investigating the moment-by-moment cognitive processes during metaphor listening and reading, it is too early to conclude that the debate about how metaphor interpretation occurs (i.e., directly or via a two-stage, *semantics*-then-*pragmatics* process) is settled, or a matter of consensus as Glucksberg (2003) claims.

A final point of contention with the direct-access (semantic) model of metaphor interpretation is the mechanism which allows listeners to access alternative meanings of the vehicle – whether they be taxonomic categories or constituent features implied by the vehicle itself. The question remains: how exactly do listeners understand what is meant by *sharks* while bypassing any consideration of sharks themselves? It is easy enough to imagine that upon hearing a dead metaphor such as “a warm welcome” (Searle, 1979, p. 98), where even dictionaries have come to associate warmth with meanings other than a certain temperature range, listeners may have direct access to this alternate meaning of “warm” in an appropriate context. But in novel metaphors, speakers have only the real meaning of the vehicle as a referent from which to glean possible figurative meanings. After all, as demonstrated by open-ended norming studies such as Roncero and de Almeida (2015), when listeners are asked to identify constituent features which link topic and vehicle in a metaphorical context, there is only a loose consensus between listeners about the words that exemplify this relationship. It is unlikely that all English speakers hold the same set of associations with a word like *sharks* – or, indeed, *lawyers* – such that they do not need to refer to their real-world knowledge about sharks and lawyers to arrive at the most salient figurative meaning of *lawyers are sharks*.

Defenders of the pragmatic theory claim that there is a moment during processing – however brief – where *sharks* really does mean sharks, the fish. Furthermore, this can be argued to be essential to access appropriate alternative understandings of a metaphor. Lepore and Stone (2014) point out that saying *lawyers are sharks* is essentially different than saying *lawyers are like sharks* in that it invites the listener to actually *imagine* lawyers as sharks rather than simply comparing their similarities. This process of imagining, or accessing imagery, may explain why simile and metaphor do not call to mind identical associations in readers (e.g. Roncero & de Almeida, 2015). It could also explain why metaphors or similes are preferred in different contexts. Theorists arguing that figurative meanings are accessed automatically have yet to explain how a listener hearing the word *shark* would even know to reject its literal meaning, if as they claim this literal meaning is not accessed at all during sentence processing. Is it possible that the word *shark* does not directly correspond to the concept of a shark in our mental lexicon?

In order to resolve the conflict in the literature between pragmatic and direct-access models of metaphor processing, it is first necessary to gain insight into exactly what happens the very moment a reader or listener encounters a vehicle intended to convey a figurative meaning. To our knowledge, no studies have actually investigated what sort of information is accessed in the moment-by-moment process of spoken comprehension of metaphors and similes using the same topic and vehicle. According to the various semantic (direct-access) models of metaphor processing, listeners should have immediate access to the figurative meaning of both similes and metaphors upon hearing the vehicle – possibly to the exclusion of literal meanings – and the preference for this interpretation should not change over time. The goal of this research is to test the alternate hypothesis, consistent with the pragmatic model of metaphor processing that listener interpretation of a metaphor does change over time, with an initial full processing of the *literal sentence meaning* of a metaphor eventually leading to a figurative understanding.

The Present Study

The aim of this study was to test what types of interpretations are accessed at the moment the vehicle is first recognized in both metaphor and simile, as well as in the moments after when comprehension of *utterance meaning* has been reached. We sought to compare the moment-by-moment comprehension of nominal metaphors in the form *X is Y* and similes in the form *X is like Y* using a cross-modal lexical decision task (CMLD; Swinney, 1979). In this task, participants listen to aurally presented sentences for comprehension and are simultaneously presented with a visual target to perform a lexical decision task (i.e., pressing "yes" if the target is a word, "no" otherwise) in which response times (RTs) are collected. The main assumption behind the technique is that RTs to targets reflect the relation between a visual target and a prime word in the sentence (here, the vehicle *Y*). Specifically, recognition of the target word should be facilitated by hearing a related prime word, and thus yield a faster reaction time compared to a target that is semantically unrelated.

For this CMLD task, we employed aurally presented sentences containing metaphors or similes with the same constituents (except for *like* in similes), following a natural speech rhythm. Our visual targets were either: a word related to a literal meaning of the vehicle, a word related to the metaphorical meaning of the vehicle, or one of two frequency- and length-matched controls used to calculate priming effects (priming = $RT_{(control)}$ - $RT_{(target)}$). We employed two probe points (i.e., when targets appear on the screen relative to the speech stream): (a) recognition point of the vehicle (determined by a gating paradigm; Zwitserlood, 1989), and (b) 500ms following recognition of the vehicle, to allow for sentence comprehension to occur. Metaphors, similes, and target words were selected from Roncero and de Almeida (2015), which obtained norms for properties like aptness, familiarity and conventionality. We selected primarily novel metaphors with high aptness but a range of familiarity ratings.

The rationale for this methodology is as follows: first, listening to spoken metaphors in an on-line lexical decision task allows for an analysis of metaphor interpretation that is both highly time-sensitive and naturalistic. Using a simple lexical decision task rather than an off-line judgment task means that participants do not base their responses on a conscious assessment of sentence meaning – indeed, they are not aware that this task is meant to test their comprehension of metaphors at all. Instead, priming for each target should reflect the interpretation of a sentence that is available at the moment visual targets are presented. Second, using similes as literal controls allows for all constituent words besides *like* (including target and vehicle) to remain identical, thus allowing for direct comparisons between literal and figurative interpretations of each topic-vehicle pair. A possible criticism of this approach is that similes are typically classed as figurative language, but as argued in the preceding paragraphs, similes differ from metaphors in that they remain literally comprehensible as comparison statements. Thus, according to the three-step model of metaphor processing detailed by Searle (1979), similes should not be identified as defective in a literal sense upon recognition of the vehicle.

Following the pragmatic model, we hypothesized that targets corresponding to the literal meaning of the vehicle word would yield greater priming effects at recognition point (a), compared to figurative targets. This should be true for both metaphor and simile if the denotative meaning of the vehicle is the only meaning initially accessed. At the later probe point (b), we expected that figurative targets would yield greater priming effects during metaphor comprehension as literal meanings were discarded. In contrast, we predicted that RTs for literal targets during simile comprehension will continue to be lower at this later probe point (b), since similes are comprehensible literally and literal meanings should not necessarily be discarded.

Alternatively, if metaphors are automatically processed using all possible constituent word properties, as suggested by Wolff & Gentner (2000), priming effects for figurative targets as well as literal targets should be seen immediately at recognition point (a). Since a direct-access account of metaphor processing suggests metaphor should be no more difficult to process than simile, according to this model no difference in priming effect should be apparent between metaphor and simile conditions at either time point, and literally related targets should generally produce smaller priming effects than figuratively related targets at the later time point (b) (for a

Table 1.

Priming effects predicted for the conditions employed in the present study, according to different metaphor processing models.

Model	Metaphor				Simile			
	(a) Early		(b) Late		(a) Early		(b) Late	
	Literal	Figurative	Literal	Figurative	Literal	Figurative	Literal	Figurative
Pragmatic	strong	weak/none	weak	strong	strong	weak/none	strong	weak
Direct Access	strong	strong	weak	strong	strong	strong	weak	strong
Dual Processing	strong	strong	strong	strong	strong	strong	strong	strong

full list of predictions, refer to Table 1). To our knowledge, no other study has investigated on-line metaphor processing using this method in order to test specific predictions of the *pragmatic* and *direct-access* models with regards to the moment-by-moment time-course of metaphor processing.

Method

Participants

Participants were recruited from Concordia University using the Psychology Participant Pool and from surrounding communities using posters and online advertisements. All participants completed a pre-screening language questionnaire which assessed their language background, self reported verbal and written fluency, and learning disabilities. Based on their questionnaire responses, we selected 37 native English speakers between the ages of 19 and 59 (M=26.32, SD=8.07; 26F) with normal or corrected-to-normal vision and hearing who met the following inclusion criteria: (1) They learned English before the age of 5 (M=1.19, SD=1.47) and identified it as their native and dominant language; (2) they rated themselves as fluent in speaking, listening, and reading English; (3) they reported no history of hearing or reading disability. Participants who were recruited via the Concordia participant pool were compensated with course credit while all other participants were compensated with $10 CAD for one hour of participation. Participants for two pretests are described along with the pretests below.

Materials

Experimental materials consisted of 32 sentences containing metaphors/similes in the form *X is (like) Y* and 160 filler sentences. Metaphor/simile sentences were selected from Roncero and de Almeida (2015), a published set of metaphor/simile sentences with accompanying norms. The sentences we selected from this set of norms were chosen on the basis of their high aptness ratings (rated above 6 on a scale of 1 to 10, with 10 being the most apt), but had a broad range of familiarity ratings. The Roncero and de Almeida (2015) norming study asked participants to generate associates/explanatory words for both the simile and metaphor versions of each sentence and for the topic and vehicle words in isolation. For use as our figuratively related targets, we selected explanatory words generated for each metaphor by the highest possible number of participants, which did not appear as associates for the vehicle word in isolation. For our literally related targets, we selected words which were generated as associates of the vehicle word by the highest possible number of participants and which did not appear as explanatory words for the metaphor on the whole (see Appendix for sentences and related targets).

Exclusion of automatic associates. To ensure that any potential priming effects were not derived from an 'automatic' association between the vehicle and target words (i.e., due to being frequently paired in speech, like 'salt and pepper'), we conducted a norming experiment where each vehicle word was read aloud to 12 native speakers of English, who were asked to say the first word that comes to mind out loud. Their responses were collected and any word which was named more than twice was excluded from selection as a target for that vehicle word.

The unrelated control words selected to calculate priming effects were chosen according to the following criteria. For each related target word, written frequency was calculated from the Corpus of Contemporary American English (COCA), a database of American English texts collected from 1990-2017 including fiction, non-fiction and academic texts. Matched (unrelated) control words were selected to have the same number of letters, same number of syllables, same morphological structure and similar frequency in the COCA database.

Sentence recording and targets. Metaphors/similes were embedded in longer sentences with explanatory contexts which we generated, with the word 'because' following each vehicle word to control for interference from explanatory contexts. Filler sentences did not repeat the topic or vehicle words of any experimental sentences. Of these, 32 followed a similar sentence structure as experimental sentences, while 128 filler sentences did not syntactically resemble experimental sentences. Visual targets for filler sentences were 64 real English words and 96 ‘nonsense’ strings of letters that did not resemble English words, of varied lengths to reflect the varied lengths of experimental targets. All sentences were read by a female native English speaker and recorded for aural presentation, with natural prosody and reading speed. Special attention was given to matching the prosody and timing of metaphor and simile pairs, to make them nearly identical except for the word ‘like.’

Recognition Times. We employed a gating paradigm to determine the recognition point of each vehicle word, following the procedure developed by Zwitserlood (1989). Recordings of each vehicle word were cut into slices increasing by 50ms each. These were played consecutively to 10 native speakers of English over noise canceling headphones. Participants were asked to write down what word they thought they were hearing after each slice was presented. Their responses were collected and recognition times for each word were defined as the moment when 80 percent of participants correctly identified the word (with or without pluralization). During the lexical decision task, the early time point was defined as 40ms prior to recognition time, to account for screen refresh rate and the fact that the word could have become recognizable anytime within the 50ms slice participants heard during the gating task. Late time points were defined as 500ms following recognition time to avoid interference from words later in the sentence.

Experimental Design

To avoid repeating experimental sentences, 16 counterbalanced lists were created following a 2 x 2 x 2 x 2 design. Each topic/vehicle pair was presented in either a metaphor- or simile-containing sentence, along with a figuratively related target, literal target or matched control target, at an early (recognition) or late time point. Each block contained two experimental sentences in each condition along with all 160 filler sentences, 20 of which were followed by comprehension questions to ensure participants were attending to aural stimuli. Each participant completed two blocks containing one list each – i.e., each participant heard both the simile and metaphor version of each sentence once in total. The sentences were randomized in order within each block of trials and participants were randomly assigned to each set of lists. Due to an error in coding which was later fixed, some participants did not see all sentences in each list, and thus 3-7 data points are missing for many of the participants.

Procedure

Participants were tested on an iMac flatscreen computer using Psyscope X B57 (Cohen, MacWhinney, Flatt, & Provost, 1993) using a button box. After voluntary consent was obtained, each participant was seated in front of the screen in a dark room, equipped with noise-cancelling headphones, and instructed to attend to both the aurally presented sentences and visual stimuli on the screen. Participants were instructed that their primary task was to identify whether the letters they saw on the screen constituted an English word and to press a button to indicate YES or NO as quickly and accurately as possible, while their secondary task was to answer comprehension questions about the sentences they heard over the headphones.

Each trial consisted of a prompt asking participants to press a button when they were ready for the next trial, followed by an aural presentation of each sentence (see Figure 1). Target

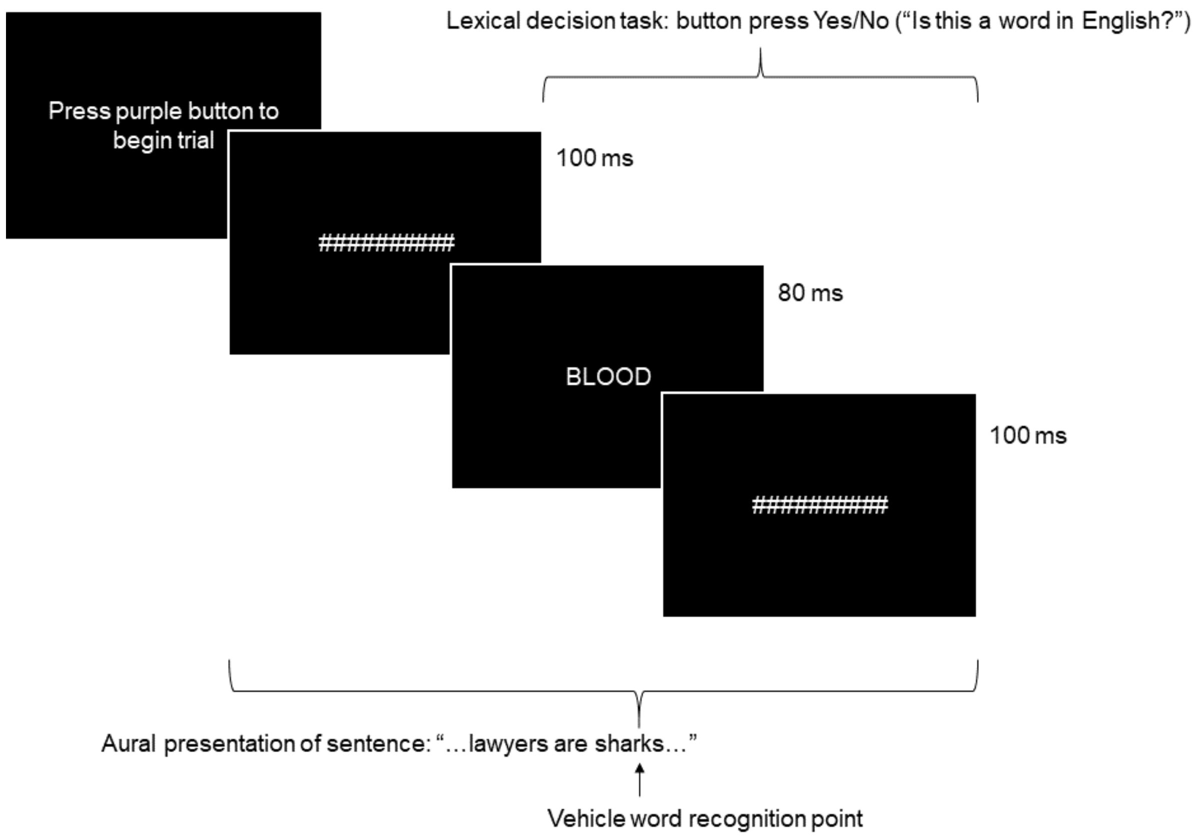

Figure 1. Time-course of events in each experimental trial. Two experimental probe points were used with this figure showing only the recognition point and not the late time point. Durations are presented in milliseconds (ms).

words appeared in white 20-point Arial font text in capital letters on a black screen for 80ms each, preceded and followed by masks which appeared for 100ms. This brief masked priming procedure was meant to reflect faster and more automatic processes of recognition rather than slower processes of judgment. Masked priming (see: Forster, 1999) reflects early processes of lexical recognition which should be uncontaminated by other semantic factors. Each participant was given five randomized practice trials, during which the experimenter answered questions and corrected mistakes.

Results

The main purpose of our analysis was to investigate priming effects, defined as the difference between unrelated and related (literal and figurative) visually presented targets, in both metaphor and simile conditions, and at both time points in order to investigate potential differences in priming effects between conditions.

Data Analysis

Analysis of reaction times (RTs) was restricted to correct trials (i.e., those where participants correctly identified the target as an English word) while incorrect trials were omitted (13% of all data points). As is standard in lexical decision paradigms (Friedmann, Taranto, Shapiro & Swinney, 2008), all reaction times above 2 seconds were discarded prior to data analysis (2% of all data points). Based on a priori decisions, we discarded blocks of trials where participants answered fewer than 70% of comprehension questions correctly (as in Friedmann et al., 2008), following the assumption that participants may not have listened carefully to aurally

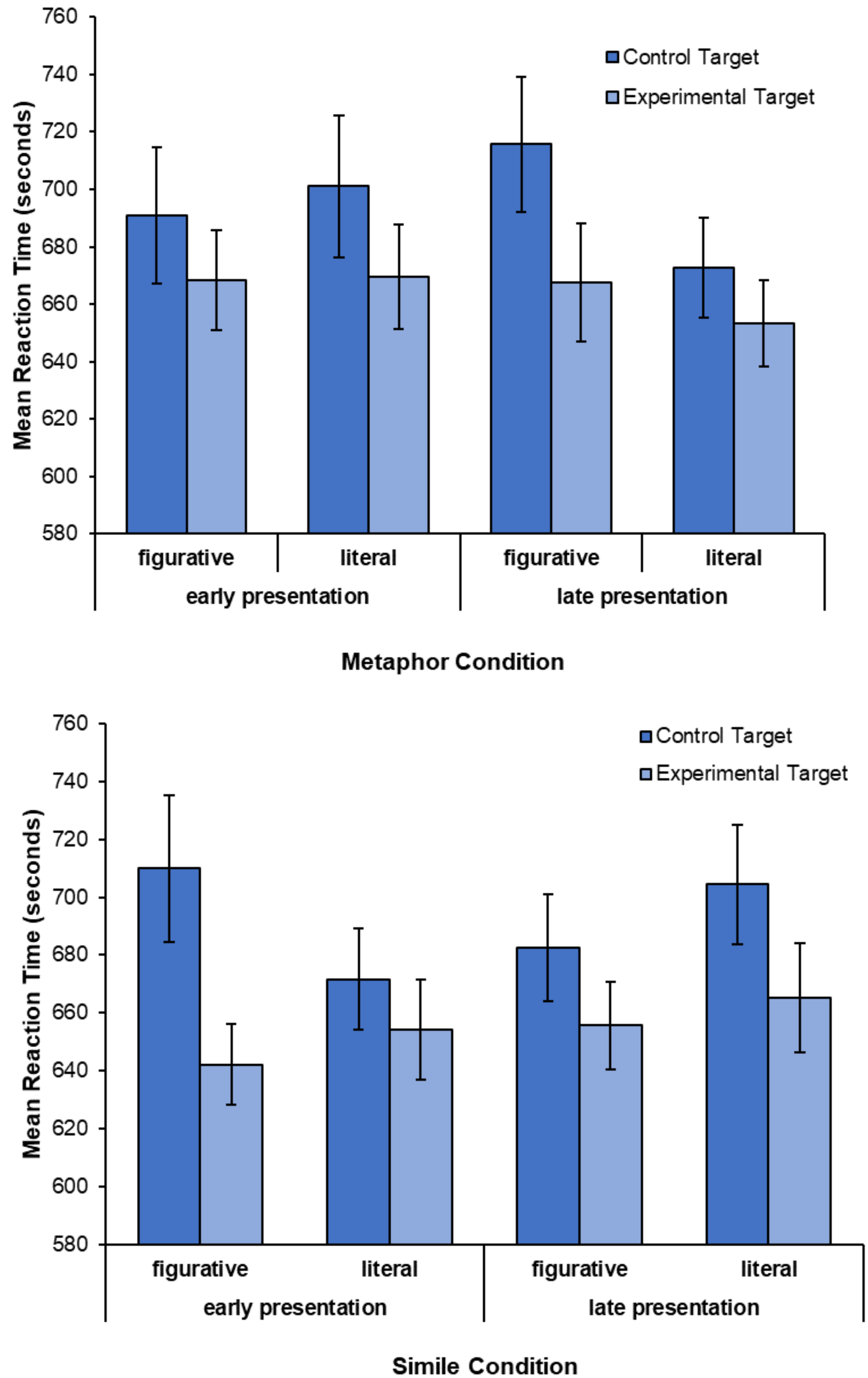

Figure 2. Mean RTs in Metaphor and Simile conditions as a function of time point and literality of visual target words and matched unrelated controls. Error bars represent SEM.

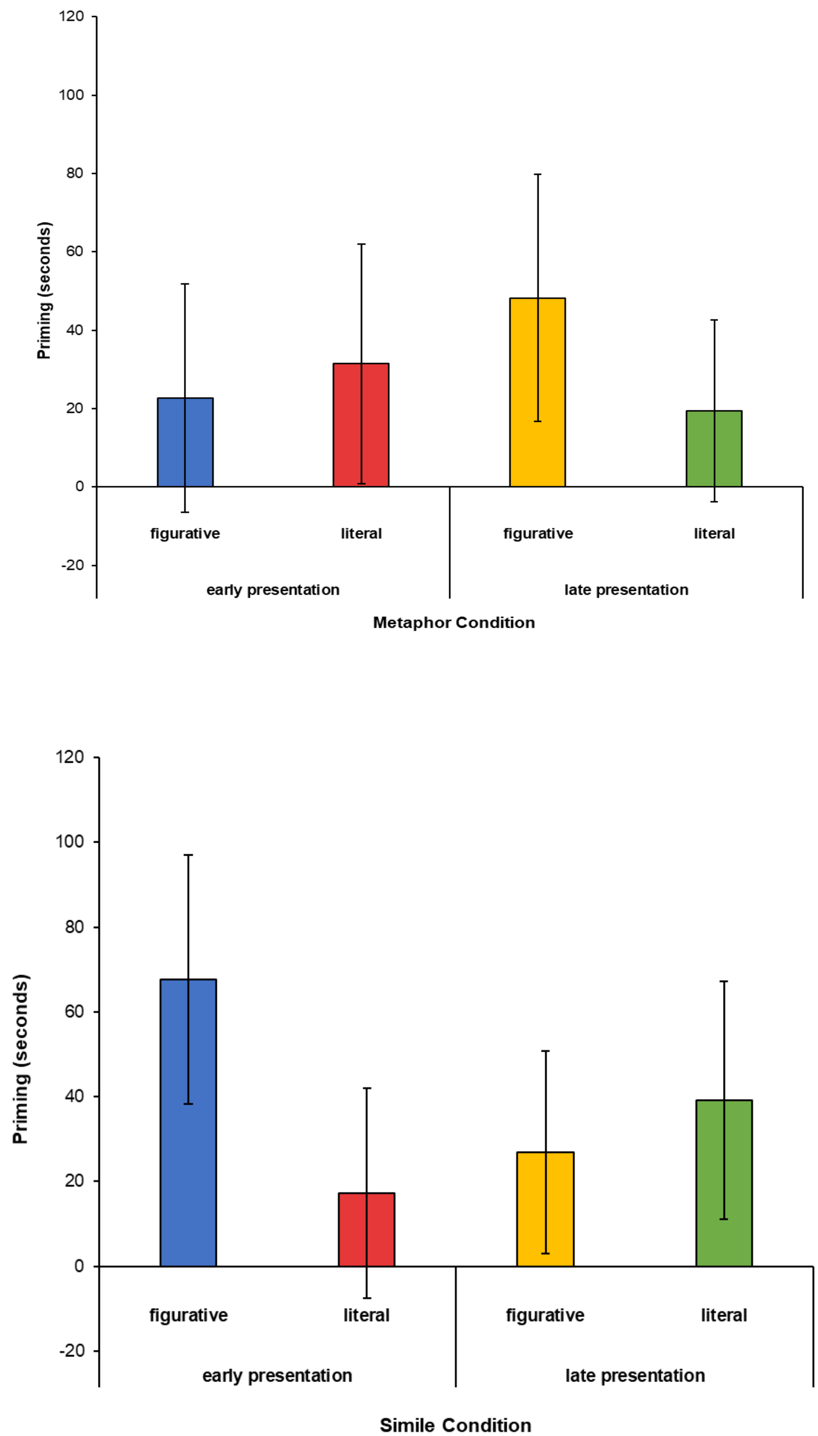

Figure 3. Mean priming effects between unmatched control words and related target words in Metaphor and Simile conditions as a function of time point and literality. Error bars represent SEM.

presented sentences; blocks where participants answered fewer than 60% of trials correctly; and participants whose mean RTs were more than 3 standard deviations away from the overall mean (following Friedmann et al., 2008). To correct for positive skew, RTs that deviated more than +/- 2 standard deviations from the mean were replaced with values 2 standard deviations from the mean prior to analysis. A visual inspection of the data confirmed that the data met assumptions of homoscedasticity but violated assumptions of normality. However, we chose to analyse raw RT scores rather than log or square root transformed data due to concerns that raw scores would be more informative about cognitive processes occurring during comprehension and transformed data can distort this data and make it more difficult to interpret (Lo & Andrews, 2015). All analyses were performed using the 'lme4' package (Bates, Maechler, Bolker, & Walker, 2015) in R (R Core Team, 2015).

Results

We performed a linear mixed-effects model regression analysis with subjects and items (vehicles) entered as random effects with random intercepts. Raw RTs were regressed on priming (control/experimental targets), sentence literality (metaphor/simile conditions), target type (figurative/literal) and time-point (early/late), as well as all first order interaction terms. For ease of interpretation, mean response times for each condition (and SEM) are presented in Figure 2. Priming effects are presented in Figure 3. The full RT model was compared to a null model including only random effects (subject and item), using the Likelihood Ratio Test to determine significance. Our model provided a better fit to the data than the null model ($\chi^2(10) = 25.70$, $p = 0.004$). We derived p-values for all main effects and interactions using the Likelihood Ratio Test to compare the full model to a model excluding the relevant term (see Table 2) and found only one significant main effect of priming.

Table 2.

Mixed-effects linear model of response times.

Predictor	Estimate	SE	t	95% CI	Null Comparison
Constant	718.19	23.88	30.08	[671.39, 765.00]	
Priming	-39.51	16.10	-2.45	[-71.06, -7.95]	$\chi^2(4)=22.38$, $p<.001$
Time-point	-1.84	16.05	-0.11	[-33.29, 29.61]	$\chi^2(4)=1.67$, $p=.80$
Target type	-16.44	16.15	-1.02	[-48.10, 15.22]	$\chi^2(4)=1.27$, $p=.87$
Sentence literality	-18.04	16.02	-1.13	[-49.44, 13.37]	$\chi^2(4)=2.31$, $p=.68$
Priming x Time-point	1.43	15.81	0.09	[-29.55, 32.40]	$\chi^2(1)=0.0083$, $p=.93$
Priming x Target type	8.09	15.87	0.51	[-23.02, 39.19]	$\chi^2(1)=0.26$, $p=.61$
Priming x Sentence literality	-4.50	15.83	-0.28	[-35.52, 26.52]	$\chi^2(1)=0.08$, $p=.78$
Time-point x Target type	2.01	15.78	0.13	[-28.93, 32.94]	$\chi^2(1)=0.02$, $p=.90$
Time-point x Sentence literality	14.28	15.77	0.91	[-16.64, 45.19]	$\chi^2(1)=0.82$, $p=.36$
Target type x Sentence literality	10.39	15.88	0.65	[-20.74, 41.52]	$\chi^2(1)=0.43$, $p=.51$

As predicted, participants took significantly longer to respond to unrelated targets than to related targets ($\chi^2(4) = 22.38$, $p < 0.001$) – overall, RTs to related targets were 40ms faster (*SEM*=23.88). While no other main terms or interaction terms reached significance, the respective means of each condition seemed to show trends which may be worth investigating with a larger sample. Specifically, in the metaphor condition, early priming values were lower for the figurative condition than for the literal condition, but priming for the figurative condition was higher at the later time point. In the simile condition, the reverse was true, with higher priming for literal targets at the late time point. Unexpectedly, the largest priming effect was observed for figuratively related targets at the early time point of the simile condition.

Discussion

Various theories have been proposed to account for how metaphors are understood. While early scholarly works on metaphor comprehension (what we have called *pragmatic* theories of metaphor processing) proposed that figurative sentences trigger "implicatures" (Grice, 1975) which allow a listener to understand a speaker's intended meaning from a literally false proposition (see: Grice, 1975; Davidson, 1978; Searle, 1979), more recent attempts to explain metaphor have assumed that figurative language is automatically understood as metaphorical. Proponents of direct-access models of metaphor comprehension (e.g., Gibbs, 1994; Wolff & Gentner, 2000; Glucksberg, 2003) have presented various accounts of strategies which would allow direct access to intended metaphorical meanings by the linguistic system, without drawing upon more nuanced *pragmatic* systems – what these accounts all have in common is the assumption that metaphorical meaning is directly accessed. However, empirical evidence given in support of this assumption has mainly been restricted to off-line judgment tasks (e.g., Glucksberg et al., 1982; Keysar, 1989; Wolff & Gentner, 2000), which fail to rule out the involvement of time-consuming *pragmatic* processes.

The goal of the present thesis was to gain insight into the moment-by-moment processing of metaphors and similes in the form *X is Y*/*X is like Y* by employing a time-sensitive cross-modal lexical decision paradigm. Following the *pragmatic* model, we hypothesized that recognition of target words would be facilitated (in both metaphor and simile conditions) at the recognition point (a) of an aurally presented vehicle word that was literally related, resulting in priming effects; however, we expected to see minimal priming effects for figuratively related target words at recognition point (a), as we hypothesized metaphorical meanings would not yet be accessible at the earliest stages of processing. Conversely, at the later time point (b), we hypothesized that priming effects would be larger for figuratively related targets than literally related targets in the metaphor condition, while literally related targets would continue to show priming effects in the simile condition (for a full list of predictions, see Table 1). The results obtained showed significant priming in all conditions and at all time points, but, contrary to our hypothesis, no statistically significant differences in priming between conditions were obtained.

This result can be interpreted as in part compatible with direct-access models of metaphor processing, as figurative targets were primed early on in both conditions. However, contrary to what both *pragmatic* and direct-access models would predict, literally related target words were still primed as much as the later time point (b) as at point (a), suggesting that even after a sentence has been fully processed (and, presumably, understood to have a non-literal intended meaning), literal representations of the vehicle word remain activated within the cognitive system. An account of metaphor processing that would best explain this result is the dual-processing model proposed by Carston (2010). According to the dual-processing model, two simultaneous processes contribute to the understanding of metaphorical language – a fast, *on-*

line formation of ad-hoc concepts linked to the metaphorical vehicle (for example, while the lexical item *shark* may conceptually represent the fish, it may also represent a concept like *aggressive*, especially for highly lexicalized metaphors like many of those used in our experiment), and a more nuanced, *off-line* process of interpreting the meaning of a metaphorical passage that relies on its literal meaning and the images the literal meaning evokes (Carston, 2010). Thus, according to Carston's (2010) model, the early priming of figuratively related targets presented at recognition point (a) could be a result of ad-hoc concept formation relating the vehicle word to figurative concepts, while the persistence of priming for literally related targets at point (b) could be explained by the persistent, simultaneous activation of literal representations.

A notable result was the lack of difference between similes (which we took to be literal) and metaphors. According to Carston’s (2010) model, ad-hoc concepts are created for metaphors in order to make their intended meaning comprehensible, but this might not be necessary for similes, which are literally comprehensible. However, we observed no difference in activation for figurative targets in the metaphor and simile conditions. A possible explanation for this result is that the word *like* in similes could lead participants to anticipate an upcoming vehicle word that is not typically literally related to the topic of the sentence. The gating paradigm used to determine recognition points tested the moment at which each word is recognized in isolation, but context could bias listeners to correctly identify the word earlier when presented within a sentence. In the context of highly familiar or lexicalized similes like *time is like money*, the word *like* could in fact trigger an assumption in the listener that the word *money* will follow, due to the frequency with which the simile is used in common speech and writing – and cause the recognition point of the vehicle word to occur earlier than anticipated. In order to test this possibility, additional experiments could be conducted relating the strength of the early figurative priming effect to the familiarity rating of each simile.

Alternatively, the word *like* in the simile phrase could lead the listener to think about things that are *like time*, and activation could reflect this rather than the vehicle word *money*. This would be less likely to occur when hearing a metaphor because there are many more plausible ways to continue the sentence (e.g., the sentence *time is passing* would be just as likely as *time is money*). Since the norms compiled by Roncero and de Almeida (2015) contained a high level of overlap between explanatory words generated for each metaphor and simile containing the same constituents, it is possible that fast response times to targets figuratively related to the vehicle reflect a relationship to the topic, rather than the vehicle, of simile sentences. This explanation is more consistent with the smaller priming effects for figurative targets at later time point (b), after the full sentence has been processed.

It is also relevant to note that, while literal targets were generally not descriptive of both the topic and vehicle of experimental sentences and reflected meanings of the vehicle in isolation, figurative targets were necessarily related to both topic and vehicle of any given phrase. This could account for general early activation of figurative targets even without the formation of ad-hoc concepts or direct access to figurative meanings, as there was a degree of literal relatedness between vehicle and figurative target in many of the sentence/target pairs we selected from Roncero and de Almeida's (2015) published norms (see Appendix 1 for a complete list of sentence/target pairs). Our rationale for using these targets was that they were commonly generated by participants in the norming studies, and thus more likely to reflect common understandings of metaphor meanings than researcher-generated targets designed to be as literally unrelated to the vehicle as possible. Despite some potential overlap in literal relatedness of the targets in each condition, we were careful not to select targets for the figurative condition which participants from the norming study (Roncero & de Almeida, 2015) had named as associates of the vehicle word in isolation – meaning that they were not what most commonly comes to mind when hearing the vehicle word, and thus their access should be facilitated to a lesser degree (in isolation) than the corresponding literally related target.

A major methodological difference between our study and other psycholinguistic experiments employing cross-modal lexical priming (e.g., Swinney, 1979; Friedmann et al., 2008) was our use of briefly presented masked visual targets. Typically, cross-modal lexical decision tasks employ an unmasked visual target presentation lasting at least 500ms (e.g. Friedmann et al., 2008), which allows for much higher response accuracy. In contrast, very brief (40-50ms) masked presentations of visual primes have mainly been employed in visual priming experiments where primes are morphologically or orthographically related to a visually presented target word (Forster, 1999). Forster (1999) explained that the use of very rapid masked primes should circumvent conscious thought processes about prime and target words and, instead, reflect unconscious processes of word association. Our use of masked visual targets presented for 80ms combined with presentation times at the recognition point of aurally presented vehicle words followed the rationale that in order to observe unconscious *on-line* access to semantically related concepts during metaphor processing, participants should not be allowed time to consciously consider either visual target or aurally presented vehicle. This created a speed-accuracy trade-off that resulted in a loss of useful data; however, the data obtained should be reflective of unconscious (*on-line*) facilitation processes.

In identity priming experiments (e.g., where the masked prime *###shark###* primes *SHARK*), priming effects of around 50ms have been consistently observed, while priming effects for morphologically or orthographically related words are generally smaller (Forster, 1999). Our masked lexical decision experiment obtained a mean priming effect of 40ms ($p < 0.001$) for related targets over unrelated controls, validating that experimental targets derived from Roncero and de Almeida's (2015) norms were strongly related to aurally presented vehicle words. Although our presentation time of 80ms was double that of morphological visual masked primes in experiments described by Forster (1999), the priming effect observed was large considering that target words were not morphologically or orthographically related to their corresponding vehicles.

It is possible that this novel rapid masked target presentation did not allow for full semantic composition of sentences (see: de Almeida & Lepore, 2017) prior to lexical decisions at the early time point (a), as the very brief presentation of target words may have forced lexical decisions before the full meaning of each metaphor or simile could be composed. If this is the case, a possible interpretation of our results is that the priming observed across early-presentation conditions reflects associations between the target word and the vehicle word in isolation, rather than the full sentence and its meaning. According to this interpretation, no difference should be observed between priming in metaphor and simile conditions, and any observed priming effects reflect automatic associations between the target and vehicle word. This interpretation is consistent with Carston's (2010) dual processing model, as associations between each vehicle and its corresponding figuratively related prime could reflect *ad-hoc concepts* formed by partial lexicalization of the metaphorical meaning of vehicle words used in the experiment, rather than an understanding that the sentence as a whole is meant to convey a metaphorical meaning as direct-access models suggest. Indeed, some evidence used to support direct-access theories is quite explicit in stating that long biasing contexts – and, therefore, awareness of an intended figurative meaning – are necessary to facilitate metaphor comprehension (e.g., Inhoff et al., 1984), while our experiment did not employ biasing contexts. If context was necessary to facilitate a figurative interpretation of a metaphor, we would not expect to observe priming in early conditions for metaphors and similes in isolation.

Conclusions

Multiple competing theories have been proposed to explain the processes involved in attaining meaning for metaphors which cannot be understood literally. While some theorists consider the issue resolved in favour of the direct-access theory of metaphor comprehension (e.g. Glucksberg, 2003), empirical investigations have failed to provide strong evidence for either *pragmatic* or direct-access accounts (de Almeida et al., 2010). Using a novel masked brief-presentation cross-modal lexical decision paradigm, the present thesis attempted to elucidate moment-by-moment cognitive processes occurring at the moment that a metaphor or simile in the form *X is (like) Y* is uttered. Results showed significant priming effects of both figuratively- and literally- related targets across all conditions (in metaphor and simile, at multiple time points), a result that we interpret as most consistent with Carston's (2010) dual-processing account of metaphor, which posits that *both ad-hoc* figurative concepts related to metaphorical vehicles and literal meanings should remain activated throughout sentence comprehension. Future directions for this line of investigation include investigating correlations between priming effects and familiarity/aptness ratings of each metaphor, to elucidate whether early activations of figurative meanings are due to lexicalization of certain metaphors or whether they are formed *on-line* according to a metaphor's explanatory value. While the results of our experiment do not provide strong evidence in favour of any particular theory of metaphor processing, they suggest that the issue of what happens in the brain during metaphor processing is far from resolved and requires more sensitive measures that tap into the earliest moments of metaphor processing.

References

Aristotle. (1926). The art of rhetoric. New York, NY: G. P. Putnam's Sons.

Ashby, J., Roncero, C., de Almeida, R. G., & Agauas, S. J. (2018). The early processing of metaphors and similes: Evidence from eye movements. *Quarterly Journal of Experimental Psychology, 71*(1), 161-168.

Bates, D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48. doi:10.18637/jss.v067.i01.

Blasko, D. G., & Connine, C. M. (1993). Effects of familiarity and aptness on metaphor processing. *Journal of Experimental Psychology, 19*(2), 295-308.

Bowdle, B. F., & Gentner, D. (2005). The career of metaphor. Psychological review, 112(1), 193-216. doi:10.1037/0033-295X.112.1.193

Carston, R. (2010). XIII—Metaphor: Ad Hoc Concepts, Literal Meaning and Mental Images. *Proceedings of the Aristotelian Society (Hardback), 110*(3pt3), 295-321. doi:10.1111/j.1467-9264.2010.00288.x

Chiappe, D., Kennedy, J.M. and Chiappe, P. (2003) Aptness is more important than comprehensibility in predicting recognition bias and preference for metaphors and similes. *Poetics, 31*, 51-68.

Columbus, G., Sheikh, N., Cote-Lecaldare, M., Häuser, K., Baum, S. R., & Titone, D. (2015). Individual differences in executive control relate to metaphor processing: An eye movement study of sentence reading. *Frontiers in Human Neuroscience, 8.*

Davidson, D. (1978). What metaphors mean. *Critical Inquiry, 5*(1), 31-47.

de Almeida, R. G., & Lepore, E. (2018). Semantics for a Module. In de Almeida, R. G. & Gleitman, L. R. (Eds.), *On Concepts, Modules, and Language*. New York: Oxford

University Press.

de Almeida, R., Manouilidou, C., Roncero, C., & Riven, L. (2010). Three tales of semantic decomposition: Causatives, Coercion, and Metaphors. In França, A. & Maia, M. (Eds.), *Papers in Psycholinguistics* (pp. 172-190).

Forster, K. I. (1999). The Microgenesis of Priming Effects in Lexical Access. *Brain and Language, 68*(1), 5-15. doi:https://doi.org/10.1006/brln.1999.2078

Friedmann, N., Taranto, G., P Shapiro, L., & Swinney, D. (2008). The Leaf Fell (the Leaf): The Online Processing of Unaccusatives. *Linguistic Inquiry, 39*(3), 355-377. DOI: 10.1162/ling.2008.39.3.355 ·

Gibbs, R. (1994). The Poetics of Mind: Figurative Thought, Language, and Understanding. New York: Cambridge University Press.

Giora, R. (2003). *On Our Mind: Salience, Context, and Figurative Language*: Oxford University Press.

Glucksberg, S. & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychological Review, 97*(1), 3-18.

Glucksberg, S. (2003). The psycholinguistics of metaphor. *Trends Cogn Sci, 7*(2), 92-96.

Glucksberg, S., Gildea, P., & B. Bookin, H. (1982). On understanding nonliteral speech: Can people ignore metaphors? *Journal of Verbal Learning and Verbal Behavior, 21*(1), 85-98.

Grice, H. P. (1975). Logic and Conversation. In P. Cole, & J. Morgan (Eds.), *Syntax and Semantics* (pp. 41-58). New York: Academic Press.

Inhoff, A. W., Lima, S. D., & Carroll, P. J. (1984). Contextual effects on metaphor comprehension. *Memory & Cognition, 12*(6), 558-567.

Janus, R., & Bever, T. (1985). Processing of metaphoric language: An investigation of the threestage model of metaphor comprehension. *Journal of Psycholinguistic Research, 14*(5), 473-487.

Keysar, B. (1989). On the functional equivalence of literal and metaphorical interpretations in discourse. *Journal of Memory and Language, 28*(4), 375-385.

Lai, V. T., Curran, T., & Menn, L. (2009). Comprehending conventional and novel metaphors: An ERP study. *Brain Research, 1284*, 145-155. doi:https://doi.org/10.1016/j.brainres.2009.05.088

LePore, E., Stone, M (2015). *Imagination and convention: Distinguishing grammar and inference in language*. UPSO eCollections (University Press Scholarship Online).

Lo, S., & Andrews, S. (2015). To transform or not to transform: using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*, 1171. http://doi.org/10.3389/fpsyg.2015.01171

Ortony, A., Schallert, D., Reynolds, R., & J. Antos, S. (1978). Interpreting metaphors and idioms: Some effects of context on comprehension. *Journal of Verbal Learning and Verbal Behavior, 17*(4), 465-477.

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. doi:https://doi.org/10.1006/brln.1996.0107

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Roncero, C., & de Almeida, R. G. (2015). Semantic properties, aptness, familiarity, conventionality, and interpretive diversity scores for 84 metaphors and similes. *Behavior Research Methods, 47*(3), 800-812. doi:10.3758/s13428-014-0502-y

Roncero, C., de Almeida, R. G., Martin, D. C., & de Caro, M. (2016). Aptness predicts metaphor preference in the lab and on the internet. *Metaphor and Symbol, 31*(1), 31-46. doi:10.1080/10926488.2016.1116908

Searle, J. (1979). Metaphor. In A. Ortony (Ed.), *Metaphor and Thought* (pp. 83-111). Cambridge University Press.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*(6), 645-659. doi:https://doi.org/10.1016/S0022-5371(79)90355-4

Tirrell, L. (1989). Extending: The Structure of Metaphor. *Noûs, 23*(1), 17-34. https://doi.org/10.2307/2215827

Wolff, P., & Gentner, D. (2000). Evidence for role-neutral initial processing of metaphors. *J Exp Psychol Learn Mem Cogn, 26*(2), 529-541.

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition, 32*(1), 25-64.

Appendix A

Experimental sentences and targets.

Metaphor/simile	Figurative target	Literal target
Genes are (like) blueprints	Maps	Construction
Salesmen are (like) bulldozers	Annoying	Big
Love is (like) a child	Innocent	Annoying
Minds are (like) computers	Intelligent	Expensive
Clouds are (like) cotton	White	Comfy
Eyelids are (like) curtains	Open	Colourful
Insults are (like) daggers	Hurtful	Blade
Typewriters are (like) dinosaurs	Antiques	Dangerous
Love is (like) a drug	High	Death
Anger is (like) fire	Scary	Flames
Families are (like) fortresses	Strong	Big
Trust is (like) glue	Binds	Liquid
Exams are (like) hurdles	Stressful	Height
Jobs are (like) jails	Boring	Dangerous
Cities are (like) jungles	Crazy	Animals
Knowledge is (like) light	Illuminates	Electricity
Music is (like) medicine	Helpful	Bitter
Time is (like) money	Important	Green

Wisdom is (like) an ocean	Vast	Blue
Beauty is (like) a passport	Advantage	Booklet
Fingerprints are (like) portraits	Unique	Beautiful
Faith is (like) a raft	Unsteady	Flat
Friendship is (like) a rainbow	Rare	Curvy
Life is (like) a river	Long	Blue
Teachers are (like) sculptors	Builders	Clay
Obligations are (like) shackles	Annoying	Heavy
Lawyers are (like) sharks	Mean	Blood
Lawyers are (like) snakes	Sneaky	Death
Memory is (like) a sponge	Soaking	Dirty
Education is (like) a stairway	Upward	Exercise
Time is (like) a thief	Quick	Dangerous
Heaven is (like) a treasure	Pleasant	Expensive