DO YOU HAVE A "STRICT PURSE"? THE ROUTES TO MEANING IN METAPHOR

By

JESSIE. E. STEWART

A thesis
submitted to the Victoria University of Wellington
in fulfilment of the requirements for the degree of
Masters of Science in Psychology

Victoria University of Wellington

2012

Abstract

Theoretically there are two processing systems through which meaning can be found for a given statement: an effortless, associative processing system (meaning retrieval), or an effortful, analytical processing system (meaning construction). The current study investigated whether or not the context in which target (loosely figurative) word-pairs are presented can influence whether a person relies on associative or analytical processing to find their meaning.

Participants were presented with target (loosely figurative) novel word-pairs and asked to judge them for meaningfulness. These target novel word-pairs were presented in different contexts: either mixed with clearly meaningful word-pairs or with additional novel word-pairs. By nature, meaning cannot be retrieved for novel word-pairs, so if a novel word-pair is to be found “meaningful,” then its meaning must usually be constructed online (via the analytical processing system).

Consistent with increased reliance on analytical processing, participants who saw target novel word-pairs mixed with additional novel word-pairs judged them meaningful more often than did participants who saw them mixed with clearly meaningful word-pairs. Participants who saw target novel word-pairs mixed with additional novel word-pairs also had more negative N400s to target novel word-pairs, indicating that they committed more semantic effort to the processing of these (again consistent with analytical processing).

Associative processing does not involve attempts to construct new meaning for given word-pairs. Consistent with increased reliance on associative processing, participants who saw target novel word-pairs mixed with clearly meaningful word-pairs judged them meaningful less often than did participants who saw them mixed with additional novel word-pairs. These participants also had less negative N400s to target novel word-pairs, indicating that they committed less semantic effort to the processing of these (again fitting with associative processing).

Further evidence for different contexts leading to differential processing of the same target novel word-pairs was provided by examination of wave morphology. Two distinctive patterns of neural activation were found in response to the same target novel word-pairs, differing depending on the context in which these appeared.

Overall, the results of the current study were consistent with the hypothesis that context can influence which processing system is relied upon to find meaning for a given statement. This finding challenges contemporary models of meaning

construction and metaphor comprehension by showing that context is essential to these processes and needs to be taken into consideration.

Table of Contents

List of Tables

List of Figures

Do You Have a “Strict Purse”? The Routes to Meaning in Metaphor

Novel statements have, by nature, no stored meanings in the semantic system. Therefore, their meaning must necessarily be constructed online (Lynott & Connell, 2010). People encounter numerous novel statements throughout their lives, and meaning construction allows them to conceptualize and make sense of these (D'Argembeau & Mathy, 2011); therefore the meaning construction process is essential to human experience. Additionally, novel concepts could not be relayed and interpreted without the meaning construction process, and therefore it is also essential to human communication (Lynott & Connell, 2010).

Because meaning dwells in the perceiver rather than in statements themselves (Radden, Kopcke, Berg, & Siemund, 2007), statements are meaningless without someone to interpret them (for a review see Cornejo, 2008). Even when there is someone to interpret a statement, how and if they sucessfully do so can rely on a variety of factors. One factor which influences the meaning construction process is context. Both semantic and non-semantic aspects of context likely have influences on meaning construction. The influences of semantic aspects of context on meaning construction have been well studied (e.g. Pynte, Besson. Robichon & Poli, 1996). However, the influences of non-semantic aspects of context on meaning construction have not been well studied. Therefore, the ways in which certain non-semantic aspects of context influence meaning construction were investigated in the current study

Metaphors

> “*Yes, metaphor. That's how the whole fabric of mental interconnections holds together. Metaphor is right at the bottom of being alive*." (Gregory Bateson, as cited in Fritjof Capra, p.76-77, 1988).

A metaphor is a form of analogy whereby one expression is used to refer to another in order to express the similarities between the two. Consider the example “*my kitten is a lioness”*. The ‘topic’ of a metaphor is the subject of interest (“*my kitten*”), and the ‘vehicle’ of a metaphor is its comparison (“*a lioness*”; Pynte et al., 1996). The ‘ground’ of a metaphor refers to the shared properties between the topic and the vehicle (Richards, 1936), and therefore indicates the inference which should be drawn. For example, the person who said the above metaphor likely had a particularly brave and ferocious kitten (attributes which are signature of a lioness). These shared attributes are the ground of this metaphor.

In the quote at the beginning of this section, Bateson rightfully implied that metaphor is essential to abstract reasoning abilities and the communication of abstract concepts for humans (Kovecses, 1988; Lakoff & Johnson, 1980; Quinn, 1987). A variety of studies have shown that metaphors are important for expressing and reasoning about different types of abstract thoughts (see Yang, Fuller, Khodaparast & Krawczyk, 2010). Abstractions including reasoning about commitment (Quinn, 1987), the learning of difficult scientific subjects (Diehl & Reese, 2010), understanding the concept of time (Clark, 1973; Gentner & Imai, 1992) and reasoning about emotions such as love (Kovecses, 1988) have all been linked to metaphor processing. The theory of embodied cognition relates metaphor to abstract thinking, and claims that much "abstract thought is generated by metaphorical projection from the domain of embodied experience" (Lakoff & Johnson, 1980).

Because they are important for the communication of abstract thoughts, metaphors are prolific in spoken language (e.g. see Graesser, Long, & Mio, 1989). Some metaphors are shared by all speakers of a language (e.g. "*time flies*"). However, more often metaphors are novel or unique. The meaning of novel metaphors must be constructed online (Bowdle & Gentner, 2005).

Theories of Metaphor

As clearly described by McGlone and Manfredi (2001), the way in which we cognitively process metaphors has been studied and discussed since the days of Aristotle's "Poetics" (around 335 BC), in which Aristotle proposed that metaphors are processed as implicit comparisons, or similes. The following section will discuss some theories of metaphor comprehension and provide a critical assessment of where we now lie in terms of understanding the underlying cognitive processes that support the understanding of metaphor.

Contemporary theories of metaphor derive from two theoretical viewpoints: direct- and indirect- access models of metaphoric meaning. Indirect- access models of metaphor processing are those which hold that only once literal meaning has already been accessed and rejected, can metaphoric meaning be processed (Grice, 1975; Searle, 1979). A pragmatic "figurative process" is employed (Goodman, 1972), and the metaphor is reinterpreted as a simile (for example the metaphor "*my kitten is a lioness*" would be reinterpreted as "*my kitten is <u>like</u> a lioness*"). This is known as comparative processing. However, there is evidence against indirect- access models: sometimes metaphoric meanings can be retrieved just as fast as literal meanings

(Harris, 1976; McElree & Nordlie, 1999). Metaphoric meanings can also interfere with literal meanings (Glucksberg, Gildea & Bookin, 1982). These findings suggest that literal meanings are not always accessed before metaphoric meanings.

Direct-access models, in contrast, hold that metaphoric meanings can be accessed directly, without having to first access literal meanings. Glucksberg and Keysar (1990) proposed the class-inclusion direct-access model of metaphor: that when a metaphor is encountered, the topic is directly processed as representing the super-ordinate category of which it is the chief exemplar. So in the example of "*my kitten is a lioness*", the vehicle ("*lioness*") is processed as the abstract category for which it best stands (*things that are brave and ferocious*), rather than as its basic-level meaning (*a female lion*). This is known as categorical processing. However, the findings of Gentner and Wolff (1997) were problematic for this view. Gentner and Wolff primed participants with either a metaphor vehicle or topic. The vehicle but not the topic of a metaphor can activate the super-ordinate category representing the metaphors ground. Therefore, if categorical processing occurs, metaphoric processing should be facilitated when participants were primed with the vehicle compared to when primed with the topic. This was found for conventional metaphors but not novel metaphors, indicating that conventional but not novel metaphors can be processed as categorizations.

Based on their findings, Gentner and Wolff (1997) proposed the Contemporary Theory of Metaphor, which combines indirect- and direct- access models, suggesting that novel metaphors are processed as comparisons (as in indirect-models) but that conventionalized metaphors are processed as categorizations (as in direct-models). Evidence supporting this theory was provided in a series of experiments by Bowdle and Gentner (2005). Participants were found to prefer the categorization form ("*my kitten is a lioness*") more for conventionalized than for novel figurative statements. However, when novel figurative statements were presented multiple times, participants were shown to shift from preferring them in the comparison to the categorization form, thus supporting the cognitive shift predicted by Contemporary Theory of Metaphor.

Several other researchers agree that metaphor can be processed in two ways: as comparisons or as categorizations, but disagree with Contemporary Theory of Metaphor on what specifically determines which process is utilized. For example, Utsumi (2007, 2011) proposed that rather than familiarity, a metaphor's "interpretive diversity" (essentially meaning semantic richness) is what determines whether it is processed comparatively or categorically. Glucksberg and Haught (2006) argued that whether comparative or categorical processes are used to process a particular metaphor depends on its aptness (its particular "semantic and referential properties", p. 935). Interestingly, Glucksberg and Haught used only novel metaphors in their study, and, challenging for the view of Gentner and Wolff (1997), found these could be processed comparatively or categorically depending on whether they had a literal or figurative referent. Literal referent metaphors are metaphors whose vehicle refers to a literal concept. For example, in the metaphor "*my kitten is a lioness",* the vehicle *"lioness"* has a literal meaning. In contrast, figurative referent metaphors are metaphors whose vehicle refers to a figurative concept. For example, in the metaphor "*my kitten is a resentful lioness*", the vehicle "*resentful lioness*" has a figurative, but no literal, meaning. This theoretical perspective shall be referred to as "aptness theory". Despite the complexity of contemporary metaphor theories, one essential factor has yet to be considered to influence the selection of metaphor processing system: context.

Context and Metaphor

While features unique to each metaphor (such as its conventionality) seem to contribute to its processing, so surely must the context in which it is presented. In real life, metaphors are always presented within a context, and so contemporary theories of metaphor must consider context as well as the lexical properties of a metaphor. Consider the following two scenarios:

Scenario one.

Lizzie: "*My kitten is six-weeks old"*

Stuart: "*My kitten is a lioness"*

Scenario two.

Lizzie: "*My kitten is very timid"*

Stuart: "*My kitten is a lioness"*

The lexical properties of the metaphor "*my kitten is a lioness*" remain the same in both of these scenarios, however, due to relevant semantic information provided in the context (Lizzie and Stuart are talking about the personality traits of their kittens), the ground of the metaphor is more apparent in scenario two. This effect has been shown experimentally. Ortony, Schallert, Reynolds and Antos (1978) found that participants were faster to respond to a metaphor when it was preceded by a long context-setting sentence than when it was preceded by a short context-setting sentence. This suggests that the amount of relevant semantic information in the context within which a metaphor is presented may contribute towards the ease of constructing its meaning. Similarly, Pynte et al. (1996) found that regardless of whether a metaphor was conventional or novel, less cognitive effort was required to process it if it was preceded by a relevant (as opposed to an irrelevant) context.

In some cases, however, the context surrounding target metaphors is less deliberate than that described above. Unless a metaphor is presented in complete isolation, then it is inevitably presented in the context of other statements, which in experimentation can be referred to as the "experimental context". Experimental context includes non-semantic aspects of context, such as the different types of stimuli, and the blocking and proportions of these within an experiment. In some experiments, experimental context is manipulated by using blocked versus mixed designs, or by changing the ratios of the different types of stimuli included within a condition (e.g. see Risko, Blais, Stolz & Besner, 2008).

Metaphor studies inevitably present metaphors in the experimental context of other types of statements. For example, in the previously mentioned study by Glucksberg and Keysar (1990), participants saw nine types of experimental sentences: "metaphors, metaphor comparison and literal comparisons, in original, noun-reversed and noun-phrase reversed order" as well as "filler items" (p. 54). While these word-pairs were not specifically semantically related to one another, they still may have influenced the processing and interpretation of one another. For example, experimental context may have lead participants to rely more on meaning construction or on meaning retrieval processes, each of which could provide different meanings for a given statement. The effects of differential experimental contexts on meaning construction for metaphors have not yet been investigated; they will therefore be the topic of the current thesis.

In the current study, different types of word-pairs were presented randomly, and none were deliberately semantically related to each other. Whether or not people interpreted the same novel (loosely figurative) word-pairs differently depending on experimental context was investigated. A review of the literature suggests that experimental context can influence the processing of metaphors in three different ways: by affecting which information processing system will be used to process them, by causing a bias in decision-making, and via process-priming. These three processes will be described in the following sections.

Information Processing

Meaning can either be retrieved or constructed (DeLong, 2010). Which of these processes is employed for a particular statement may depend on the experimental context in which it is presented. A related theory was proposed by William James (1890/1950): the dual processing theory. This theory holds that there are dual information processing systems: one which is associative, effortless and fast; the other which is analytical, effortful and slow (James, 1890). James proposed that the associative system comes into play when a person can access stored past experiences, allowing them to compare the current situation with images or abstractions related to those past experiences. Therefore, when processing conventional metaphors, people may employ the associative system. In contrast, James proposed that the analytical system comes into play when true reasoning “helps us out of unprecedented situations” (p.330). In essence, this means that James’s analytical system is required for meaning construction, and therefore may be employed when processing novel metaphors.

James’s dual-coding theory has been reinterpreted and extended in various ways. For example, the two information processing systems have been redefined as “probability-governed” versus “rule-based” systems (Sloman, 1996), as “intuition” versus “reasoning” (Kahneman, 2003), and in terms of language processing as “linguistic” versus “simulated” systems (Lynott & Connell, 2010). Regardless of subtle differences, dual-processing theories of information processing maintain that there are dual systems: one effortless, the other effortful.

Various factors can influence how confident a person is with a decision, and thus how willing they are to employ the more effortful information processing strategy rather than the effortless system. For example, factors such as personal relevance (e.g., Chaiken, Giner-Sorolla, & Chen, 1996; Petty & Cacioppo, 1986) and accountability (Tetlock & Lerner, 1999) have been shown to affect whether or not a person will bypass the effortless strategy in favour of the effortful strategy. It is possible that experimental context may also influence how willing a person is to employ the effortful rather than the effortless system.

One product of experimental context has already been proposed to influence which system a person will utilize. Alter, Oppenheimer, Epley and Eyre (2007) suggested that the level of fluency that is experienced in a decision will influence the selection of information processing strategy. Fluency refers to the ease with which something is processed (Oppenheimer, 2008). Essentially, the easier something is to process, the more "fluent" it is. Fluency can be manipulated in a variety of ways- many of which cause differing experimental contexts to arise (e.g. degree of text legibility). Across four studies, Alter et al. (2007) demonstrated that participants were more likely to engage the analytical system in conditions which were disfluent compared to in conditions that were relatively fluent. The dichotomization of fluency-disfluency in these experiments was achieved in several different ways. One manipulation of fluency was "ease-of-reading" (by changing fonts, contrasts and font-size). Another manipulation of fluency was related to embodied cognition: participants were asked either to furrow their brows (body language consistent with disfluency), or to puff out their cheeks (unrelated to fluency). The last manipulation of fluency involved manipulating the apparent competency of a person presenting the information to be evaluated. Across the studies by Alter et al. (2007), engagement with a particular information processing system was measured by participants' answers. The experiments were designed so that participants would have to disengage from the associative system and use analytical thinking to respond correctly. Across all experiments, participants who were in disfluent conditions gave more correct answers suggesting that they engaged their analytical system more often. Alter et al. (2007) reasoned that because more elaborate reasoning processes are required for more difficult tasks, participants who found their tasks hard used disfluency as a signal to engage the analytic processing strategy.

In a proposition fitting with the results of Alter et al. (2007), Lynott and Connell (2010) suggested that task demand can influence which strategy a person engages for understanding language. Lynott and Connell developed a model that specifically addresses dual processes in language, the Embodied Conceptual Combination model (ECCo). In the ECCo model, Lynott and Connell proposed that there are two information processing systems with which we process language: the linguistic (effortless) system and the simulation (effortful) system. Lynott and Connell apply their model to the example of a task involving relatedness judgements to word-pairs. According to the tenets of ECCo, if a person is to make a "relatively shallow" judgement about word-pairs, then the linguistic system can provide a "quick and dirty shortcut" (Lynott & Connell, 2010, p. 4). This shortcut is a heuristic based on the statistical history of how often two words have been seen in combination in the past. If the two words have been frequently paired, then the linguistic heuristic will cause a person to find them sensible. In contrast if two words share no “statistical distributional history” then the linguistic system will not find them sensible (p. 4).

If deeper processing is required due to higher task demand, then according to ECCo, a person will rely on their simulation system rather than their linguistic system. According to ECCo, the simulation system activates neural processes involved with motor, affective, and perceptual information, and thereby simulates the experience of the two concepts which are being combined. Simulation involves meshing (the integration of key concepts) and affordances (“ways in which a concept offers opportunities for meshing with other concepts”, Lynott & Connell, 2010, pg. 5). A deconstructive process or non-deconstructive process is applied to the concepts being combined. A deconstructive process strips some information away from one of the concepts being meshed (e.g., “whale seal” strips the word “whale” down to simply meaning “large”). A non-deconstructive process leaves all information about both of the concepts being meshed intact. Deconstructive and non-deconstructive processes constrain the amount, and quality, of potential interpretations of any given word-pair. According to ECCo, semantic aspects of context will influence meaning construction by simulating relevant neural pathways prior to stimulus onset, so that stimulus processing is speeded if context is semantically relevant. Non-semantic aspects of context (specifically task demand) will influence which system (linguistic or simulation) is relied on.

In the current study it is possible that experimental context may affect how easy (or fluent) a person perceives the task of judging a word-pair for meaningfulness. In turn, which system is used to process a word-pair may depend on how easy (or fluent) it is to judge its meaning. If the task demands of judging meaningfulness for a word-pair are low or fluent (e.g. if there was little pressure to find meaning), then in accordance with Alter et al.’s findings and the premises of the ECCo model, a more effortless system of processing might be used. In contrast, if the task demands for judging meaningfulness for a word-pair are high or disfluent (e.g. if there was high pressure to find or construct a meaning), a more effortful system of processing might be used.

If an effortless processing system is applied to a metaphor, then according to ECCo, simple heuristics will be applied and the metaphor will be considered meaningful if the two concepts being blended share "statistical distributional history", but meaningless if they do not. Descriptively, this linguistic system seems to share some properties with the categorization process proposed by Gentner and Wolff (1997) to be used in the processing of conventionalized metaphors. Gentner and Wolff suggest that if a vehicle has been previously processed metaphorically enough times, it will become associated with a super-ordinate category meaning. Therefore, when processing categorically, a person simply activates the link between the vehicle and super-ordinate category, rather than constructing a new meaning online. The idea of this link between the vehicle and super-ordinate category proposed by Gentner and Wolff (1997) may correspond to the idea of shared "statistical distributional history" proposed by ECCo.

In contrast, if an effortful system of processing is applied to metaphor, then the metaphor would be processed analytically. ECCo predicts that a person who applied an effortful system of processing to a metaphor would need to simulate the experience of both concepts being blended. Successfully blending the simulations of both concepts in a metaphor would mean that the metaphor would be judged meaningful, however, if simulations of these concepts are not successfully blended then the metaphor would be judged not-meaningful. In terms of metaphor processing, perhaps the simulation system shares some properties with the comparison process purported by Gentner and Wolff (1997) to be used when processing novel metaphors. This comparative process is said to involve aligning representations of a metaphor's topic and vehicle, and searching for commonalities. This idea may be equivalent to the idea that the simulation process attempts to mesh, or integrate, key concepts. Novel metaphors should be judged more meaningful when the simulation rather than the linguistic system is relied on, because they should less commonly contain words which share "statistical distributional history" than they should contain words which have potential for meshing.

Despite the fact that James's dual processing theory and that of the ECCo model have their differences, they largely make the same predictions in terms of constructing meaning for novel metaphors. For ease, the effortful systems of both will therefore be referred to as the analytical system, and the effortless systems of both will be referred to as the associative system. When predictions do differ between these two theories, the authors will be mentioned to allow readers to identify which theory is being discussed.

Biases in Decision Making

Another process which may influence how novel word-pairs are assessed occurs at the decision-making stage. As previously discussed, interpretations of novel word-pairs may be influenced by experimental context that leads to the selection of different information processing strategies. Experimental context may also, or alternatively, cause people to use different decision-making criteria when assessing novel word-pairs for meaningfulness. For example, imagine you are participating in an experiment in which you have to decide if each item you see is meaningful or not. The items are dubious; none appear to be particularly meaningful. However, the experimenter has led you to believe that some of the items should be meaningful. You would probably feel under pressure to say “meaningful” to the dubious items some of the time. In contrast, imagine if you saw these same dubious items combined with other, obviously meaningful items. In this case you could just say “meaningful” to the obviously meaningful items, and therefore not feel compelled to call the dubious items “meaningful”. Therefore, although you may perceive the dubious items in the same way in both scenarios (as “not-meaningful”), your response to them is likely to differ depending on the experimental context in which they are presented. In this example, although there is not likely to be a change in interpretation between scenarios, a shift in decision-making criterion is expected. Such a decision-making shift is not always conscious. Shifts in decision making are often referred to as biases (e.g. see Green & Swets, 1966), so that a person might be biased toward the “meaningful” or “not-meaningful” response. It is possible that a decision-making bias may occur in the meaning construction process due to experimental context.

A theory that deals with decision-making bias is Signal Detection Theory (SDT; Green & Swets, 1966). SDT was originally designed to investigate perceptual sensitivity (Green & Swets, 1966), although it has since been applied to a variety of domains including recognition memory (e.g. Rhodes, Parkin, & Tremewan, 1993), and metaphor processing (e.g. Mashal & Faust, 2008). Classical cognitive experiments measure both reaction time and accuracy, and consider a person to be accurate if they correctly detect the presence of a signal (a ‘hit’), and to be inaccurate if they do not correctly identify the presence of a signal (a ‘miss’). However, this measure of accuracy does not account for the fact that people are also accurate if they correctly identify the absence of a signal (a ‘correct rejection’), and that they are also inaccurate if they say a signal is present when it is not (a ‘false alarm’). This means that whether a person's performance reflects sensitivity (the ability to discriminate between signal present and signal absent trials) or bias (the increased likelihood to select one response over the other) cannot be determined from traditional measures alone. Application of SDT to data is useful because it considers ‘false alarms’ as well as ‘hits’, and therefore can tell researchers what type of errors (and correct responses) are made, and therefore what kinds of decision making strategies participants are using.

Signal detection theory holds that if a signal is present, then the corresponding neural activation will be made up of noise (random variations in activation) and signal (activation that occurs to indicate that signal is present). If no signal is present, then the corresponding neural activation will be made up purely of noise. Therefore, in a competent participant, neural activation to signal-present trials should usually exceed activation to signal-absent trials. In a non-biased participant, a criterion threshold should be set exactly half-way between the ‘signal’ and ‘noise’ distributions so that the participant answers ‘signal’ to anything which causes activation above this criterion threshold, and ‘noise’ to anything which does not cause activation above this criterion threshold. See the graph attached in the Appendix A for a pictorial explanation. This strategy would maximize the number of hits, and minimize the number of false alarms. However in a biased participant, the position of the criterion threshold relative to the two distributions is shifted. This means that a person is more likely to select the response opposite to the direction of the shift. If the criterion is shifted towards the ‘noise’ distribution then a participant is more likely to say ‘signal’, and therefore will make more hits and more false alarms. If the criterion is shifted towards the ‘signal’ distribution than a person is more likely to say ‘noise’, and therefore will make fewer hits and fewer false alarms. This increased likelihood of selecting one response over the other is the participant’s ‘bias’. For a more detailed explanation see Lauwereyns, (2010).

For the current study, it is important to note the fact that SDT is a model of decision-making *strategies.* This means that the participant need not be 100% sure of their answer, but that they can adopt a certain criterion threshold level (based on a number of rational factors) and use this to come up with the best possible answer to avoid errors. As well as SDT, other models of decision-making exist (e.g. the LATER model, Reddi & Carpenter, 2000), In all of these models, however, decision-making can be biased, and in all of these models, bias leads to the preference for one response over another. Decision-making theories are also consistent in their premise that bias is established before the required response is known, and often even before the relevant stimuli have been presented. In the current study, as well as affecting information processing strategy selection, experimental context might also influence participants' criterion for meaningfulness, and therefore cause participants to approach novel word-pairs more predisposed (or "biased") to classify them as meaningful or not-meaningful. For example, if experimental context involves a lot of word-pairs that are only loosely connected and a participant is therefore repeatedly judging them all as not-meaningful, they may shift their criterion so that they require less evidence to call a word-pair meaningful.

Process Priming

When a stimulus (such as a metaphor) is encountered, a particular processing strategy is activated. If a similar stimulus is then encountered, its processing might be facilitated because the appropriate strategy remains activated (process-priming; Inhoff, Lima & Carroll, 1984). For example, if process priming exists then participants should be faster to process the metaphor "*my kitten is a lioness*" after the metaphor "*my baby is a lamb*" than after the literal statement "*my baby is young*". In metaphor experiments (as was previously described in the example of the study by Glucksberg & Keysar, 1990), target metaphors are often presented in blocks mixed with a variety of other types of statements (the experimental context). If the experimental context includes metaphors which facilitate the processing of subsequently presented metaphors, then process priming would present an example of how experimental context can influence the meaning construction process.

Although the existence of process priming is intuitively appealing, it has little empirical support. Some support for metaphoric process-priming was found by Inhoff et al. (1984) because in their study, metaphoric target statements were read more quickly when they followed metaphoric than literal context-setting paragraphs. The same results were found when context-setting single sentences were used instead of paragraphs. As less schematic information is available in single sentences than in paragraphs, Inhoff et al., (1984) suggested that their results supported process- rather than semantic- priming. However, Inhoff et al. ran an additional analysis to see if metaphoric context setting sentences could be primed by a metaphoric target that was presented in the previous trial. Since these sentences and targets were not in the same trial, their meanings were unrelated and so any facilitation could not attributed to semantic priming (and therefore must represent process priming). It was found that metaphoric context sentences which were preceded by metaphoric targets were read faster (245ms per word) than those preceded by literal primes (264ms per word). However, this difference was not statistically significant, and Inhoff et al. concluded that while denunciation of the process priming hypothesis was premature, future research was required.

Only one study has followed up that of Inhoff et al. (1984). Blasko and Briihl (1997) presented participants with unrelated literal, unrelated metaphoric, related literal or related metaphoric sentences prior to target metaphors. Participants were found to be significantly facilitated at reading target metaphors when they first read related metaphoric but not when they first read unrelated metaphoric sentences. The authors therefore conclude that the "results do not support a pure process-priming hypothesis" (p. 268). Nevertheless, the total reading times for target metaphors when participants first read unrelated metaphoric sentences were intermediate between the reading times for target metaphors when they were first preceded by related metaphoric and when they were first preceded by unrelated literal sentences. Therefore, metaphoric process-priming may have caused "some benefit" (p. 268).

Studies of metaphoric process-priming are scarce and whether or not it is a true phenomenon remains unresolved. Previous research has focused on immediate effects of process-priming: whether processing a metaphor will facilitate processing of an immediately subsequent metaphor. However, it is possible that reading metaphors that are interspersed throughout an experiment (i.e. the presence of metaphors in an experimental context) will affect the way in which other stimuli in the experiment are processed via a more global version of process-priming (getting into a "metaphoric way of thinking"). This concept of global process-priming will be investigated in the current thesis.

Goals

As was demonstrated in the quote "*my kitten is a lioness*", whether or not a meaning is constructed for a given phrase can depend on the context in which it is presented. This is because meaning is subjective and largely influenced by context (e.g. Pynte et al., 1996). However, the influence of non-semantic aspects of context on meaning construction has been largely neglected in the literature. Therefore, the present study was conducted to examine the influences of experimental context (one particular non-semantic aspect of context) on the meaning construction process and on metaphor processing. Participants in both experiments one and two of this study were presented with target novel word-pairs interspersed amongst other types of word-pairs which provided the experimental context. Critically, since the other word-pairs were not related to the target novel word-pairs in any systematic way, any influences of experimental context were non-semantic. Participants were asked to judge all word-pairs for meaningfulness. Specifically investigated were: 1) whether contextual effects of task demand or fluency can act as a mechanism through which information processing strategies are selected, and therefore influence whether meaning construction is accomplished; 2) whether experimental context can shift the criterion for meaningfulness; and 3) whether context can create a "figurative way of thinking" that leads participants to attempt figurative (rather than literal) meaning construction. These three possibilities are not mutually exclusive. Behavioural data was recorded in experiment one, and to follow up the results of experiment one, electrophysiological data was recorded using electroencephalography (EEG) in experiment two.

Experiment One

The first study was conducted to determine if experimental context could affect whether or not, and if so in which way, people create meaning for novel word-pairs. Participants saw novel and other word-pairs, and judged them for meaningfulness. The novel word-pairs used were randomly paired adjective- noun couplets, and a fair amount of effort was required to interpret these (if they even could be interpreted at all). Because no literal meaning existed for these word-pairs, a "meaningful" judgement necessarily reflects metaphoric meaning construction. The novel word-pairs were presented interspersed amongst different types of other word-pairs in different conditions. These other types of word-pairs therefore presented different experimental contexts.

Participants were assigned to different conditions. In each of these conditions the participants saw the same 75 novel word-pairs (e.g. "strict purse") randomly mixed with 25 other word-pairs. In the first condition, called the "conventional condition" the 25 other word-pairs were conventional metaphoric pairs (e.g. "old flame"). In the second condition, called the "literal condition" the 25 other word-pairs were literal pairs (e.g. "curly hair"). Participants in these two conditions were told that "In the past people have found *half* of the word-pairs meaningful". Participants in the third condition saw the same 75 novel word-pairs with an additional 25 novel word-pairs and were given the same verbal information. This condition will be referred to as the “novel50% condition”. In the fourth condition participants saw the same word-pairs as participants in the third condition. These participants were told that “In the past people have found *one-third* of the word-pairs meaningful”. Therefore, this condition will be referred to as the “novel33% condition” The reason that participants were given different verbal information will be discussed below.

First, in order to examine the role of fluency/ task demand on selection of the information processing strategy used to find meaning, comparisons were made across all conditions. Conditions where novel word-pairs were mixed with clearly meaningful word-pairs (conventional and literal conditions) involved the presence of some easy meaningfulness decisions. This is because the clearly meaningful word-pairs used should have been experienced by participants previously, and so their meanings should have been easy to retrieve. The presence of clearly meaningful word-pairs may therefore have provided a model of how easy a decision should feel/ how much evidence is needed to judge a novel word-pair “meaningful”. The presence of clearly meaningful word-pairs may therefore have eased meaningful decisions to novel word-pairs, as it provides a neat comparison for their judgement. The presence of easy meaningfulness decisions also meant participants were under less pressure to find novel word-pairs meaningful (since they could respond meaningful to the clearly meaningful word-pairs), making their task of judging novel word-pairs easier/ more fluent. If this increase of ease/ fluency caused these participants to employ associative information processing strategies, then (as predicted by ECCo) they should have used heuristics and judged the target novel word-pairs as less meaningful than participants who saw novel word-pairs alone (because these word-pairs should not share distributional history). These participants should also have been faster in their judgements, as the associative process is effortless and quick.

The absence of easy (or fluent) meaningfulness decisions (in the novel50% and novel33% conditions) may have caused participants to have no obvious threshold for meaningfulness and to feel more pressure to find meaning in the target novel word-pairs. If this decreased ease/ fluency caused participants to employ analytical information processing strategies then they should have been more likely to judge target novel word-pairs as meaningful than participants who saw target novel word-pairs with clearly meaningful word-pairs. This is because the analytical system (according to ECCo) combines simulations of both key concepts, and is therefore capable of online metaphoric meaning construction. Because analytical processing is effortful, these participants should have been slower in their judgements.

Second, the role of decision making biases was investigated, because, by nature the different conditions of this experiment encouraged different meaningfulness thresholds to the novel word-pairs. While the same target number of meaningful responses was implied (verbally) for participants in the conventional, literal, and novel50% conditions, their implicit criteria may have differed. Participants in the conventional and literal conditions should have found most (if not all) of their 25 experimental-context-setting (conventional metaphoric and literal) word-pairs meaningful. Therefore, to meet the verbally implied target of *one-half*, these participants only needed to find 25 of the 75 target novel word-pairs meaningful. This would set an implicit meaningfulness criterion for target novel word-pairs at one-third (25/75 = 33%). In contrast, the participants in the novel50% condition (who only judged novel word-pairs and who heard the same verbally implied target as participants in the conventional and literal conditions), saw no clearly meaningful word-pairs, and were therefore just as likely to find any of these word-pairs meaningful as each other. Therefore, these participants might have an implicit meaningfulness criterion for target novel word-pairs of one-half (50%). Overall, although the implied targets remained the same (one-*half*), it was possible that implicit criteria of meaningfulness for target novel word-pairs differed between the conventional and literal conditions, and the novel50% condition.

To control for differences in implicit criteria, the novel33% condition was included in this study. This condition was exactly the same as the novel50% condition, except that participants in this condition were told that "in the past, participants have found *one-third* of these word-pairs to be meaningful". Therefore, the implicit criterion in the novel33% condition matched those in the conventional and literal conditions (although it was based on a different verbally implied target). To determine whether or not participants used verbally implied targets to set implicit criteria, responses in the novel50% and novel33% conditions were compared. If responses between these two conditions did not differ, then this would indicate that something other than verbally implied targets (potentially experimental context) was driving participant's implicit criteria. In this case, participant's implicit criteria would be unknown, and therefore whether or not experimental context affected bias would need to be further investigated.

A third and final enquiry addressed in experiment one involved global metaphoric process-priming. Participants' judgements of novel word-pairs were compared in the two conditions in which novel word-pairs were mixed with clearly meaningful word-pairs (the conventional and the literal conditions). In the conventional condition, the process of metaphoric interpretation in response to the conventional metaphors may have facilitated figurative thinking and therefore aided meaning construction for target novel word-pairs (which can only be interpreted metaphorically). If such process-priming occurred, then it would be found that the target novel word-pairs were judged as "meaningful" more often in the conventional condition than in the literal condition.

Method

Participants

Participants were 180 undergraduate students from Victoria University of Wellington. The mean age of participants was 20.43 years (*SD* = 4.02), and the age range was 18-61 years. In total there were 119 women and 61 men. Of these, 45 (27 women) were in the conventional condition, 47 (28 women) were in the literal condition, 44 (32 women) were in the novel50% condition, and 44 (32 women) were in the novel33% condition.

Materials

Adjectives (126) and nouns (124) that were balanced for frequency and length were taken from the Affective Norms for English Words (Bradley & Lang, 1999). All adjectives that were colour words and all proper nouns were excluded. The remaining 222 words were paired randomly using an experiment programmed in E-prime (Schneider, Eschman, & Zuccolotto, 2002), such that the first word from each pair was always an adjective, and the second word was always a noun. The novel word-pairs generated by this program were then used to make a screening questionnaire. This screening questionnaire was completed by 15 participants (8 women). The average age was 23.18 years (*SD* = 1.72). The screening questionnaire (Appendix B) required participants to indicate how meaningful they found the 111 novel word-pairs on a five-point scale, with 5 representing "very meaningful", and 0 representing "not meaningful at all". In general, participants found these novel word-pairs to be low in meaning (*M* = 2.04, *SD* = 1.12). However, there was some variability among word-pairs, with the most meaningful being rated on average as 3.87 out of 5, and the least meaningful being consistently rated as 0 out of 5. The eleven word-pairs with the lowest ratings were excluded. All remaining word-pairs and their average ratings are available in Appendix C. Of these word-pairs, 75 were randomly selected to be the target novel word-pairs which appeared in each condition. The remaining 25 were the additional novel word-pairs that were used in the novel33% and the novel50% conditions.

Finally, 25 conventional word-pairs, and 25 literal word-pairs were taken from those used by Faust and Mashal (2007; see Appendix D) for use in the conventional and literal conditions

Design

Participants undertook a five-to-ten minute experiment on computers. They saw 106 adjective-noun word-pairs (there were six practise and 100 experimental trials), and had to decide whether each pair was meaningful or not.

This experiment had four conditions, with condition as a between-subjects variable. Each condition had six practise trials that were comprised of novel word-pairs. Conditions were as follows:

1. Conventional: The 100 experimental word-pairs were made up of 75 target novel word-pairs and 25 conventional metaphoric word-pairs. Participants in this condition heard the following verbal instructions:

"In the past people have found around *half* of these word-pairs to be meaningful, but don't worry, it is a subjective decision so there are no right or wrong answers".

2. Literal: The 100 experimental word-pairs were made up of the same 75 target novel word-pairs as in the conventional condition, and 25 literal word-pairs. Participants in this condition heard the same verbal information as participants in the conventional condition.

3. Novel50%: The 100 experimental word-pairs were made up of the same 75 target novel word-pairs as in the above two conditions, and an additional 25 novel word-pairs. This condition controlled for verbal instruction, because participants in this condition heard the same verbally implied targets as did those in the conventional and literal conditions. However, as explained earlier, this condition might give participants a higher implicit criterion for meaningful responses to novel word-pairs (50%) than participants in conventional and literal conditions (33%).

4. Novel33%: The 100 experimental word-pairs were made up of the same word-pairs as in the novel33% condition. Participants in this condition heard the following verbal instructions: "In the past people have found around *one-third* of these word-pairs to be meaningful, but don't worry, it is a subjective decision so there are no right or wrong answers". As previously described, these instructions mean that this condition controlled for criterion, with participants in this condition presumably having the same implicit criterion for meaningful responses to novel word-pairs as participants in the conventional and literal conditions.

Procedure

Participants either remained behind after their (second year) cognitive psychology labs had finished, or met the experimenter in the experiment room. They were first given a consent form to fill out (Appendix E) and then asked if they had any questions. If not, they were given the appropriate verbal instructions and then began the experiment. Participants initially saw an instruction screen indicating which button to press if they thought a word-pair was meaningful and which button to press if they thought a word-pair was non-meaningful. The 'meaningful' and 'not-meaningful' keys were the '1' and '2' keys (on the number pad), and these were counterbalanced across participants. Then (after pressing the space bar) participants saw the first practise trial. Practise trials were set out in the same way as experimental trials. During each trial, participants saw the following: the first word of the pair for 200ms, a blank screen for 800ms, and then the second word for 200ms. Participants then saw another blank screen and had to decide if the word-pair was meaningful or not. Presentation of this screen was ended by a button press (either the '1' or the '2' key). After six practise trials, participants were reminded of which keys were their meaningful and non-meaningful keys, and then they proceeded for 100 experimental trials. Their response times to the second word of each pair, and their responses were recorded. Afterwards, participants were debriefed (Appendix F), and given either a course credit or a chocolate fish.

Results

All statistical analysis was performed in Excel and SPSS (Statistical Package for Social Sciences) 17.0 for windows, and used an alpha level of .05. In all analyses, any trials in which responses were made in less than 200ms were excluded because they were considered too fast to be in response to stimuli. Any response that took longer than 5000ms probably did so due to the participant being distracted from the task and so were also excluded. In total, 92 out of 18851 trials were excluded.

Subject analysis

To determine if participants judged the target novel word-pairs differently depending on the condition in which they appeared, the proportion of “meaningful” responses to target novel word-pairs for each participant was calculated (the number of times each participant chose a meaningful response to a target novel word-pair/ their total number of responses to target novel word-pairs). A one-way ANOVA with an independent factor of condition (conventional, literal, novel33% or novel50%), and a dependent variable of proportion of “meaningful” responses for each participant was calculated. The resulting comparison was significant, $F(3, 176) = 4.86$, $p = .003$, $\eta^2_p = .076$. Post-hoc tukey tests revealed that participants were more likely to call a target word-pair “meaningful” in the two conditions in which only novel word-pairs were presented (the novel50% and novel33% conditions) than in the other two conditions (conventional and literal). However, participants did not significantly differ in their proportion of “meaningful” responses between the two conditions in which only novel word-pairs were presented or between the two conditions in which novel word-pairs were presented mixed with other types of word-pairs. See Table 1 for the mean proportion of novel word-pairs judged as “meaningful”.

Table 1

Mean proportion of times that subjects judged target novel word-pairs as "meaningful".

Condition	Mean	Standard deviation
Conventional	.18[a]	.13
Literal	.19[a]	.14
Novel50%	.27[b]	.17
Novel33%	.27[b]	.18

Letters indicate that means do not significantly differ; all other comparisons differ significantly at the $p < .05$ level.

To determine if response times to the “meaningful” versus “not-meaningful” responses differed between conditions, a repeated measures ANOVA was conducted on response times, with condition (conventional, literal, novel50% or novel33%) as a between-subject factor, and response type (“meaningful”/“not-meaningful”) as a within-subjects factor. No significant main effects or interactions were found.

Item Analysis

The proportion of “meaningful” responses to each of the target word-pairs in each condition was calculated. Using these proportions, a repeated measures ANOVA with condition (conventional, literal, novel33% or novel50%) as a within-subjects factor was calculated. This was found to be significant, $F(3, 219) = 55.24$, $p < .001$, $\eta^2_p = .431$.

Follow-up paired sample t-tests with a Bonferroni adjusted alpha level of .01 showed results parallel to those in the subject analysis. Target novel word-pairs were significantly more likely to be judged “meaningful” in the conventional condition than the novel50% condition, $t(74) = -8.64$, $p < .001$, $d = 1.007$[1], or in the novel33% condition, $t(74) = -10.34$, $p < .001$, $d = 1.198$. Target novel word-pair were also significantly more likely to be judged “meaningful” in the literal than in the novel50%, $t(74) = 6.93$, $p < .001$, $d = .806$, or the novel33% condition, $t(74) = -8.90$, $p < .001$, $d = -1.023$. The proportion of “meaningful” responses to target novel word-pairs did not significantly differ between conventional and literal conditions. The proportion of “meaningful” responses to target novel word-pairs also did not significantly differ between novel33% and novel50% conditions. See Table 2 for means and standard deviations.

[1] Cohen’s d was calculated for all paired sample t-tests in this study by dividing the population mean change by the standard deviation of change scores (Gibbons, Hedeker & Davis, 1993). Cohen’s d was calculated for all independent sample t-tests in this study by dividing the difference between the means of the two populations by the pooled standard deviation (Hedges, 1981, 1982).

Table 2

Mean proportion of times that target novel word-pairs were judged as "meaningful."

Condition	Means	Standard deviations
Conventional	$.18_a$	.14
Literal	$.19_a$	.13
Novel50%	$.27_b$	.16
Novel33%	$.28_b$	.14

Letters indicate that means do not significantly differ; all other comparisons differ significantly at the $p < .05$ level.

In order to determine if the response times to “meaningful” versus “not-meaningful” responses differed across conditions, a repeated measures ANOVA for median response times was calculated with the within-subjects factors of condition (conventional, literal, novel50% or novel33%) and response type (“meaningful”/“not-meaningful”). A significant main effect of response was found, $F(1, 74) = 76.12$, $p<.001$, $\eta^2_p = .507$. “Not-meaningful” responses took less time than “meaningful” responses. A main effect of condition was also found, $F(3, 222) = .59$, $p= <.001$, $\eta^2_p = .074$. Follow-up paired sample t-tests with a Bonferroni adjusted alpha level of .01 showed that, regardless of response type, target novel word-pairs in the novel50% condition were judged more slowly than target novel word-pairs in the conventional, $t(74) = 3.20$, $p < .001$, $d = .377$, literal, $t(74) = 3.99$, $p < .001$, $d = -.425$, or novel33% conditions, $t(74) = 3.65$, $p < .001$, $d = -.462$. No other significant differences between conditions were found. Condition and response did not interact. See Table 3 for averaged median reactions times and standard deviations.

Table 3

Median response times (ms) for "meaningful" and "not-meaningful" responses to the same target novel word-pairs (for participants in different conditions).

Response	Condition	Mean	Standard deviation
"Meaningful"	Conventional	1016	453
	Literal	1034	331
	Novel50%	1144	248
	Novel33%	1023	254
"Not-meaningful"	Conventional	846	65
	Literal	856	54
	Novel50%	926	88
	Novel33%	878	80

Correlation Analysis

To determine whether or not the same word-pairs were judged "meaningful" in each condition, correlations were calculated between the proportions of meaningful responses to target novel word-pairs for each condition. All correlations were found to be strong, positive and significant, indicating that if a word-pair was likely to be judged "meaningful" in one condition, then it was likely to be judged "meaningful" in other conditions. This result implies that similar word-pairs were judged as meaningful across conditions, suggesting that participants were not responding randomly. See Table 4 for the correlation coefficients.

Table 4

Correlations between proportions of "meaningful" responses for the target novel word-pairs in different conditions.

Condition	Conventional	Literal	Novel50%	Novel33%
Conventional				
Literal	.86**			
Novel50%	.83**	.80**		
Novel33%	.84**	.83**	.85**	

** Correlation is significant at the .001 level (2-tailed).

Discussion

Results showed that participants in the two conditions in which novel word-pairs were presented alone (the novel50% and novel33% conditions) were not significantly different in their likelihood of judging target novel word-pairs "meaningful". Similarly, participants in the two conditions in which novel word-pairs were presented alongside other types of word-pairs (the conventional and literal conditions) were not significantly different in their likelihood of judging target novel word-pairs "meaningful." However, participants in the two conditions in which novel word-pairs were presented alone were more likely to judge target novel word-pairs "meaningful" than participants in the two conditions in which novel word-pairs were presented alongside other types of word-pairs.

In the item analysis, "meaningful" responses took longer than "not-meaningful" responses, indicating that successfully finding a novel word-pair meaningful takes longer than rejecting it as meaningless. Participants in the novel50% condition took longer to respond than participants in all other conditions. These results fit with the idea that these participants were using their implied target (50%), to set their implicit criterion for meaningfulness, trying to find meaning in a target 50% of target novel word-pairs, and therefore taking longer to respond than other participants (who should have had a implied criterion of 33%).

However, because responding did not differ for participants in the novel50% and novel33% conditions (who had different implied targets), the fact that participants took longer to respond in the novel50% condition clearly did not affect their responses. Therefore, even if these participants were aiming for their implied targets to set meaningfulness criteria, they did not resort to responding randomly to meet this.

Further support that participants were not randomly responding to meet implied targets was provided by correlations between the proportions of times that target novel word-pairs were judged meaningful across conditions. These correlations were positive and highly significant which indicates that similar target novel word-pairs were judged as meaningful across conditions. Therefore, the possibility that participants were responding randomly to meet a meaningfulness threshold can be ruled out.

Metaphor

The current results showed that participants who saw target novel word-pairs with additional novel word-pairs were more likely to judge them "meaningful" than participants who saw them mixed with other types of word-pairs. This finding of differential responding to the same novel (loosely figurative) word-pairs is consistent with Glucksberg and Haught's suggestion (2006) that novel metaphors can be processed in more than one way. Glucksberg and Haught suggest that novel metaphors can be processed via comparative or categorical processing. The current results do not specify whether or not observed processes are comparative and categorical processing. Nevertheless, the types of responses found are consistent with comparative and categorical processing. Participants using comparative processing should often find novel word-pairs meaningful because comparative alignment should often result in finding a shared metaphor ground. Consistent with this, participants thought to use analytical processing (which may share properties with comparative processing) found more novel word-pairs meaningful. In contrast, the vehicle of novel word-pairs should (by nature) not be linked to a super-ordinate category meaning, and therefore participants using categorical processing should rarely find novel word-pairs meaningful. Consistent with this, participants thought to use associative processing (which may share properties with categorical processing) found less novel word-pairs meaningful.

However, the current results cannot be completely explained by Glucksberg and Haught's aptness theory, because this would predict that that "semantic and referential properties" would determine the way in which novel word-pairs are processed. Therefore, proponents of aptness theory would predict that the same novel word-pairs would be judged in the same way across all conditions. However, participants in some conditions judged more novel word-pairs as "meaningful" than participants in other conditions. Therefore, novel word-pairs were not all judged in the same way across conditions, and so aptness cannot explain current findings. Overall, the current results show that the processing of novel (loosely figurative) word-pairs can differ depending on experimental context, even when factors such as salience and aptness are kept constant.

Information Processing

The differences found in the proportion of "meaningful" responses to novel word-pairs between conditions are consistent with predictions derived from the fluency/task-demand and information processing strategy selection literature. Alter et al. (2007) established that participants were more likely to employ the analytical information processing system when their task was less fluent, and the associative information processing system when their task was more fluent. Similarly, originators of the ECCo model (Lynott & Connell, 2010) predicted that during a language task, people will be more likely to use a heuristic-based linguistic system if task demands are low and a simulation system if task demands are high. In experiment one, participants who saw novel word-pairs alone did not encounter as many easily meaningful decisions as participants who saw novel word-pairs with clearly meaningful word-pairs. Therefore, participants who saw novel word-pairs alone had harder/ less fluent tasks overall. It is possible that based on this increase in task difficulty/disfluency, participants in conditions which included only novel word-pairs applied a more analytical information-processing system to the novel word-pairs than did participants who saw novel word-pairs with clearly meaningful word-pairs. The increased use of this analytical system could explain why participants who only saw novel word-pairs judged them as more meaningful than did participants who saw novel word-pairs with other types of word-pairs, as these novel word-pairs only made sense in a loose figurative fashion (if at all), and thus could only be found meaningful when assessed in depth online.

Biases in Decision-Making

Responses did not significantly differ between the two conditions in which only novel word-pairs were presented (the novel50% and novel33% conditions). Such a finding implies that participants were not using the verbally implied targets when judging novel word-pairs; therefore their meaningfulness criteria remain unknown. It is possible that experimental context may have biased their results. The correlations conducted in the item analysis show that there were strong relationships between how often target novel word-pairs were judged “meaningful” between conditions. Since these correlations were positive, this shows that participants tended to call the same word-pairs “meaningful” across conditions, which indicates that they were not biased in the respect that they simply responded randomly to meet a fixed criterion of “meaningful” responses. However, this finding does not rule out the possibility that participants moved their criteria for meaningfulness depending on experimental context. Participants who judged novel word-pairs “meaningful” more often (those in the novel50% and novel33% conditions) may have done so because their experimental context led them to lower their meaningfulness criteria, and therefore require less evidence for meaning than participants in other conditions. Therefore, it is still a possibility that different experimental contexts caused different biases in meaning construction.

Process Priming

No significant difference was found between responses in the conventional and literal conditions. As previously explained, this indicates that process priming did not occur in the current experiment. Past experimenters who have tested process priming found only questionable evidence for its existence (Basko & Briihl, 1997; Inhoff et al., 1984). Past studies considered the immediate effects of a metaphor on a subsequent sentence, whereas the current study investigated a more global version of process priming (whether or not a person could get into a “metaphoric way of thinking” across the course of an experiment). Perhaps then, process priming is only “effective within a relatively short time interval” as Inhoff et al. initially proposed (Inhoff et al., 1984, p. 564). Alternatively, because conventional and novel metaphors have been suggested to be processed differently (e.g. Gentner & Wolff, 1997), perhaps the conventional metaphors used in the current experiment did not provide relevant primes for the novel (loosely figurative) targets.

Experiment two

The aims of experiment two were first to replicate the results of experiment one, and second to determine whether these results arose due to bias, the use of different information processing systems, or both. Experiment one showed that participants who saw novel word-pairs alone found these more meaningful than participants who saw the same novel word-pairs with clearly meaningful word-pairs. Replication of these results is necessary to verify the robustness of this effect. In experiment two, a different set of novel word-pairs were used in a different experimental setting to experiment one. Therefore, if results of experiment one are replicated then this would confirm the strength of the findings of experiment one. Experimental context affects meaning construction and metaphor processing (see experiment one); how this occurs is unresolved. Therefore, further investigation is significant to the understanding of meaning construction. In order to further the understanding of the results observed in experiment one, the technique of event-related potentials was recruited.

Event-Related Potentials (ERPs)

ERPs are differences in ongoing electric brain activity that occur in response to specific stimuli. Neurologically speaking, they are the summed post-synaptic potentials that arise when multiple neurons respond in synchrony to the same stimulus (for a review, see Rugg & Coles, 1995). ERPs are recorded via electrodes placed on the scalp and can be measured to determine differences in response, response effort, and response preparation (Lai, Curran, & Menn, 2009). Different ERPs (reflected in different waveforms or “components”) index qualitative differences in ongoing neurophysiological processes (Coulson & Van Petten, 2002). Differences in the magnitude or latencies of components index quantitative differences in ongoing neurophysiological processes (Coulson & Van Petten, 2002). ERP is a particularly useful technique when response time differences are not sensitive enough to differentiate underlying processes. In experiment one of the current study, differences in response and response time were predicted for participants in different conditions because they were expected to use strategies which differed in cognitive effort. The item analysis revealed differences in response times, however these differences did not map clearly onto the differences in response found. Equivalent response times do not always reflect the amount of effort spent (Coulson & Van Petten, 2002). For example, it would take just as long to lift a 2 kg weight as a 4 kg weight, although the amount of effort required would differ. Therefore, experiment two will employ the more sensitive measure of ERP in the attempt to reveal differences in cognitive effort which explain differential responding between conditions. Of use to this investigation are two ERP components, the N400 and the Lateralized Readiness Potential (LRP). In experiment two, the N400 component will be used to determine whether the results of experiment one arose due to differences in interpretation based on information processing strategy. The LRP component will be used to determine whether the results of experiment one arose due to differences in meaningfulness criteria.

The N400

The N400 is a particularly well studied language-related ERP component (Arzouan, Goldstein, & Faust, 2007; Corenjo et al., 2009; Coulson & Van Petten, 2002; Iakimova, Passerieux, Laurent, & Hardy-Bayle, 2005; Kazmerski, Blasko, & Dessalegn, 2003; Kutas & Federmeier, 2000, 2011; Kutas & Hillyard, 1980; Lai et al., 2009; Pynte et al., 1996; Sakamoto, Matsuishi, Arao, & Oda, 2003; Tartter, Gomes, Dubrovsky, Molholm, & Stewart, 2002). It is visible as a negative deflection in ongoing electrical activity from the brain which peaks around 400 milliseconds after the presentation of a stimulus (Kutas & Hillyard, 1980). The amplitude of the N400 has consistently been shown to vary with the degree of semantic conflict that a particular stimulus presents, becoming more negative as semantic conflict increases (for a review, see Kutas, Van Petten, & Kluender, 2006). More specifically, the N400 has been said to be an index of the difficulty or ease involved with retrieving or constructing the meaning of a stimulus, and/or involved with the semantic integration of the stimulus meaning within a context (Kutus & Federmeier, 2000; Kutas & Van Petten, 1994; Lai et al., 2009). The N400 was first observed in response to semantically incongruent sentences (Kutas & Hillyard, 1980), but has since been observed in response to various other types of stimuli including non-words (Bentin, McCarthy & Wood, 1985), music (Besson & Macar, 1987), pictures (Barrett & Rugg, 1990), and most importantly to this study, metaphors (Arzouan et al., 2007; Coulson & Van Petten, 2002; Iakimova et al., 2005; Lai et al., 2009; Pynte et al., 1996; Tartter et al., 2002).

Usually, N400s are found to be more negative to metaphoric than to literal phrases, indicating that the semantic processing of metaphors is more difficult than the semantic processing of literal phrases (Arzouan et al., 2007; Coulson & Van Petten, 2002; Iakimova et al., 2005). Novel metaphors elicit more negative N400s than conventional metaphors indicating that the semantic processing of novel metaphors is more difficult than the semantic processing of conventional metaphors. (Arzouan et al., 2007). Lai et al. (2009) found that when the N400 time-window is broken down further, although both conventional and novel metaphors elicit greater negativity than literal phrases initially, the activity in response to conventional metaphors converges with that of literal phrases in the second half of this window, while that to novel word-pairs remains more negative. These findings indicate that while novel and conventional metaphors initially require similar semantic effort, this is sustained for novel, but not conventional, metaphors. Astoundingly, past ERP studies of metaphor have not considered the effects of experimental context on the processing of novel metaphors. In the current study, it was of interest to see if the same novel (loosely figurative) word-pairs would elicit N400s of different magnitudes depending on the context in which they were presented.

In the current study, experiment one was insufficient to resolve how the experimental contextual presence of random literal or conventional metaphoric word-pairs caused participants to judge target novel word-pairs as less meaningful than did the contextual presence of additional novel word-pairs. Observation of the N400 in response to these novel word-pairs will help in elucidating whether these results arose because novel word-pairs were processed with different degrees of semantic effort in different conditions.

Dual processing theories predict that participants in the conditions in which only novel word-pairs were presented had a more difficult/disfluent task, and therefore should have relied more on an analytical processing strategy. In contrast, participants in the conditions where novel word-pairs were presented amongst clearly meaningful word-pairs had a relatively easier/ fluent task, and therefore should have relied more on an associative processing strategy. Because the analytical processing strategy involves simulation and meshing of concepts, it requires increased semantic effort relative to the associative (linguistic) processing strategy, which is heuristic-based and shallow. If dual processing predictions are true, then participants in conditions in which only novel word-pairs were presented should have more negative N400s in response to target novel word-pairs than participants in conditions in which these were presented mixed with clearly meaning word-pairs.

Morphological differences in ERP

Contemporary theories of metaphor processing suggest that there are two processes with which metaphor can be assessed: comparative and categorical processes. The selection of these processes has been attributed to several factors including metaphor salience (Gentner & Wolff, 1997), interpretive diversity (Utsumi, 2007, 2011) and metaphor aptness (Glucksberg & Haught, 2006). Qualitative morphological differences in ERP represent differential processing (Coulson & Van Petten, 2002). If experimental context is another factor which influences which process a metaphor is assessed with, then qualitative rather than quantitative morphological differences in brain activity are expected between the two conditions in response to the same novel word-pairs. For the same reasons, qualitative differences in ERP between conditions would also provide further support for dual processing predictions.

The Lateralized Readiness Potential (LRP)

The LRP will be used in experiment two to investigate bias. The LRP is a measure of response preparedness (e.g. see Gratton, Coles, Sirevaag, Eriksen & Donchin, 1988). It is indexed by an increase in negative scalp potentials over the motor cortex on the side contralateral to the hand which will be used to respond (Kutas & Donchin, 1974, 1977, 1980). The more negative the LRP, the more prepared someone is to respond with their contralateral hand. Therefore, if one response is assigned to one hand, and another response is assigned to the other hand, then one can compare cortical LRPs to determine whether participants are more prepared to respond with one hand (that is with one response) than the other. The LRP begins once response selection has been completed, at the beginning of motor preparation (Masaki, Wild-Wall, Sangals, & Sommer, 2004). Therefore, if one locks the ERP signal to response, then ERP waveform just prior to response should be more negative over the motor cortex contralateral to response hand (response-locked LRP).

A recent study by Steinhauser, Hubner and Druey (2009) showed that if a manipulation causes a person to be biased towards one response relative to the other, then the LRP will begin prior to the onset of the stimulus. Consequently, if the ERP signal is locked to stimulus onset (stimulus-locked LRP), then the bias will be revealed in ERP waveform just prior to stimulus onset. In each trial, Steinhauser et al. (2009) presented their participants with a cue followed by a digit. Participants had two tasks which were randomized throughout trials: to either judge whether the digit was odd or even; or whether it was greater than or less than five. Responses were also randomised. Therefore, sometimes participants would complete the same task for two trials in a row (task repetition) and sometimes they would switch tasks (task-switch). Also sometimes participants would press the same response twice in a row (response repetition) and sometimes they would need to switch responses (response-switch). Steinhauser et al. focused on the LRP in the interval between each cue and the presentation of the following stimulus (the cue-stimulus interval: CSI). This interval was focused on so that it could be determined whether response preparation manifested before the stimulus to which participants needed to respond was even presented. The results showed that during the CSI, the LRP drifted from the side corresponding to the previous response to the opposite side, reflecting a response-switch bias. This LRP was not affected by whether the task itself switched or not.

Steinhauser et al. (2009) demonstrated that the LRP can show that a participant is ready to respond in a certain *way before the stimulus is even presented.* Therefore, Steinhauser et al. showed that the LRP can be adequately used to determine whether or not a bias reflected in motor preparation is occurring. In the current study, response bias was a possible explanation for results, and therefore needed to be investigated. Therefore, experiment two was designed so that any potential response bias would be reflected in motor preparation, and therefore could be seen in participants' LRPs.

For each participant, the right hand was assigned to press the “meaningful” key, and the left hand to press the “not-meaningful” key. Therefore, if a participant is predisposed to say “meaningful” or “not-meaningful” due to a decision-making bias, then they should be predisposed to use their left or their right hand to respond before they even see the stimulus. Such a bias in responding would be reflected in corresponding LRPs during the CSI (time between the first and second word of each pair). Importantly, a long CSI was available between each word in the pair, so that the LRP would have time to build. In terms of the decision-making models described earlier, such a stimulus-locked LRP prior to the onset of the second word in a word-pair would reflect the criterion shift away from one distribution in SDT leading to participants being more likely (or “biased”) to make one response over the other. Response-locked LRPs were also inspected in experiment two to assure that the analysis of the LRP was correctly conducted (LRPs over contralateral hemispheres should occur before response regardless of whether participants are biased or not, and so this provides a dependable analysis check).

The LRP electrophysiological measure of bias was used rather than a traditional measure of bias because of the particular stimuli involved and the question being addressed. If traditional measures of bias had been used, then it would be necessary to determine whether each word-pair was meaningful or not so that hits and false alarms could be assigned. However, since meaningfulness is subjective, assigning a meaningful response as a hit or false alarm is impossible. Usually, a ratings experiment would provide answers as to whether each word-pair is meaningful or not. However, rating experiments themselves inevitably contain certain experimental contexts (either the novel word-pairs are rated alone or with the other types of word-pairs). Since experimental context is the factor predicted to influence responding in this study, a rating experiment would not provide a fair measure of whether each word-pair was meaningful or not-meaningful. Using the LRP allows for a more objective measure of bias in the current experiment, because it highlights the contrast between participants’ response preparation in different conditions without making necessary the comparison with ratings which may themselves be contaminated by bias.

In summary, experiment two was conducted to determine whether the differing effects of experimental context on meaning construction observed in experiment one occurred due to a shift in interpretation based on information-processing strategy (as reflected in the N400) and/or due to a shift in criterion (as reflected in the LRP). If the predicted differences in N400 magnitude are found, this would indicate that participants judge novel word-pairs differently depending on the information processing strategy that experimental context led them to employ. If the predicted differences in the LRP are found, then this would indicate that different experimental contexts lead participants to set different meaningfulness criteria. If both differences in the N400 and the LRP are found, then this would indicate that different experimental contexts lead participants to employ different information processing strategies, and to adopt different meaningfulness thresholds.

Because ERP is a relatively expensive technique (both financially and time-wise), only two conditions were included in experiment two. As responding in the conventional and literal conditions was not found to significantly differ in experiment one, only the conventional condition will be used in experiment two. Similarly, as responding in the novel50% and novel33% conditions was not found to significantly differ in experiment one, only the novel33% condition will be used in experiment two.

Method

Participants

Participants were 46 (24 women) right-handed undergraduate students from Victoria University of Wellington. Of these, 23 (12 women) were in the conventional condition, and 23 (12 women) were in the criterion control condition. However, due to technological problems with one of the EEG recording caps, 17 participants were excluded from analyses, leaving 16 participants (9 female) in the conventional condition, and 13 participants (5 females) in the novel33% condition. Participants all spoke English as their first language and had normal or corrected to normal vision. Participants were asked to wear glasses rather than contact lenses if their vision needed correcting, as contact lenses increase blink artifacts in EEG recording (see Luck, 2005). Participants had no history of neurological disorder.

Materials

For the preliminary study 606 adjectives and 605 nouns that were balanced for frequency and length were taken from the Affective Norms for English Words (Bradley & Lang, 1999). All adjectives that were colour words and all proper nouns were excluded. The remaining 1146 words were paired randomly using an experiment programmed in E-prime (Schneider et al., 2002), such that the first word from each pair was always an adjective, and the second word was always a noun. No words were repeated. The first 240 novel word-pairs generated by this program were then used as novel word-pairs in the experiment. Of these, 180 were randomly designated to be target novel word-pairs, and the remaining 60 were designated to be the additional novel word-pairs seen in the novel33% condition. These word-pairs can be seen in Appendix G.

Finding an extensive list of metaphoric word-pairs previously rated as conventional proved difficult. Therefore, various linguistic databases were scoured, and 221 metaphoric word-pairs (that the experimenter considered conventional) were found. These were assessed in a screening experiment to verify their conventionality. In the screening experiment, 8 participants (4 women) simply read the conventional metaphoric word-pairs and then answered three questions: “Does this pair make sense?” “Is it metaphoric?” “Is it a familiar word-pair?” The proportion of yes response across answers was calculated for each metaphoric word-pair, and the 60 word-pairs with the highest proportions were selected as conventional metaphors for the conventional condition. These word-pairs and the proportion of times they were judged as familiar, metaphoric and meaningful can be seen in Appendix H. The proportion of times these word-pairs were rated as making sense ranged from 0.86 to 1 (M = 0.96, SD = 0.07). The proportion of times these word-pairs were rated as metaphoric ranged from 0.57 to 1 (M = 0.78, SD = 0.23). The proportion of times these word-pairs were rated as familiar ranged from 0.43 to 1, (M = 0.92, SD = 0.16).

Design

Participants were randomly assigned to one of two conditions. Conditions were as follows:

1. **Conventional:** As in experiment one; the experimental word-pairs were made up of 75% (180) target novel word-pairs and 25% (60) conventional metaphoric word-pairs. Participants in this condition heard the following verbal information: “In the past people have found around *half* of these word-pairs to be meaningful, but don’t worry, it is a subjective decision so there are no right or wrong answers”.

2. **Novel33%:** The 240 experimental word-pairs were made up of 100% (240) novel metaphoric word-pairs (75% of which were those presented in the conventional condition). As in experiment one, this condition controlled for implicit criterion because participants in this condition heard the following verbal information: “In the past people have found around *one-third* of these word-pairs to be meaningful, but don’t worry, it is a subjective decision so there are no right or wrong answers”.

Procedure

In each trial a fixation cross was first presented for 500ms. The purpose of fixation was to focus participants attention on the centre of the screen so that they would not make an eye movement when the word-pair appeared (which would cause eye-movement artifacts in the ongoing electrical activity). If the first word of a word-pair appeared directly after the fixation then it might provide a backwards mask for fixation. Masking fixation might mean that it would not hold participant's attention, thereby increasing eye-movement artifacts. Therefore, a blank screen followed fixation for 100ms with the purpose of preventing backwards masking. Then the first word (from each word-pair) was presented for 200ms. This was followed by a blank screen for 800ms (giving the LRP time to build), and then the second word for 200ms. Finally a blank response screen came up, and this was terminated when participants made a response. Between each trial, an inter-trial interval which lasted anywhere between 800 and 1200ms was presented to prevent synchronisation of EEG oscillations to trials.

The experimental procedure was as follows: Participants completed a consent form (Appendix I) and then were fitted with the electrode Quik-Cap (Compumedics NeuroMedical Supplies) which was connected to a BrainAmp amplifier (Brain Products GmbH). They then undertook a twenty minute long experiment presented by E-Prime (Schneider et al., 2002). In the experiment, participants initially saw an instruction screen which informed them that their task was to decide whether each word-pair was meaningful or not. Participants were told to press the 'm' key on the keyboard if the pair was meaningful and the 'z' key if the pair was not-meaningful. Participants pressed a keyboard button to move on from the instructions screen. They then saw six practise trials. After the practise trials, participants saw a screen which reminded them which keys to press and told them to ask the experimenter any questions that they had. This screen was terminated by a button press, and then participants saw the 240 experimental trials. Experimental and practise trials followed the same formula (described above). Participants' response times to the second word of each pair, their responses, and their ongoing electrical brain activity was recorded Afterwards, participants were fully debriefed (Appendix J).

Hands were not counter-balanced for "meaningful" and "not-meaningful" response. Usually in experimentation, response hands are counterbalanced because each hand is controlled by a different hemisphere of the brain. Each hemisphere might use different processes for responding and provide unique noise. Counterbalancing eliminates this noise so that hemispheric effects do not mask experimental effects (Osman & Moore, 1993). However in the current experiment, between-condition contrasts were of interest. If participants in all conditions have the same hands assigned to the same responses, then when comparing results for the two conditions, participants in each should be equally affected by hemispheric noise, and therefore any differences found could not be attributed to this. Keeping response hand constant for all participants also decreased noise in the ERP, and therefore maximised the power of the experiment to detect between-condition effects.

ERP Recording

Participants were seated in a sound attenuated, electrically shielded room for the duration of recording. EEG activity was recorded continuously from 26 Ag/AgCl electrodes embedded in a lycra Quik-Cap (Compumedics NeuroMedical Supplies) referenced online to linked mastoids (M1+M2). The system used for recording was BrainVision Recorder (Brain Products GmbH). Vertical eye movements were measured by electrodes placed above and below the left eye, and horizontal eye movements were recorded by two electrodes located lateral to the left and right eyes. The impedances for scalp electrodes were kept below 5kΩ. Impedances for eye electrodes were kept below 10kΩ. Artifact rejection was used to exclude trials with excessive eye artifacts (specified as occurring when a voltage step of more than 50Hz occurred, if activity less than .05µV occurred in 100ms intervals, or if more than a difference of 300µV occurred within a given segment). In total 18% of trials were excluded for this reason (15% of trials in the novel33% condition and 20% of trials in the conventional condition). The signal was sampled at 500Hz. Markers were set for the onset of the second word for each word-pair, and for both “meaningful” and “not-meaningful” responses.

N400. Consistent with the study by Pynte et al. (1996) and with the recommendations of Luck (2005), the signal was filtered offline with a low pass frequency of 40Hz and a high pass filter of 0.01Hz. A notch filter of 50Hz was used to attenuate extraneous noise from electrical equipment. Epochs stretching from 200ms before the onset of the second word in each word-pair to 800ms after the onset were created. The 200ms before the onset of the second word of each word-pair was used as a baseline. Consistent with the analysis of Arzouan et al. (2007), the mean area under the curve between 350 and 450ms after the second word was investigated.

LRP. Consistent with the studies by Masaki et al. (2004), the signal was filtered offline with a low pass frequency of 30Hz. Following the recommendations of Luck (2005), a high pass filter of 0.10Hz and a notch filter of 50Hz were applied. Epochs stretching from 200ms before the onset of the fixation cross which began a trial, until 200ms after the onset of the second word of each word-pair (totaling 1800ms) were created for the stimulus-locked LRP. Epochs stretching from approximately 200ms before the onset of the fixation cross which began a trial, until 500ms after response (totaling 3000ms) were created for the response-locked LRP. The 200ms before the onset of the fixation cross was used as a baseline, as the time period of interest (for the stimulus-locked LRP) occurred just before the onset of the second word in each pair, and therefore this interval could not have been used as a baseline as it was used in the N400 analysis. The mean areas in the 300ms prior to the onset of the second word (for the stimulus-locked LRP), and the 300ms prior to response (for the response-locked LRP) were investigated.

Results

All statistical analysis was performed in Excel and SPSS (Statistical Package for Social Sciences) 17.0 for windows, and used an alpha level of .05. EEG filtering and data averaging was completed in Brain Vision Analyzer (Brain Products GmbH). In all analyses, any trials in which responses were made in less than 200ms were excluded because they were considered too fast to be in response to the stimuli. Any response that took longer than 5000ms probably did so due to the participant being distracted from the task and so were also excluded. In total 57 out of 8160 trials were excluded for these reasons.

Behavioural Data

Subject analysis. To determine if participants judged the target novel word-pairs differently depending on the condition in which they appeared, the proportion of "meaningful" responses for each participant were calculated (the number of times they made a meaningful response to a target novel word-pair/their total number of responses to target novel word-pairs). Next an independent samples t-test with an independent factor of condition (conventional or novel33%), and a dependent variable of the proportion of "meaningful" responses for each participant was calculated. As in experiment one, a significant difference between the proportion of "meaningful" responses in each condition was found, $t(42) = -3.20$, $p = .003$ $d = -1.023$. Despite the use of different novel word-pairs from experiment one, participants again called novel word-pairs "meaningful" significantly more often in the novel33% condition (*M* = .36, *SD* = .15) than in the conventional condition (*M* = .23, *SD* = .11).

Because 17 participants were excluded from ERP analysis due to technological faults, and because it was relevant to compare the ERP data with the behavioural data from the same people, the above analysis was re-run including only data from participants who were kept in ERP analyses. An independent samples t-test with an independent factor of condition (conventional or novel33%), and a dependent variable of the proportion of "meaningful" responses for these participants revealed the same results as above, $t(25) = -4.10$, $p < .001$, $d = -1.758$. These participants also called novel word-pairs "meaningful" significantly more often in the novel33% condition (*M* = .43, *SD* = .16) than in the conventional condition (*M* = .23, *SD* = .11).

In order to determine if the response times to the different response types ("meaningful"/"not-meaningful") differed across conditions, a repeated measures ANOVA was calculated with between- subjects factors of condition (conventional or novel33%) and a within-subjects factor of response types ("meaningful"/"not-meaningful"). A main effect of response type was found, and this was qualified by a Response type by Condition interaction, $F(1, 42) = 8.20$, $p = .007$, $\eta^2_p = .163$. Follow-up independent sample t-tests revealed that participants in the novel33% condition were slower to say "not-meaningful" (*M* = 995ms, *SD* = 332ms) than participants in the conventional condition, (*M* = 815ms, *SD* = 192ms), $t(42) = 2.17$, $p = .036$, $d = 0.711$. No effect was found for "meaningful" responses.

When this analysis was re-run to include only participants whose ERP data was kept, the same results were found, $F(1, 25) = 17.47$, $p < .001$, $\eta^2_p = .411$. Participants in the novel33% condition were significantly slower to say "not-meaningful" to novel word-pairs (*M* = 1074ms, *SD* = 306ms) than participants in the conventional condition (*M* = 813ms, *SD* = 204ms), $t(25) = -2.67$, $p = .013$, $d = 1.151$.

Item analysis. The proportion of "meaningful" responses to each of the target word-pairs was calculated. Using these proportions as a dependent variable, a paired samples t-test with condition (conventional or novel33%) as a within-subjects factor, was calculated. As in the subject and item analyses of experiment one, and the subject analysis of experiment two, this was found to be significant, $t(179) = 14.88$, $p < .001$, $d = 1.114$. Target word-pairs were more likely to be called "meaningful" in the novel33% condition (*M* = .36, *SD* = .21) than in the conventional condition (*M* = .23, *SD* = .19).

When this analysis was re-run to only include participants whose ERP recording was kept, the same results were found in a paired samples t-test , $t(179) = 15.16$, $p < .001$, $d = .790$. Target word-pairs were more likely to be called "meaningful" in the novel33% condition ($M = .43$, $SD = .23$) than in the conventional condition ($M = .23$, $SD = .20$).

In order to determine if the response times to the different response types ("meaningful"/"not-meaningful") differed across conditions (novel33% and conventional), a repeated measures ANOVA was conducted with within-subject factors of response and condition. Main effects of response and condition were found. These were qualified by a Response by Condition interaction, $F(1,165) = 31.53$, $p < .001$, $\eta^2_p = .160$. To further explore this interaction, paired sample t-tests were calculated. Target novel word-pairs were judged as "not-meaningful" more slowly in the novel33% condition (M = 924ms, SD = 127ms) than in the conventional condition (M = 784ms, SD = 112ms), $t(178) = 11.35$, $p < .001$, $d = .838$. No significant differences were found for "meaningful" responses.

This analysis was re-run to only include participants whose ERP recording was kept. A main effect of response and a main effect of condition were found. These were qualified by a Response by Condition interaction $F(1, 156) = 47.04$, $p < .001$, $\eta^2_p = .232$. To further explore this interaction, paired sample t-tests were conducted. Again, a significant main effect of condition was found for "not-meaningful" responses, $t(176) = 10.98$, $p < .001$, $d = .890$. Target novel word-pairs were judged "not-meaningful" more slowly in the novel33% condition (M = 1042ms, SD = 277ms) than in the conventional condition (M = 799ms, SD = 126ms). Additionally, a significant main effect of condition was found for "meaningful" responses, $t(159) = -2.52$, $p = .013$, $d = -.191$. Target novel word-pairs were judged "meaningful" more slowly in the conventional condition (M = 1169ms, SD = 372ms) than in the novel33% condition (M = 1069ms, SD = 372ms).

Correlation analysis. To determine whether or not it was the same word-pairs that were being judged "meaningful" across conditions, a correlation between conditions was calculated for the proportion of meaningful responses for each target novel word-pair. As in experiment one, the resulting correlation was found to be strong, positive and significant. For all participants run in this experiment the correlation was $r(180) = .83$, $p < .001$; for only the participants whose ERP data was kept the correlation was $r(180) = .69$, $p < .001$. The correlations indicate that the same novel word-pairs tended to be judged "meaningful" across conditions, and therefore that participants were not just responding randomly to meet an implied target.

ERP Analysis

N400. The N400 has been repeatedly found over central and parietal sites (for a review see Kutas & Federmeier, 2011). For this reason, many researchers only include central and parietal sites in N400 analysis (for an example of some prominent ERP researchers who have used this method see Federmeier, Van Petten, Schwartz & Kutas, 2003). In the current experiment, the caps used had two frontal and one temporal electrode with technological problems. Therefore, these electrodes could not be included in analyses. However, since only central and parietal areas are crucial to N400 analysis, this problem was resolved by only including central and parietal sites in the N400 analysis of the current study. Mean amplitudes in the 350-450ms time window (consistent with the time window inspected by Arzouan et al., 2007) after the onset of the second word of each word-pair were investigated for the "meaningful" responses only. These mean amplitudes at individual electrodes were averaged to create mean amplitudes for Left-Central (C3, CP3), Left-Posterior (P3, P7), Medial-Central (Cz, CPz), Medial-Posterior (Pz), Right-Central (C4, CP4) and Right Posterior (P4, P8) positions.

Two analyses were conducted. The first was a between-conditions analysis to investigate whether participants differed in responses to target novel word-pairs in the N400 time period. Only "meaningful" responses were considered which means that only differences in the successful strategies were investigated. ECCo predicts that participants in the novel33% condition should process novel word-pairs analytically and therefore have more negative N400s, while participants in the conventional condition should process novel word-pairs associatively and therefore have less negative N400s. The second analysis was a within-conventional-condition analysis to see if participants in the conventional condition differed in their "meaningful" responses to novel versus conventional word-pairs within the N400 time period. ECCo predicts that since both types of word-pairs should undergo similar associative processing, no differences should be found.

Mean amplitudes from the 350-450ms time window for the previously mentioned six positions were entered into a 2 condition (conventional/novel33%) × 3 laterality (left/medial/right) × 2 caudality (central/posterior) repeated measures ANOVA to investigate whether participants in different conditions differed in responses to target novel word-pairs in the N400 time period. A main effect of caudality (central/posterior) was significant, $F(1, 54) = 6.54$, $p = .017$, $\eta^2_p = .195$. As expected for the N400, mean amplitudes were more negative for central (*M* = -2.80, *SD* = .62) than for posterior (*M* = -1.53, *SD* = .48) sites.

A Laterality (left/medial/right) by Condition (conventional/ novel33%) interaction approached significance, $F(2, 54) = 2.99$, $p = .059$, $\eta^2_p = .100$. Because a condition effect was predicted, this interaction was further explored. Data was collapsed across caudality (central/posterior) and follow up independent sample t-tests were calculated to determine whether mean amplitudes in the 350-450ms time window differed by condition in left, medial, and/or right locations. The difference between mean areas in the different conditions approached significance for medial locations, $t(27) = 1.80$, $p = .082$, $\eta^2_p = .108$, with participants in the novel33% condition having more negative N400s (*M* = -3.49, *SD* = 3.15) than participants in the conventional condition (M = -.91, SD = 4.31). No other findings approached significance. For visual examination of the N400 effect, see Figure 1. Means and standard deviations for areas (μV.ms) used in these analyses can be seen in Table 5.

Table 5

Means and standard deviations for areas (μV.ms) of the N400 in different lateral (left/central/right) and caudal (central/posterior) positions for target novel word-pairs judged as “meaningful” in the conventional and novel33% conditions.

	Conventional condition		Novel33% condition	
Locations	Mean	SD	Mean	SD
Left-central	-2.56	3.15	-3.40	2.39
Medial-central	-1.20	5.14	-4.20	2.97
Right-central	-1.92	3.89	-3.51	2.69
Left-parietal	-1.59	3.66	-1.33	2.18
Medial-Parietal	-.62	3.96	-2.78	3.46
Right-Parietal	-.85	2.59	-1.99	2.36

Figure 1.

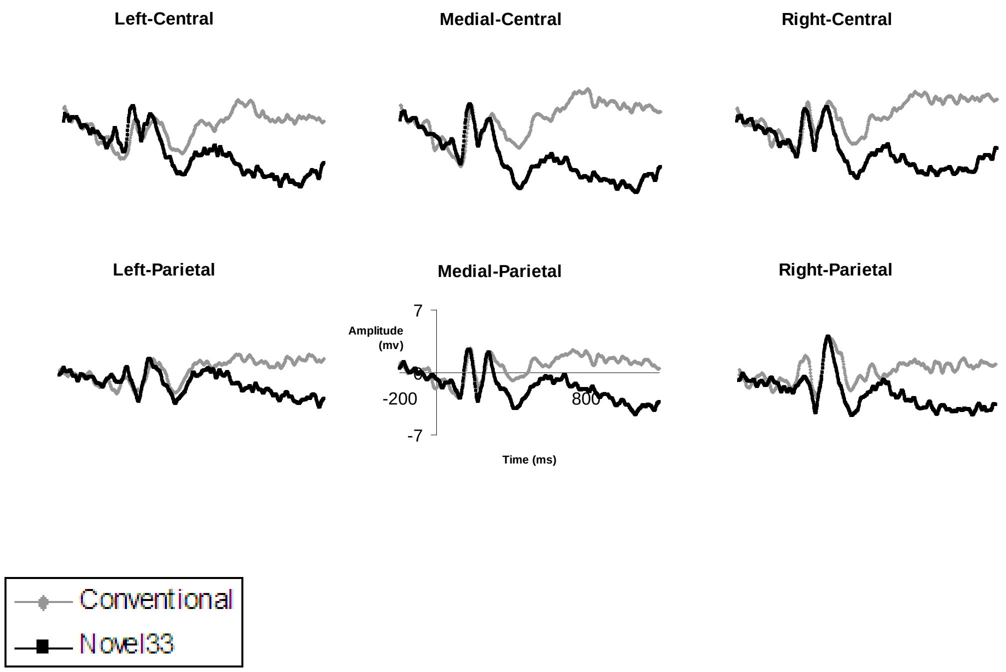


Figure 1. ERP data elicited by novel word-pairs was averaged across participants and is shown for each condition. Onset of the second word for each novel word-pair occurred at 0 ms. At some sites the N400 effect can be seen at around 400ms, and at all sites a late component can be seen with an onset of around 500ms.

To determine if participants in the conventional condition differed in their processing “meaningful” novel and “meaningful” conventional metaphoric word-pairs, a repeated measures of ANOVA with within-subject factors of word-pair type (conventional metaphoric/novel), laterality (left/medial/right) and caudality (central/posterior) was calculated. A main effect of word-pair type was found, $F(1, 13) = 5.16$, $p = .041$, $\eta^2_p = .284$, N400 areas were more negative in response to novel (M = -1.01, SD = .74) than conventional metaphoric (M = 1.63, SD = 1.16) word-pairs. See Table 6 for these means and standard deviations.

Table 6

Means and standard deviations for areas (μV.ms) of the N400 in different lateral (left/central/right) and caudal (central/posterior) positions for different types of word-pairs (conventional metaphors/ novel) judged as "meaningful" in the conventional condition.

	Conventional Metaphors		Novel word-pairs	
Locations	Mean	SD	Mean	SD
Left-central	1.07[a]	4.49	-2.15[a]	3.71
Medial-central	.99	5.46	-1.76	3.83
Right-central	.1.58	4.84	-1.38	3.45
Left-parietal	1.60[b]	3.15	-.63[b]	1.91
Medial-Parietal	2.20	5.66	-.47	3.42
Right-Parietal	2.35	4.96	.32	2.53

Letters indicate a significant difference at the .05 level

Figure 2.

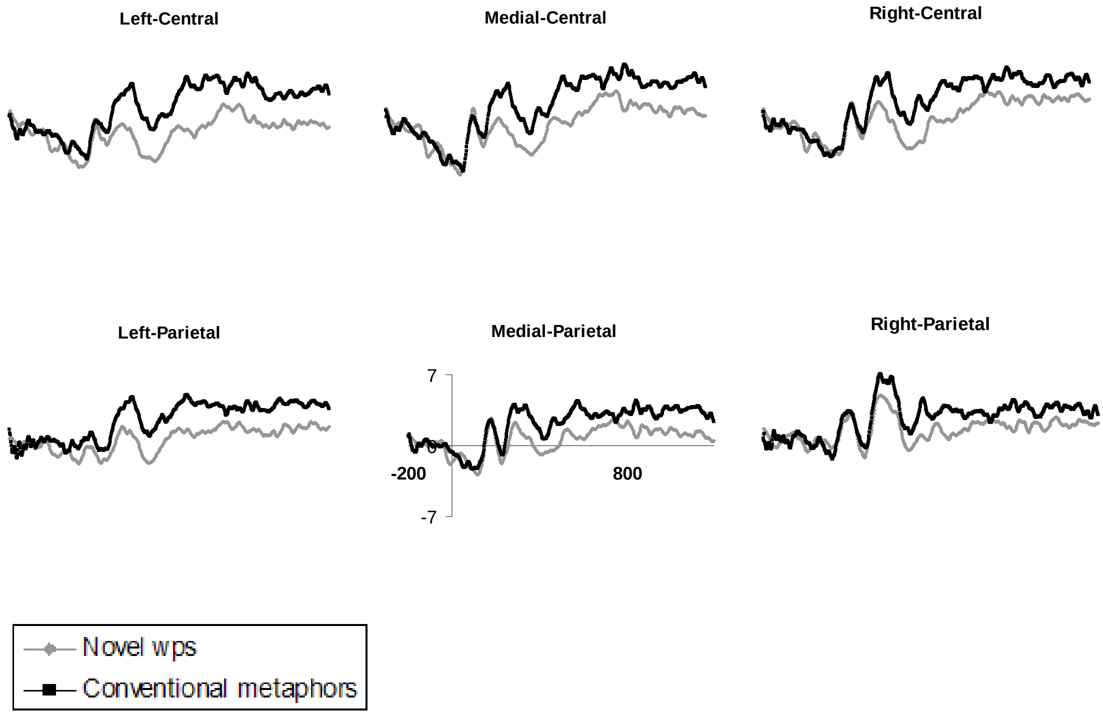


Figure 2. ERP data elicited by novel and conventional metaphoric word-pairs was averaged across participants in the conventional condition and is shown above. Although quantitatively different, the neural response to novel versus conventional metaphoric word-pairs appears qualitatively similar within this condition.

Late Effects. Visual inspection of the waves created for the N400 analysis revealed differences in the late time window between conditions (see Figure 1). Therefore, two exploratory analyses were conducted: a between-conditions analysis to investigate if participants in the two conditions differed in responses to target novel word-pairs at the late time period, and a within-conventional-condition analysis to see if participants in the conventional condition differed in their responses to novel versus conventional metaphoric word-pairs in the late time period.

To examine differences between meaningful responses in the two conditions, an initial 2 condition (conventional/novel33%) × 3 laterality (left/medial/right location) × 2 caudality (central/posterior) repeated measures of ANOVA was calculated for the mean amplitudes in the time window between 500ms and 900ms after the onset of the second word in each target novel word-pair. This time window was used because most researchers who investigate late effects analyze a similar time period (see Arzouan et al., 2007; Coulson &Van Petten, 2002; DeGrauwe, Swain, Holcomb, Ditman, & Kuperberg, 2010; Ibanez et al., 2011; Kazmerski et al., 2003; Pynte et al., 2006; Regel, Gunter & Friederici, 2010). A main effect of condition was found, $F(1, 21) = 5.26$, $p = .032$, $\eta^2_p = .200$. Participants in the novel33% condition showed more negative late effects in the 500-900ms time window ($M = -1.81$, $SD = .83$) than participants in the conventional condition ($M = .94$, $SD = .87$).

Condition interacted marginally with both laterality, $F(2, 42) = 2.83$, $p = .071$, $\eta^2_p = .119$, and caudality, $F(1, 21) = 4.05$, $p = .057$, $\eta^2_p = .162$. Follow up analyses indicated that the condition effect was observed primarily over right, $F(1, 21) = 4.98$, $p = .033$, $\eta^2_p = .192$, and medial positions, $F(1,21) = 7.36$, $p = .013$, $\eta^2_p = .260$, and in central sites, $F(1, 21) = 5.78$, $p = .025$, $\eta^2_p = .216$. Again, mean amplitudes for participants in the novel33% condition were more negative in these positions than mean amplitudes for participants in the conventional condition. See Table 7 for these means and standard deviations. See Figure 1 for visual inspection of the late effect.

Table 7

Means and standard deviations for areas (μV.ms) of the late effect in different caudal positions (central/posterior) and lateral positions (left/medial/right) for participants in different conditions.

		Conventional		Novel33%	
Central		Mean	SD	Mean	SD
	Left	.14	1.17	-2.49	1.12
	Medial	1.24	1.14	-3.36	1.09
	Right	1.25	1.19	-2.76	1.14
Posterior					
	Left	.96	.77	.01	.74
	Medial	1.24	.82	-1.18	.78
	Right	.80	.80	-1.09	.76

Whether or not the novel and conventional word-pairs judged as “meaningful” were processed in a similar fashion to each other in the conventional condition was investigated. A repeated measures ANOVA for mean areas in the 500-900ms time window was calculated with within-subject factors of word-pairs type (conventional metaphoric/ novel), lateral position (left/medial/right) and caudal position (central/posterior). A main effect of word-pair type approached significance, $F(1, 13) = 3.69$, $p = .077$, $\eta^2_p = .221$. Because a word-pair effect was predicted, this data was collapsed across caudality, and repeated measures ANOVAs were conducted with word-pair type as a within-subjects factor for each of the three lateral positions. There was a significant main effect of word-pair type at left, $F(1, 13) = 7.82$, $p = .015$, $\eta^2_p = .375$, but not at other, lateral positions. Neural activation was more negative in response to conventional metaphoric (*M* = 3.38, *SD* = .76) than novel (*M* = 1.00, *SD* = .74) word-pairs. See Figure 2 for visual inspection of this effect.

LRP. In order to check that the analysis was appropriate for measuring the LRP effect (in terms of filtering, electrode placement and data reduction), a within-subjects response-locked comparison was conducted for “meaningful” versus “not-meaningful” responses. The amplitude over the left and right motor cortices from 300ms before response execution was calculated for each participant for “meaningful” and “not meaningful” responses separately. Next, average amplitudes over the right motor cortex were subtracted from average amplitudes over the left motor cortex for each type of response (C3-C4), and the resulting difference waves were compared in a paired samples t-test. Results of this test were significant $t(39) = -5.27$, $p < .001$, $\eta^2_p = .416$, with “meaningful” responses having more negative response-locked LRPs (*M* = -1.77, *SD* = 2.50) than “not-meaningful” responses (*M* = .46, *SD* = 1.65). The right hand (left hemisphere) was always used to press the “meaningful” key, and the left hand (right hemisphere) was always used to press the “not-meaningful” key. The C3-C4 comparison should render waves more negative for increased relative negativity over the left hemisphere, and more positive for relative negativity over the right hemisphere. Therefore, these results are consistent with a LRP, as they showed increased negativity over the hemisphere which was preparing a response. Because this is a classic response-locked LRP effect, it can be assumed with confidence that the analysis was suitable to measure differences in LRP waves. See Figure 3 for a visual representation of this data.

Figure 3. The response-locked LRP

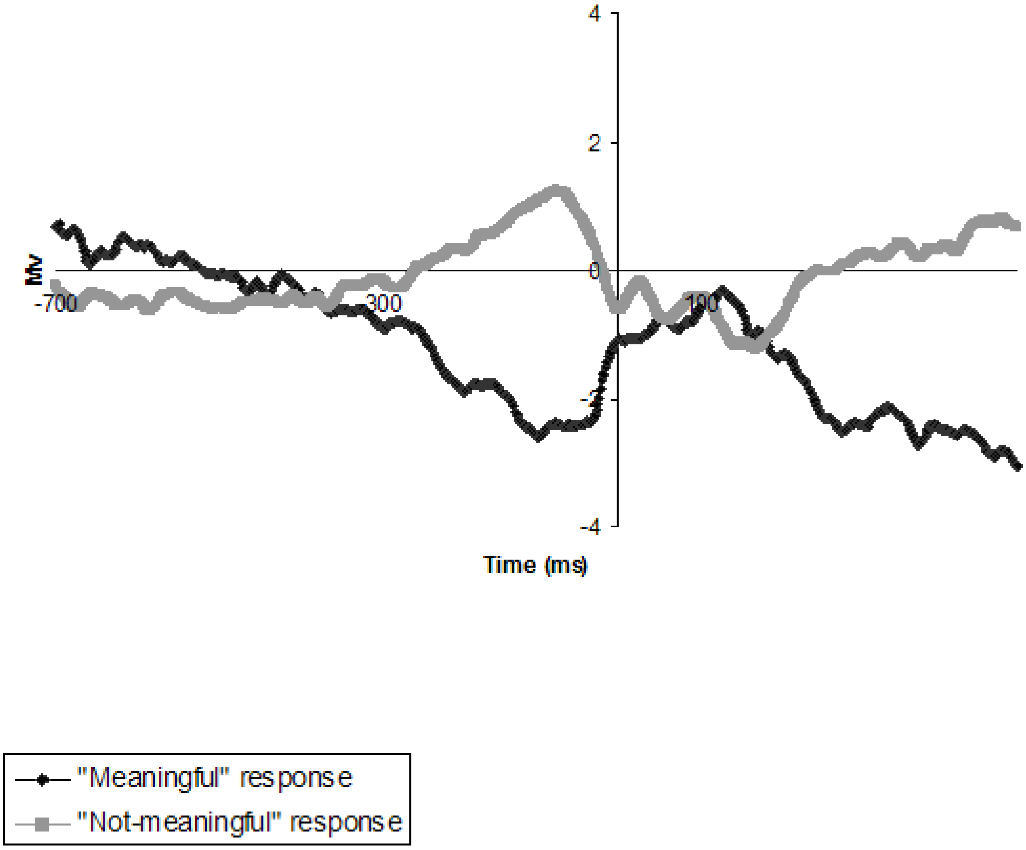


Figure 3. ERP data elicited prior to response is shown in the form of a C3-C4 comparison. Negative waves show increased activity over the left hemisphere relative to the right hemisphere. Positive waves show increased activity over the right hemisphere relative to the left hemisphere. Responding occurred at 0 ms on the graph. Results consistent with contralateral response preparation are shown.

To calculate whether the LRP began before the onset of the second word of each word-pair (which was investigated as a measure of response bias), the stimulus-locked amplitude from 300 to 0ms before the onset of the second word of each pair was calculated over the left and right motor cortices for each participant. Because participants did not know what type of word-pair they would be presented with before they saw the second word, their LRP measured prior to the second word should not differ for different word-pair types, and so the 300ms before response for all word-pair types were included. If participants in the novel33% condition were anticipating a "meaningful" response more than participants in the conventional condition (thus explaining behavioural results), then their LRP prior to the onset of the second word of each word-pair is expected to be more negative over the left hemisphere. Average amplitudes over the right motor cortex were subtracted from average amplitudes over the left motor cortex for each participant (C3-C4). The remaining difference waves were then compared for participants between conditions in an independent samples t-test. No significant differences were found, $t(27) = -1.06$, $p = .299$, $\eta^2_p = .040$, indicating that participants did not differ in their response preparation between conditions before they saw the second word of each word-pair. Participants in both conditions showed positive LRPs which reflect their preference to select the "not-meaningful" answer more than the "meaningful" answer, as reflected in the behavioural results. The stimulus-locked LRP can be seen in Figure 4.

Figure 4. The stimulus-locked LRP

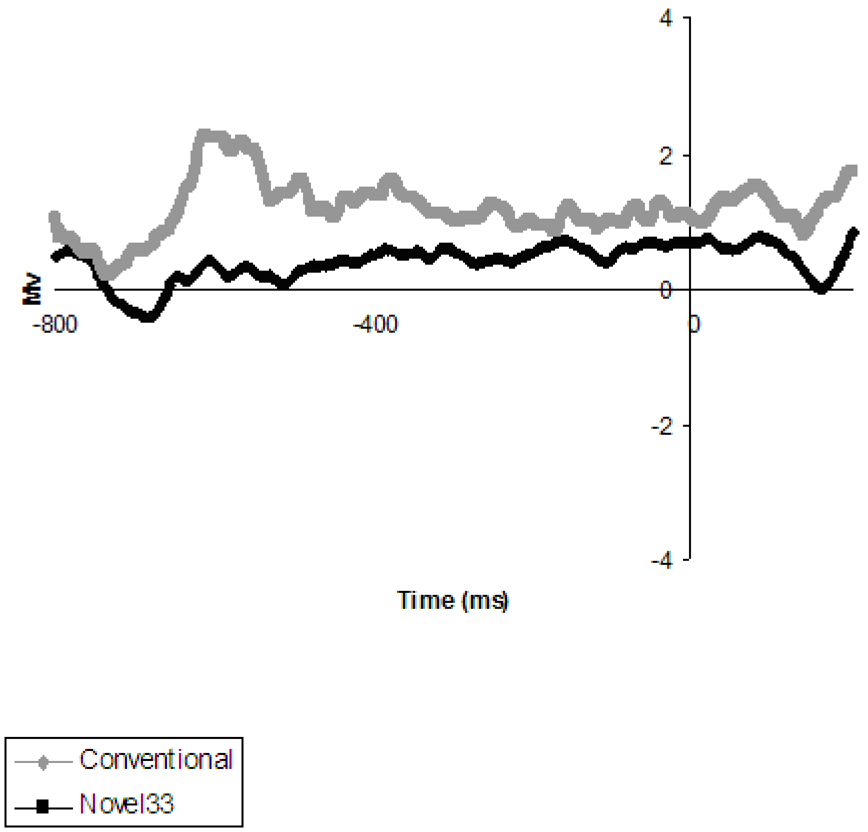


Figure 4. ERP data elicited prior to the second word of each word-pair is shown in the form of a C3-C4 comparison. Negative waves show increased activity over the left hemisphere relative to the right hemisphere. Positive waves show increased activity over the right hemisphere relative to the left hemisphere. Stimulus onset of the second word occurred at 0 ms on the graph. There were no significant differences in response preparation between conditions prior to the onset of the second word.

Discussion

Behavioural Data

Behavioural results were similar regardless of whether the data for participants whose ERP data was excluded was kept in analyses or not. Therefore, (with one exception) behavioural results will not be discussed separately for participants who were, and who were not, kept in ERP analysis.

As in experiment one, the behavioural results of experiment two showed that participants who saw novel word-pairs alone (in the novel33% condition) found them more meaningful than participants who saw novel word-pairs mixed with other types of word-pairs. This replication of results is significant because it reveals the robustness of the effect. This replication of experiment one supports the dual processing hypothesis that experimental context can lead participants to select different processing strategies when assessing novel word-pairs, and thereby influence whether or not they attempt to create meaning for them. This replication of experiment one also supports the idea that experimental context can influence which process (comparative or categorical) is used to process a metaphor. Comparative/categorical processing is supported because current results suggest that a novel metaphor can be processed in two qualitatively different ways leading to different judgements (note that novel word-pairs which were judged “meaningful” in the current study could only be interpreted figuratively and therefore can be called novel metaphors). However, results do not specifically divulge whether or not the two processes observed are categorisation, comparison, or some other processes.

Participants in the novel33% condition were significantly slower to judge novel word-pairs as “not-meaningful” than participants in the conventional condition. This result indicates that participants in the novel33% condition tried harder (and thus spent more time) trying to find meaning in novel word-pairs. This result thereby fits with analytical processing, as increased response times can reflect increased effort, and analytical processing was predicted to take more semantic effort than associative processing.

In contrast, participants in the novel33% condition (whose ERP data was kept) were significantly faster to judge novel word-pairs as “meaningful” than participants in the conventional condition (whose ERP data was kept). This finding is somewhat concerning for a dual processing interpretation, as it would predict that participants in the novel33% condition should be relatively slower to respond due to the taxing nature of analytical processing. However, the word-pairs used were adjective-noun couplets. Adjectives and nouns should have many affordances for coupling because adjectives are, by definition, words that are used to modify nouns. Perhaps then, adjectives and nouns are usually successfully meshed in a specific fashion and so this is the first way in which the analytical process attempts to mesh novel word-pairs. Consequently, using the analytical processing strategy should lead to quick responses when this adjective-noun strategy is successful. If the given word-pairs cannot be meshed in this particular way then the analytical process (which is more motivated than the associative process) might keep trying other ways of meshing, thus explaining longer "not-meaningful" responses. One way to test this speculative interpretation would be to re-run the current experiment, using noun-noun couplets as novel word-pairs. Regardless of which response they make ("meaningful"/"not-meaningful"), participants in the novel33% condition are expected to be slower to judge novel noun-noun word-pairs than participants in the conventional condition. Such a finding (combined with current results), would indicate that high task demand does cause people to rely on the analytical processing strategy, and that this strategy is usually slow, but that the meshing of adjective-noun couplets presents a special case.

N400

Investigation of the N400 was conducted in order to examine the dual processing hypothesis. Because it should involve more semantic effort, participants processing novel word-pairs using the analytical system should have more negative N400s than participants processing the novel word-pairs using the associative system. It was predicted that if experimental context influences the processing strategy that is used (with harder task demands/ less fluent contexts giving preference to the analytical system), then participants in the novel33% condition (which was harder/ less fluent because it involves fewer easy decisions) should have more negative N400s to novel word-pairs than participants in the conventional condition. A laterality (left/medial/right) by condition (conventional/ novel33%) interaction reached marginal significance. This indicates topographical differences in N400s effects for the two different conditions, which (in line with dual processing) may imply different underlying neural sources (Nunez, 1981). Follow-up tests showed that the difference in mean amplitudes in the different conditions was largest at medial locations, which is consistent with a typical N400 effect. In line with dual-processing predictions and corroborating behavioural data, N400 magnitudes to novel word-pairs at this location were more negative for participants in the novel33% condition than for participants in the conventional condition. Because this effect was in the predicted location (Federmeier, Wlotko, Ochoa-Dewald, & Kutas, 2007), predicted time window, and because this effect occurred in the predicted fashion, it is therefore interpreted as denoting the N400 effect. However, because the effect of condition did not reach significance at any single location, these results are not fully reliable. The lack of statistical significance is likely to represent a lack of statistical power due to a small sample (because many participants had to be excluded from analyses). To further support these results, more participants should be run in this experiment. Meanwhile, N400s to "meaningful" novel word-pairs are at least indicative of support for the dual processing prediction that participants with a higher task demand (in the novel33% condition) use a more analytical processing strategy, and that participants with a lower task demand (in the conventional condition) use a more associative processing strategy.

Participants within the conventional condition were found to differ in their N400 magnitudes in response to novel versus conventional metaphoric word-pairs. N400s to novel word-pairs were significantly more negative than those to conventional metaphoric word-pairs, replicating the finding of Arzouan et al. (2007), and indicating that novel metaphors take more semantic effort to process than conventional metaphors.

Because differences in the N400 for novel versus conventional metaphoric word-pairs within the conventional condition differed over left sites, but differences in the N400 for novel word-pairs between conditions differed over medial sites, this shows marked differences in the N400 effect distribution for the within- versus between- subjects comparisons. These marked differences are likely to have arisen because (as can be seen by visual inspection of the waves), processing of target novel word-pairs in the conventional condition differed *qualitatively* from the processing of target novel word-pairs in the novel33% condition, but processing of novel and conventional metaphoric word-pairs within the conventional condition differed *quantitatively*. This difference may have occurred because different neural sources were recruited for processing word-pairs in the two conditions.

Often, (e.g. Lai et al., 2009), but not always (e.g. Arzouan et al., 2007) experiments investigating the N400 delay responding so that response preparation does not interfere with these waves. Responding was not delayed in the current study, as delayed responding would have interfered with the LRP analysis (participants would soon learn that they did not need to respond immediately after seeing the second word in each word-pair, so if responding was delayed, bias would not likely appear in the time window prior to the onset of the second word). A follow-up study in which responding is delayed would indicate whether morphology of the waves found in the current study was affected by response preparation.

Late Effects

No predictions about late effects were made in the current study. Nevertheless, visual inspection of the waves revealed a difference in late effects between groups, and so this was statistically examined. Late unpredicted positivities in waves were found in the between- condition analysis for participants in the conventional relative to the novel33% condition, and in the within-conventional condition in response to conventional relative to novel word-pairs analysis. Often positivity over late time windows in response to figurative statements is interpreted as the "semantic P600". Although the P600 was originally thought to index syntactic anomalies and the N400 was thought to index semantic anomalities (See Kutas & Hillyard, 1980 for an original description of the N400, and Osterhout & Holcomb, 1992, 1993 for an original description of the P600), recent research has shown the line between semantic and syntactic processing to be more blurred (for an overview of unexpected N400 and P600 effects, see Kutas, Van Petten & Kluender, 2006). There are a variety of different theoretical interpretations of the "semantic P600" (for a review of these see Bornkessel-Schlesewsky & Schlesewsky, 2008). Conservatives like Kim and Osterhout (2005) hold that the semantic P600 occurs when syntactically correct phrases (e.g. "*the meal was devouring…")* are re-interpreted to make semantic sense (e.g. to assume that it meant "*the meal was devoured*"). This reinterpretation means that the original phrase is presumed to have been presented in the syntactically incorrect form, thereby initiating the P600. Other researchers hold that the semantic P600 arises when thematic role assignment is costly (e.g. Hoeks, Stowe & Doedens, 2004). Recent theorists proposed that the semantic P600 arises when parallel processing streams come to different interpretations (Bornkessel-Schlesewsky & Schlesewsky, 2008; Kuperberg, 2007). Alarmingly, these so called "semantic P600s" often differ in morphology (see the "P600s" found by Frenzel, Schlesewsky & Bornkessel-Schlesewsky, 2011; and compare these to the "P600s" found by Regel et al., 2010).

In other studies, late positivities found in response to figurative statements are referred to as the Late Positive Complex (LPC) although there seems to be no clear way that these positivities differ from those referred to as the semantic P600. For example, DeGrauwe, et al. (2010), found what they termed a “late positive component” to be larger to target metaphors than to target literal words in sentences. Ibanez et al. (2011) found what they termed a “LPC-like” component to be more negative in response to metaphoric than literal gestures.

The late negativity found by Ibanez et al. (2011) reflects another disparity in the literature- the valence of late effects to figurative statements is inconsistent. Similar to the negativity found by Ibanez et al., the current study found a late negative going late wave (in the novel33% condition) to word-pairs which were expected to be processed figuratively. Also fitting with results of the current study, Arzouan et al. (2007) found a late negativity to novel metaphoric word-pairs when compared to conventional and literal word-pairs. Arzouan et al. referred to their late effect as an “LPC” although at other points they referred to it as a “late negativity” and in one paragraph as the “LPC (P600)” (Arzouan et al., 2007, p. 76). Arzouan et al. proposed that their late negativity might reflect the additional semantic elaboration required to find meaning in these word-pairs. Consistent with this proposal, Pynte et al. (1996) found that regardless of whether target stimuli were conventional or novel metaphors, later components were more negative if they were presented in an irrelevant context compared to when they were presented in a relevant context. Similarly, Friederici, Steinhauser and Frish (1999) also reported a late negativity in both semantically incongruent and syntactically + semantically incongruent conditions of their experiment, but not in syntactically incongruent conditions. They pronounced this late negativity to be distinct from the P600 which they found only in syntactically incongruent and syntactically + semantically incongruent conditions, and which was reduced by the negative effect in the syntactically + semantically incongruent condition. Therefore, Friederici et al. proposed that two distinct neural sources may produce components which overlap in the time domain (one positive, the other negative).

Overall, a variety of late effects that differ in morphology and valence, and which are inconsistent in the way they behave, have been found in response to figurative statements. These late effects are referred to using different terms, and terminology is not consistent across morphology, valence or behaviour. It has been proposed that these ambiguities may arise because several different ERP components occur during the same time period, although sufficient evidence to dissociate these components is not yet available. The current study was not designed to differentiate between possible explanations of the late effect. Better understanding of late effects to figurative language is necessary before interpretation of the late effects in the current study is feasible. Nevertheless, the finding that there were large late effects after the N400 is consistent with some of the research on figurative language, and suggests that experimental context has profound effects on late stages of meaning construction.

Wave Morphology

The dual processing hypothesis predicts that participants in the novel33% condition should process novel word-pairs analytically, that participants in the conventional condition should process novel word-pairs associatively and therefore that there should be a qualitative difference in waveforms to target novel word-pairs between conditions. Visual inspection of the neural activation supports this prediction, with a negative going late wave in the novel33% condition, and a positive going late wave in the conventional condition (a difference which is significant at medial and right positions).

The dual-processing hypothesis also predicts that because participants in the conventional condition should use the associative processing strategy, conventional metaphors and novel metaphors within this condition that are judged as “meaningful” should be processed in a similar manner. Results show that wave forms for these two word-pair types within the conventional condition looked morphologically similar across all time periods (both are positive in the late period). However, both N400 and late analyses showed a significant difference in the magnitude of these waves over left (central and parietal) sites, indicating a quantitative difference in processing between conventional metaphoric and novel word-pairs.

This qualitative difference in ERP morphology shown in the between-condition comparison also supports the hypothesis that which process (comparative or categorical) a novel metaphor is processed with can be influenced by experimental context. The novel word-pairs in this study required figurative processing to be found meaningful, therefore, novel word-pairs which were judged “meaningful” can be considered novel metaphors. However, while ERP morphology reveals these qualitative differences between conditions, it does not reveal whether these separate processes are categorization and comparison, or not. Therefore, while this data is consistent with experimental context determining comparative/categorical processing selection, it does not establish it as fact.

LRP

Response-locked LRPs were calculated and compared for "meaningful" versus "not-meaningful" responses for all participants. A significant difference was found, indicating increased negativity over the left hemisphere when a person was about to use their right hand to select "meaningful", and increased negativity over the right hemisphere when a person was about to use their left hand to select "not-meaningful" responses. These results fit with an LRP definition, and indicate that the analyses used for LRP effects in the current study were appropriate for measuring LRP effects.

The LRP prior to onset of the second word of each word-pair was also analysed. If participants were already preparing a response in the interval between words, then they must have been biased towards one response over the other because they had not yet seen the stimulus to which they should respond. Stimulus- locked LRPs of participants in the conventional condition were compared to those of participants in the novel33% condition for all word-pairs. No difference was found between conditions, indicating that the different experimental contexts do not cause people to set different criteria for meaningfulness. This consequently rules out bias as an explanation for why people who see novel word-pairs with clearly meaningful word-pairs judge them as meaningful less often than people who see them with additional novel word-pairs.

It must be noted, however, that the current response-locked LRP analysis may have been more sensitive than the current stimulus-locked LRP analysis. There is likely to be less variance in motor preparation in the 300ms prior to responding than in motor preparation in the 300ms before stimulus onset. Essentially, stimulus-locked LRP's are more variable than response-locked LRP's. Therefore, failure to find a difference in stimulus-locked LRP's between conditions could reflect a power issue. Some other researchers who have looked at stimulus-locked LRPs have used measures that produce larger activation over contralateral motor cortices than a button press. For example, Steinhauser et al. (2009) used dynamometers (hand-held devices used for measuring the force of grip strength) in their experiment, with participants squeezing either the dynamometers held in the left or in the right hand to respond. In another example, participants in the study by Masaki et al. (2004) used force sensitive key-releases to respond. Therefore, to further rule out bias as a cause for the behavioural results in experiment one and two, experiment two should be repeated using a measure which produces greater contralateral activation during response preparation.

General Discussion

Overall, this study examined the influence of experimental context on the meaning construction process, and on metaphor processing. Target word-pairs were presented mixed with different types of other word-pairs in different conditions. It was investigated whether or not the contextual presence of these other types of word-pairs (the "experimental context") would influence if and how meaning was constructed for target word-pairs. A literature review suggested that experimental context could have influenced the meaning construction process in three different ways- by influencing information processing strategy selection, bias, and process priming.

Fitting with the predictions based on the dual processing and fluency literature, it was found that participants were more likely to call target novel word-pairs meaningful when they saw them in the contextual presence of additional novel word-pairs, than when they saw them in the contextual presence of clearly meaningful word-pairs. This provides evidence that different experimental contexts elicit different task demands/fluency, which in turn influence which information processing strategy a person uses to assess a novel word-pair. The analytical processing strategy (presumably used by participants who saw only novel word-pairs) involves meaning construction, whereas the associative processing strategy (presumably used by participants who saw novel word-pairs with other types of word-pairs) does not; therefore results indicate that different contexts can influence whether or not the meaning construction process is relied upon.

However, response time data in experiment two did not map onto dual processing predictions. For participants whose ERP data was kept, a response type by condition interaction was found, with participants in the conventional condition being faster to say "not-meaningful", and participants in the novel33% condition being faster to say "meaningful". It is suggested that people should know the way in which novel adjective-noun couplets most frequently successfully combine, and thus when processing analytically, should attempt this type of meshing before other less common ways. Successful combination via this process should result in swift "meaningful" responses. However, if this process is not successful, these participants may attempt other forms of meshing, thereby having slowed “not-meaningful” responses relative to participants who process novel word-pairs associatively, and who do not try to mesh them at all.

The dual processing explanation was bolstered by ERP results- participants hypothesised to use the analytical strategy had more negative N400s to target novel word-pairs than participants hypothesized to use the associative strategy. This provides evidence that participants were using the hypothesized strategies, as the analytical strategy should involve more semantic effort and therefore make more negative the N400 when compared to the associative strategy. Also, wave morphology revealed distinctive processing for novel word-pairs in the two conditions, but qualitatively similar processing for conventional and novel word-pairs in the conventional condition. Wave morphology thus provides further evidence for dual processing and fluency hypotheses, as these predict that “meaningful” responses to novel word-pairs should be processed in a qualitatively different manner in novel33% versus conventional conditions (analytical versus associative processing); but that “meaningful” responses to novel word-pairs and conventional word-pairs should be judged in a qualitatively similar manner within the conventional condition (associatively).

It was of concern for dual processing predictions that results may have arisen due to participants randomly responding to meet an implied target number of “meaningful” responses. However, positive correlations between the proportions of times each word-pair was judged “meaningful” were found across conditions suggesting that similar word-pairs were called “meaningful” across conditions, thus indicating that participants were not simply responding randomly. It was still possible, however, that the results occurred due to different contexts causing participants to adjust their criteria for meaningfulness differently (i.e. causing them to differ in how much evidence they required before they judged something as “meaningful”). Experiment one could not rule this possibility out, as responses did not differ in the novel33% and novel50% conditions (where implicit criteria was different but context was identical), and so participants must not have been conforming to implied targets. Therefore, the criteria that participants were using and whether or not it depended on experimental context remained unknown. However, the finding in experiment two that stimulus-locked LRPs did not differ between conditions showed that the participants did not differ in their response preparation before seeing the second word of each word-pair. Therefore, participants in different conditions in experiment two did not differ in bias/meaningfulness criteria. Finally, no evidence for process priming was found, although this may be due to experimental design.

Metaphor

ERP wave morphology showed that while novel metaphors (the novel word-pairs that were judged meaningful) were processed in a qualitatively different manner to conventional metaphors in some contexts, they were processed in a qualitatively similar manner to conventional metaphors in other contexts. This is challenging for Glucksberg and Keysar (1990) who proposed that novel metaphors are always processed in a different manner to conventional metaphors. Current findings also challenge the findings of Glucksberg and Haught (2006), which indicate that while novel metaphors can be processed in two different ways, the way in which novel metaphors are processed depends on their lexical aptness. The current results show that the same novel metaphors (which therefore are equally lexically apt across conditions) can be processed in different ways depending on experimental context. This finding therefore indicates that theories of metaphor processing should consider context as a significant factor in determining how a novel metaphor is processed. While it appears that an easy context will elicit reliance on metaphor categorization (which corresponds best to associative processing) and a relatively difficult context will elicit reliance on metaphor comparison (which corresponds best to analytical processing), the current results can not specify whether or not the two processes found truly represent categorization and comparison.

An experiment is proposed to determine whether or not the two processes observed are categorisation and comparison. Firstly experiment one from the current study should be re-run using noun-noun metaphors (for reasons that will be obvious below). If similar findings to the current study are found, then a follow-up study should be conducted to determine whether categorical and comparative processes can explain these findings. There should be four conditions. The first condition should have the same word-pairs and verbal information as the conventional condition of noun-noun experiment proposed above. However, these word-pairs should be presented in the following sentence form: "___ *is a* ___" (the noun-noun word-pairs are necessary as adjective-noun pairs would not fit into this type of sentence). The second condition should have identical stimuli and information to condition one except that the word-pairs should be presented in the following sentence form: “___ *is like a* ___”. The third condition should have the same word-pairs and verbal information as the novel33% condition of the noun-noun experiment proposed above. However, these word-pairs should be presented in the following sentence form: “___ *is a* ___”. The fourth condition should have identical stimuli and information to condition three except that the word-pairs should be presented in the following sentence form “___ *is like a* ___”. If experimental context can influence whether the same novel word-pairs are processed categorically or comparatively, then participants in the second condition of the proposed study should have more negative N400s to the final word of each target novel word-pair relative to participants in the first condition of the proposed study. This is because participants in these conditions should be processing target novel word-pairs categorically (due to their experimental context), and therefore find them easier to process in the metaphor form (which is presented in condition one but not two). Participants in the third condition of the proposed study should have increased N400s to the final word of each target novel word-pair relative to participants in the fourth condition of the proposed study. This is because participants in these conditions should be processing target novel word-pairs comparatively (due to their experimental context), and therefore find them easier to process in the simile form (which is presented in condition four but not three). Also, wave morphology for the two processes should be distinct from each other. As in the current study, a negative going late effect is expected for participants in the third and fourth (novel33%) conditions, while a positive going late effect is expected for participants in the first and second conditions. If the predicted results are found, then this would indicate that the easier task provided by the context of the conventional condition causes participants to judge novel word-pairs using categorical processing, and that the harder task provided by the context of the novel33% condition causes participants to judge novel word-pairs using comparative processing.

If the results of the proposed experiment support the hypothesis that differential task ease (initiated by differential experimental context) can influence whether people use associative or analytical processing to process novel metaphors, then past metaphor experiments might also be explained in terms of task ease. Bowdle and Gentner (2005) found participants to prefer to process conventional metaphors as categorisations, and novel metaphors as comparisons. Relative to the conventional metaphors, the task of processing novel metaphors would have been difficult.

Therefore, perhaps the participants of Bowdle and Gentner simply found categorical processing less easy for the novel than the conventional metaphors, and stated their preference accordingly. Utsumi (2007, 2011) suggested that people process metaphors which are high in interpretive diversity categorically, and metaphors which are low in interpretive diversity comparatively. Interpretive diversity might be another factor which eases or makes harder the task, causing metaphors with high interpretive diversity to be considered "easier to process", and therefore to be processed categorically. Glucksberg and Haught (2006) found that aptness determines the manner in which metaphors are processed, with literal-referent metaphors being processed categorically, and figurative-referent metaphors being processed comparatively. However, figurative-referent metaphors are intuitively harder to process than literal-referent metaphors. Therefore, this finding may have also risen due to the inherent relative ease/difficulty in processing the two metaphor types, rather than because of the style of their referents. Overall, it is possible that past metaphor findings, as well as current metaphor findings, can be explained by the relative ease experienced in processing a metaphor. Perhaps, in general, feelings of ease lead toward categorical processing, and feelings of difficulty lead towards comparative processing.

Dual Processing

Results suggest that experimental context influences the meaning construction process by affecting which information processing strategy is used to process novel word-pairs. The current study found evidence suggesting that when task demand was low (as in the conventional condition), participants rely on an effortless associative processing strategy. This strategy is suggested to be "linguistic" and based on statistical heuristics such as whether or not words share distributional history (Lynott & Connell, 2010). Participants in the current study who were predicted to use this strategy found few novel word-pairs meaningful, which is fitting because novel word-pairs should rarely share distributional history. When these participants did find novel word-pairs "meaningful", their neural responses to these were morphologically similar to those to the conventional metaphoric word-pairs which they found "meaningful". This suggests that similar processing occurs across word-pair type within the conventional condition, which again supports the hypothesis that these participants rely on the associative strategy

The current study also found evidence that when task demand was high (as in the novel50% and novel33% condition) participants relied on a more effortful, analytical processing strategy (the "simulation" strategy according to Lynott & Connell, 2010). Participants in the current study who were predicted to use this strategy had a relatively high number of "meaningful" responses to target novel word-pairs. This is fitting because these participants were predicted to simulate and attempt to mesh neural activations related to the different words in each pair, and therefore to find many novel word-pairs meaningful because adjectives and nouns have many "affordances" ("ways in which a concept offers opportunities for meshing with other concepts", Lynott & Connell, p. 5). Consistent with analytical processing, participants in the novel33% condition had more negative N400s to target novel word-pairs than participants in the conventional condition, signifying increased processing effort. These participants also had neural responses to target novel word-pairs which were qualitatively distinctive from those of participants in the conventional condition to the same word-pairs. This provides further evidence for dual process predictions, as it indicates that participants in the two conditions process target novel word-pairs in distinct ways.

These results bolster the predictions of the ECCo model that task demand will influence the manner in which people process language (Lynott & Connell, 2010, p.4). Other dual-processing theories also predict that similar types of dual processes exist, (e.g. James, 1890; Kahneman, 2003; Sloman, 1996), but these do not predict how task demand will influence which strategy is relied upon, and so ECCo proves to be a stronger model in this respect.

A follow-up study to further test the ECCo model, and to explore the consequences of analytical processing, is proposed. In this study, participants should complete the meaningfulness decision task of experiment one of the current study, and subsequently complete a recognition task. Stimuli in the recognition task should include the novel word-pairs from the current study, and novel re-combinations of these, as well as additional novel word-pairs. If, as predicted, participants in the novel33% condition analytically process novel word-pairs then they should often mesh the two concepts successfully. In accordance with ECCo, participants who process novel word-pairs analytically should therefore be less susceptible to combination errors (errors in which novel re-combinations of old words are mistakenly called old pairs) than participants who process the novel word-pairs associatively. This is because participants who process novel word-pairs analytically should form a new concept in their semantic memories and therefore remember this combined concept rather than the individual word meanings. If it is found that participants in the novel33% and novel50% conditions make fewer combination errors than participants in the conventional and literal conditions, then this would provide converging evidence that different task demands driven by different experimental context can cause participants to use different processing strategies to assess novel word-pairs. In addition, Lynott and Connell (2010) describe analytical processing as deeper than associative processing. Craik and Lockhart (1972) showed that the more deeply something is processed, the more likely it is to be remembered. Therefore, another prediction of the proposed experiment would be that if participants in the novel33% condition are processing target novel word-pairs more analytically than participants in the conventional condition, then they should successfully recognise old word-pairs better. The suggested findings would indicate that experimental context can influence memory.

Fluency

The current results are also consistent with the fluency literature. Fluent processing has been shown to elicit associative strategy reliance, and disfluent processing has been shown to elicit analytical strategy reliance (Alter et al., 2007). The fluency literature is advanced compared to ECCo and other language-related dual processing models because it has already shown some specific aspects of experimental context (such as reading ease and experimenter competency, see Alter et al., 2007) can influence which processing strategy is relied upon. The current study adds to this by demonstrating that “word-pair type” is another aspect of experimental context which influences fluency, and therefore influences information processing strategy reliance. Future examination into other aspects of experimental context, such as proportions and blocking of different word-pair types, would further elucidate the influences of it on dual processing reliance, and therefore on the meaning construction process.

The current study assumes that the task of judging a novel word-pair as meaningful in the novel33% condition is more difficult, or “disfluent”, than the task of judging a novel word-pair as meaningful in the conventional condition. While this makes logical sense, there is no empirical evidence supporting this. In order to provide empirical evidence that the task is more disfluent in the novel33% than in the conventional condition, a follow-up study is proposed. Participants should do the same task as in experiment one and two of the current study, however, once completed they should respond to a series of "Moses Illusion"- like questions. The Moses Illusion was first explored by Erickson and Mattson (1981), and is the finding that when asked questions such as "*How many animals of each kind did Moses take on the Ark?*" participants tend to answer "*two*", despite knowing that Noah rather than Moses took the animals on the arc. This effect has been shown to be augmented by fluent conditions (Song & Schwarz, 2008). Therefore, if the conventional condition is more fluent than the novel33% condition, then subsequent to task completion, participants in the conventional condition should be more prone to Moses Illusion-like errors than participants in the novel33% condition.

Other Effects of Experimental Context

As well as dual processing influences, two other factors of experimental context which might influence the meaning construction process were considered in the current thesis: bias and process priming. No evidence was found for either. In experiments one and two, similar word-pairs were judged as "meaningful" across conditions suggesting that participants did not randomly respond to meet the implied number of "meaningful" responses. In experiment two, no difference in stimulus-locked LRP was found between conditions indicating that participants did not differ in response preparation before seeing the second word of each pair, and therefore that they did not shift their criteria for meaningfulness based on experimental context. However, participants responded via button presses, and it is possible that this did not initiate strong enough contralateral brain activity to reveal bias. In future this should be examined to determine whether the measure used was sensitive enough.

Participants in the conventional and literal conditions of experiment one did not differ in their proportion of "meaningful" response to novel word-pairs, and therefore the presence of conventional metaphors did not uniquely facilitate figurative processing for novel (loosely figurative) word-pairs. As previously mentioned, conventional metaphors are proclaimed by some researchers to be processed in a different manner to novel metaphors, (although the current study shows that sometimes they can be processed in a similar manner). Therefore, it is possible that conventional metaphors are not a relevant prime for the processing of novel metaphors.

Effect Strength

Although the results of experiment two showed the replication of the effect of condition found in experiment one, the overall proportion of “meaningful” responses appeared to differ between experiments. To test this statistically, two subject analyses were conducted with experiment as a between subjects factor. Firstly, a 2 condition (conventional/novel33%) by 2 experiment (experiment one/ experiment two) ANOVA was calculated for the proportion of meaningful responses to target novel word-pairs. Main effects of condition, $F(1,133) = 17.60$, $p = .001$, $\eta^2_p = .117$, and experiment, $F(1, 133) = 7.45$, $p = .007$, $\eta^2_p = .053$, were found. As behavioural results of the two experiments showed, participants in the novel33% conditions of both experiments had a higher proportion of meaningful responses to novel word-pairs (*M* = .32, *SD* = .19) than participants in the conventional conditions (*M* = .20, *SD* = .19). Interestingly, participants in experiment two had a higher proportion of meaningful responses to novel word-pairs (*M* = .30, *SD* = .02) than did participants in experiment one (*M* = .22, *SD* = .02). However, no interaction between condition and experiment was found, indicating that effects of condition were robust across the two experiments. It is possible that this difference between experiments arose because participants in experiment two were more invested in the situation, and therefore more likely to “try harder”. Participants in experiment two were paid for their time and spent a while talking with the experimenter beforehand, while participants in experiment one quickly completed the experiment at the end of a two hour laboratory. This increase in personal investment could explain stronger effects in experiment two because it means that participants were more likely to be engaging the same cognitive task. No other effects or interactions were found.

A repeated-measures ANOVA with condition (conventional/novel33%) and experiment (experiment one/experiment two) as between-subject factors, and with response (“meaningful”/“not-meaningful”) as a within-subjects factor was calculated for reaction times to target novel word-pairs. A main effect of response was found that was qualified by a Condition by Experiment by Response interaction, $F(1, 129) = 6.10$, $p = .015$, $\eta^2_p = .045$. Follow-up condition (novel33%/conventional) by response (“meaningful”/”not-meaningful”) ANOVAs conducted for experiment one and experiment two separately, revealed a significant interaction of response time by condition for experiment two, $F(1, 42) = 8.20$, $p = .007$, $\eta^2_p = .163$, but not for experiment one. In experiment two, participants in the novel33% condition were slower than participants in the conventional condition to make “not-meaningful” responses, but participants in the conventional condition were slower than participants in the novel33% condition to make “meaningful” responses (see results section). The same pattern of results was seen in experiment one, although this did not reach significance. Again, this difference between experiments may have occurred due to differential investment in the task.

Participants in experiment one were paid in course credit or chocolate fish, whereas participants in experiment two were paid in movie vouchers, which are financially more valuable. Also, participants in experiment one spent only a few minutes conversing with the experimenter prior to task initiation, while participants in experiment two spent between 30 and 60 minutes during ERP set-up. In addition, participants in experiment one were run in groups, spent only about five to ten minutes on the experiment, and completed 106 trials; while participants in experiment two were run individually, spent about 90 to 120 minutes on the experiment and completed 246 trials. Moreover, different sets of novel word-pairs were used in the two experiments. Therefore, the fact that similar effects of condition were found in both experiments, despite all these differences in personal investment and experience, demonstrates the strength of this effect.

Implications for Experimentation

Many past researchers have compared and contrasted their experiments without considering that experimental context may have influenced their results in different ways. This is not just the case for metaphor studies (which have been the focus of this thesis) but more generally for all studies that incorporate different types of stimuli. Past studies with results that differed or were comparable may have obtained these results due to influences of experimental context and yet have attributed them to something else. For example, Arzouan et al. (2007) compare their results that metaphor familiarity affects negativity of the N400 to the results of Pynte et al. (1996) which suggest the opposite. However, Arzouan et al. presented their novel metaphors mixed with conventional metaphoric, semantically related, and semantically unrelated phrases, while Pynte et al. presented their novel metaphors mixed with conventional metaphoric and semantically related phrases. The fact that Pynte et al. did not present semantically unrelated phrases, while Arzouan et al. did, means that their novel metaphors were presented in different experimental contexts. However Arzouan et al. did not consider this as potential reason why results differed between the two experiments. The results of the current study show that reconsideration of past studies is imperative. The current results also suggest that future researchers need to be considerate of experimental context when designing and interpreting their experiments, so as not to confound experimental manipulations with effects of experimental context.

Conclusions

The current study found evidence that two types of processing can occur when people make successful meaningfulness judgements, and that which one a person uses depends on experimental context. One of these processes was found to lead to more “meaningful” judgements than the other, indicating that experimental context can determine the likelihood of a person finding a novel word-pair meaningful. This process is consistent with ECCo’s simulation process which constructs new meanings for given phrases. Therefore, the current study suggests that experimental context can influence whether or not the meaning construction process is utilized. In terms of metaphor, the current study showed two processes with which a novel metaphor can be assessed, that were influenced by experimental context. Whether or not these processes correspond to the categorization and comparison processes discussed in contemporary theories of metaphor should be further examined. Bias and process priming were not found to affect results in the current study, although they cannot be ruled out as influential aspects of context because they might influence the meaning construction process under different circumstances.

References

Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology*, *136*(4), 569-576. doi: 10.1037/0096-3445.136.4.569

Aristotle, (1932). Rhetoric. In L. Cooper (Trans.), *The rhetoric of Aristotle.* New York: Appleton-Century-Crofts.

Arzouan, Y., Goldstein, A., & Faust, M. (2007). Brainwaves are stethoscopes: ERP correlates of novel metaphor comprehension. *Brain Research*, *1160*, 69-81. doi: 10.1016/j.brainres.2007.05.034

Barrett, S., & Rugg, M. (1990). Event-related potentials and the semantic matching of pictures. *Brain and Cognition*, *14*, 201-212. doi: 10.1016/0278-2626(90)90029-N

Bentin, S., McCarthy, G., & Wood, C.C. (1985). Event-related potentials, lexical decision and semantic priming. *Clinical Neurophysiology*, *6*, 343-355. doi: doi:10.1016/0013-4694(85)90008-2

Besson, M., & Macar, F. (1987). An event-related potential analysis of incongruity in music and other non-linguistic contexts. *Psychophysiology*, *24*, 14-25. doi: 10.1111/j.1469-8986.1987.tb01853.x

Blasko, D.G., & Briihl, D.S. (1997). Reading and recall of metaphorical sentences: Effects of familiarity and context. *Metaphor & Symbol*, *12*, 261-285. doi: 10.1207/s15327868ms1204_4

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on “semantic P600” effects in language comprehension. *Brain Research Reviews*, 55-73. doi: 10.1016/j.brainresrev.2008.05.003

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

Bradley, M.M., & Lang, P. J. (1999). Affective norms for English words (ANEW): Stimuli, instruction manual and affective ratings. (Technical report C-1), Gainesville, Florida: The University of Florida Center for Research in Psychophysiology.

Chaiken, S., Giner-Sorolla, R., & Chen, S. (1996). *Defence and impression motives in heuristic and systematic information processing*. New York: Guilford Press.

Clark, H. (1973). Space, time, semantics, and the child. In T. Moore. (Ed.), *Cognitive development and the acquisition of language* (pp. 27–63). New York: Academic Press.

Cornejo, C. (2008). Intersubjectivity as co-phenomenology: From the holism of meaning to the being-in-the-world-with-others. *Integrative Psychological and Behavioral Science, 42*(2), 171–178. doi: 10.1007/s12124-007-9043-6

Coulson, S., & Van Petten, C. (2002). Conceptual integration and metaphor: an event-related potential study. *Memory and Cognition*, *30*(6), 958-968.

Craik, F., & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behaviour, 11*, 671-684. doi: 10.1016/S0022-5371(72)80001-X

D'Argembeau, A., & Mathy, A. (2011). Tracking the construction of episodic future thoughts. *Journal of Experimental Psychology*, *140*(2), 258-271. doi: 10.1037/a0022581

DeGrauwe, S., Swain, A., Holcomb, P., Ditman, T., & Kuperberg, G. (2010). Electrophysiological insights into the processing of nominal metaphors. *Neuropsychologia*, *48*, 1965–1984. doi: 0.1016/j.neuropsychologia.2010.03.017

DeLong, K. (2010). Electrophysiological explorations of linguistic pre-activation and its consequences during online sentence processing. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, *70*(7), 4504.

Diehl, V., & Reese, D. (2010). Elaborated metaphors support viable inferences about difficult science concepts. *Educational Psychology, 30*(7), 771-791. doi: 10.1080/01443410.2010.504996

Erickson, T., & Mattson, M. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behavior, 20*, 540-551. doi: 10.1016/S0022-5371(81)90165-1

Faust, M., & Mashal, N. (2007). The role of the right cerebral hemisphere in processing novel metaphoric expressions taken from poetry: A divided visual field study. *Neuropsychologia*, *45*(4), 860-870. doi: 10.1162/jocn.2008.20005

Federmeier, K., Van Petten, C., Schwartz, T., & Kutas, M. (2003). Sounds, words, sentences: Age-related changes across levels of language processing. *Psychology and Aging*, *18*(4), 858-872. doi: 10.1037/0882-7974.18.4.858

Federmeier, K., Wlotka, E., De Ochoa, E., & Kutas, M. (2007). *Brain research, 1146*, 75-84. doi: 10.1016/j.brainres.2006.06.101

Frenzel, S., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2011). Conflicts in language processing: A new perspective on the N400–P600 distinction. *Neuropsychologia, 49*, 574–579. doi: 10.1016/j.neuropsychologia.2010.12.003

Friederici, A., Steinhauser, K., & Frisch, S. (1999). Lexical integration: Sequential effects of syntactic and semantic information. *Memory & Cognition*, *27*(3), 438-453. doi: 10.1016/j.neuropsychologia.2010.12.003

Fritjof, C. (1988). *Uncommon wisdom: Conversations with remarkable people.* New York: Bantam Books.

Gentner, D., & Imai, M. (1992). Is the future always ahead? Evidence for system-mappings in understanding space-time metaphors. *Proceedings of the Fourteenth Annual Meeting of the Cognitive Science Society*, 510–515.

Gentner, D., & Wolff, P. (1997). Alignment in the processing of metaphor. *Journal of Memory and Language, 37*, 331-355. doi: 10.1006/jmla.1997.2527

Gibbons, R.D., Hedeker, D.R., & Davis, J.M. (1993). Estimation of effect size from a series of experiments involving paired comparisons. *Journal of Educational Statistics*, *18*(3), 271-279. doi: 10.2307/1165136

Glucksberg, S., Gildea, P., & Bookin, H. (1981). On understanding nonliteral speech: can people ignore metaphors? *Journal of Verbal Learning and Verbal Behavior*, *21*(10), 85-98. doi: 10.1016/S0022-5371(82)90467-4

Glucksberg, S., & Haught, C. (2006). Can Florida become like the next Florida? *Association for Psychological Science*, *17*(11), 935- 938. doi: 10.1111/j.1467-9280.2006.01807.x

Glucksberg, S., & Keysar, B. (1990). Understanding metaphoric comparisons: Beyond similarity. *Psychological Review*, *97*(1), 3-18. doi: 10.1037/0033-295X.97.1.3

Goodman, N. (1972). Seven strictures on similarity. In N. Goodman (Ed.), *Problems and Projects* (pp. 437-447). New York: Bobbs-Merrill.

Graesser, A., Long, D., & Mio, J. (1989). What are the cognitive and conceptual components of humorous texts? *Poetics, 18,* 143-164. doi: 10.1016/0304-422X(89)90026-0

Gratton, G., Coles, M., Sirevaag, E., Eriksen, C., & Donchin, E. (1988). Pre- and post-stimulus activation of response channels: a psychopsyhsiological analysis. *Journal of Experimental Psychology, 14*(3), 331-344. doi: 10.1037/0096-1523.14.3.331

Green, D. M., Swets J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Grice, H. (1975). Logic and conversation. In P. Cole., & J. Morgan (Eds.), *Syntax and semantics 3: Pragmatics* (pp. 41-58). Academic Press, New York

Harris, R. (1976). Comprehension of metaphors: A test of the two-stage processing model. *Bulletin of Psychonomic Society, 8*, 312–314.

Hedges, L.V. (1981). Distribution theory for Grass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*, 107-128. doi:10.2307/1164588

Hedges, L.V. (1982). Estimation of effect sizes from a series of independent experiments. *Psychological Bulletin, 92*, 490-499. doi:10.1037/0033-2909.92.2.490

Hoeks, J., Stowe, L., & Doedens, G. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading. *Cognitive Brain Research, 19*, 59–73. doi: 10.1016/j.cogbrainres.2003.10.022

Iakimova, G., Passerieux, C., Laurent, J., & Hardy-Bayle, M. (2005). ERPs of metaphoric, literal, and incongruous semantic processing in schizophrenia. *Psychophysiology, 42*(4), 380-390. doi: 10.1111/j.1469-8986.2005.00303.x

Ibanez, A., Toro, P., Cornejo, C., Hurquina, H., Manes, F., Weisbrod, M., & Schroder, J. (2011). High contextual sensitivity of metaphorical expressions and gesture blending: A video event-related potential design. *Psychiatry Research: Neuroimaging, 191*, 68–75. doi: 10.1016/j.pscychresns.2010.08.008

Inhoff, A., Lima, S., & Carroll, P. (1984). Contextual effects on metaphor comprehension in reading. *Memory and Cognition, 12*(6), 558-567. doi: 10.3758/BF03213344

James, W. (1950). *The principles of psychology, volume two.* New York: Dover Publications Incorporated (Original work published 1890).

Kahneman, D. (2003). A perspective on judgement and choice: Mapping bounded rationality. *American Psychologist*, *58*(9), 697-720. doi: 10.1037/0003-066X.58.9.697

Kazmerski, V., Blasko, D., & Dassalegn, B. (2003). ERP and behavioral evidence of individual differences in metaphor comprehension. *Memory and Cognition*, *31*(5), 673-689.

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, *52*, 205-225. doi: 10.1016/j.jml.2004.10.002

Kovecses, Z. (1988). *The language of love: the semantics of passion in conversational English.* Cranbury, United States: Bucknell University Press.

Kuperberg, G.R. (2007). Neural mechanisms of language comprehension: challenges to syntax. *Brain Research*, *1146*(1), 23–49. doi: 10.1016/j.brainres.2006.12.063

Kutas, M., & Donchin, E. (1974). Studies of squeezing: The effect of handedness, the responding hand and response force on the contralateral dominance of the readiness potential. *Science*, *186*, 545- 548. doi: 10.1126/science.186.4163.545

Kutas, M., & Donchin, E. (1977). The effects of handedness, of responding hand, and of response force on the contralateral dominance of the readiness potential. In J. Desmedt (Ed.), *Attention, voluntary contraction and event-related cerebral potentials* (189- 210). Basel: Karger.

Kutas, M., & Donchin, E. (1980). Preparation to respond as manifested by movement-related brain potential. *Brain Research*, *202*, 95-115. doi: 10.1016/S0006-8993(80)80037-0

Kutas, M., & Federmeier, K.D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463-47. . doi: 10.1016/S1364-6613(00)01560-6

Kutas, M., & Federmeier, K.D. (2011). Thirty years and counting: finding meaning in the N400 component of the Event-Related-Brain Potential (ERP). *Annual Review of Psychology*, *62*, 621-647. doi: 10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S., A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*, 203–205. doi: 10.1126/science.7350657

Kutas, M., & Van Petten, C. (1994). Psycholinguistics electrified: Event related bran potential investigations. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 83-143). San Diego: Academic Press.

Kutas, M., Van Petten, C., & Kluender, R. (2006). Psycholinguistics electrified II: 1994–2005. In M. Traxler & M. Gernsbacher (Eds.), *Handbook of psycholinguistics,* (2nd ed., pp. 659–724). New York: Elsevier.

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

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* Chicago: University of Chicago.

Lauwereyns, J. (2010). *The anatomy of bias. How neural circuits weigh the options.* Cambridge: MIT Press.

Luck, S. (2005). *An introduction to the event-related potential technique.* Cambridge: Massachusetts Institute of Technology Press.

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

Masaki, H., Wild-Wall, N., Sangals, J., & Sommer, W. (2004). The functional locus of the lateralized readiness potential. *Psychophysiology*, *41*, 220-230. doi: 10.1111/j.1469-8986.2004.00150.x

Mashal, N., & Faust, M. (2008). Right hemisphere sensitivity to novel metaphoric relations: Application of the signal detection theory. *Brain and Language*, *104*(2), 103-112. doi: 10.1016/j.bandl.2007.02.005

McElree, B., & Nordlie, J. (1999). Literal and figurative interpretations are computed in equal time, *Psychonomic Bulletin & Review, 6*(3), 486–494. doi: 10.3758/BF03210839

McGlone, M., & Manfredi, D. (2001). Topic-vehicle interaction in metaphor comprehension. *Memory and Cognition*, *29*(8), 1209-1219.

Nunez, P.L. (1981). Electric Fields of the Brain: *The Neurophysics of EEG.* New York: Oxford University Press.

Oppenheimer, D. (2008). The secret life of fluency. *Trends in Cognitive Science, 12*(6), 237-241. doi:10.1016/j.tics.2008.02.014

Ortony, A., Schallert, D.L., Reynolds, R.E., & Antos, S.J. (1978). Interpreting metaphors and idioms: Some effects of context upon comprehension. *Journal of Verbal Learning and Verbal Behavior, 17*, 465-477. doi: 10.1016/S0022-5371(78)90283-9

Osman, A., & Moore, C. (1993). The locus of dual-task interference: Psychological refractory effects on movement-related brain potentials. *Journal of Experimental Psychology: Human Perception and Performance*, 19(6), 1292-1312. doi:10.1037/0096-1523.19.6.1292

Osterhout, L., & Holcomb, P. (1992). Event-Related Brain Potentials elicited by semantic anomaly. *Journal of Memory and Language*, *31*, 785-806. doi: 10.1016/0749-596X(92)90039-Z

Osterhout, L., & Holcomb, P. (1993). Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes*, *8*(4), 413. doi: 10.1080/01690969308407584

Petty, R.E., & Cacioppo, J.T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*, (vol. 19, pp. 123-205) New York: Academic press.

Pynte, J., Besson, M., Robichon, F., & Poli, J. (1996). The time-course of metaphor comprehension: An event-related potential study. *Brain and Language*, *55*, 293-316. doi: 10.1006/brln.1996.0107

Quinn, N. (1987). Convergent evidence for a cultural model of American marriage. In D. Holland & N. Quinn (Eds.), *Cultural models in language and thought* (pp. 173-192). New York: Cambridge University Press.

Radden, G., Kopcke, K., Berg, T., & Siemund, P. (2007). *Aspects of meaning construction*. Philadelphia: John Benjamins North America.

Reddi, B.A., & Carpenter, R.H. (2000). The influence of urgency on decision time. *Nature Neuroscience. 3*, 827–830. doi: 10.1038/77739

Regel, S., Gunter, T., & Friederici, A. (2010). Isn't it ironic? An electrophysiological exploration of figurative language processing. *Journal of Cognitive Neuroscience*, *23*(2), 277–293. doi: 10.1162/jocn.2010.21411

Rhodes, G., Parkin, A.J., & Tremewan, T. (1993). Semantic priming and sensitivity in lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 154-165. doi: 10.1037/0096-1523.19.1.154

Richards, I. (1936). *The philosophy of rhetoric.* Oxford: Oxford university press.

Risko, E., Blais, C., Stolz, J., & Besner, D. (2008). Nonstrategic contributions to putatively strategic effects in selective attention tasks. *Journal of Experimental Psychology*, *34*(4), 1044-1052. doi: 10.1037/0096-1523.34.4.1044

Rugg, M., & Coles, M. (1995). *Electrophysiology of mind: Event-related brain potentials and cognition*. Oxford: Oxford University Press.

Sakamoto, T., Matsuishi, K., Arao, H., & Oda, J. (2003). An ERP study of sensory mismatch expressions in Japanese. *Brain and Language*, *86*, 384–394. doi: 10.1016/S0093-934X(02)00539-4

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime Suite (Version 1.1). Pittsburgh: Psychology Software Tools Inc.

Searle, J.R. (1979). *Expression and meaning: Studies in the theory of speech acts*. Cambridge: Cambridge University Press.

Sloman, S. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*(1), 3-22. doi: 10.1037/0033-2909.119.1.3

Song, H., & Schwarz, N. (2008). Fluency and detection of misleading questions: Low processing fluency attenuates the Moses illusion. *Social Cognition*, *26*(6), 791–799. doi: 10.1521/soco.2008.26.6.791

Steinhauser, M., Hubner, R., & Druey, M. (2009). Adaptive control of response preparedness in task switching. *Neuropsychologia*, *47*(8-9), 1826-1835. doi: 10.1016/j.neuropsychologia.2009.02.022

Tartter, V., Gomes, H., Dubrovsky, B., Molholm, S., & Stewart, R. (2002). Novel metaphors appear anomalous at least momentarily: evidence from N400. *Brain and Language*, *80*, 488-509. doi: 10.1006/brln.2001.2610

Tetlock, P.E., & Lerner, J.S. (1999). The social contingency model: Identifying empirical and normative boundary conditions for the error and- bias portrait of human nature. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 571–585). New York: Guilford Press.

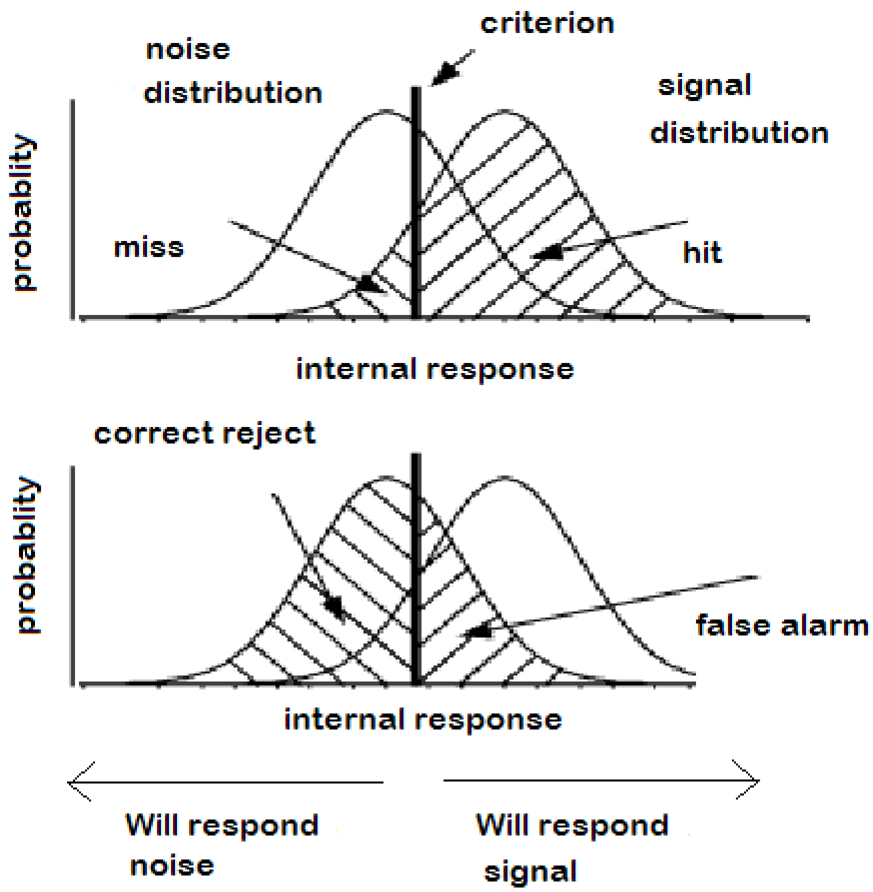
Utsumi, A. (2007). Interpretive diversity explains metaphor-simile distinction. *Metaphor and Symbol*, *22*(4), 291-312.

Utsumi, A. (2011). Computational exploration of metaphor comprehension processes using a semantic space model. *Cognitive Science*, *35*, 251-296. doi: 10.1111/j.1551-6709.2010.01144.x

Yang, F., Fuller, J., Khodaparast, N., & Krawczyk, D. (2010). Figurative language processing after traumatic brain injury in adults: A preliminary study. *Neuropsychologia*, *48*, 1923–1929. doi: 10.1016/j.neuropsychologia.2010.03.011

Appendix A

Criterion in Signal Detection Theory



Appendix B

The Screening Questionnaire of Experiment One

On a scale of 1 to 5 (with 1 meaning "not related at all", and 5 meaning "very related"), how related do you think each of the following word pairs are?

	Not related at all				Very related
sheer drain	1	2	3	4	5
fleet stole	1	2	3	4	5
flat chuck	1	2	3	4	5
neat gloom	1	2	3	4	5
void tense	1	2	3	4	5
plain brand	1	2	3	4	5
blind spark	1	2	3	4	5
tense blast	1	2	3	4	5
frank print	1	2	3	4	5
sweet quote	1	2	3	4	5
meek tract	1	2	3	4	5
salt creek	1	2	3	4	5
worst blade	1	2	3	4	5
deaf waist	1	2	3	4	5
dear spade	1	2	3	4	5
swiss gauge	1	2	3	4	5
wise cliff	1	2	3	4	5

plus curse	1	2	3	4	5
blank boots	1	2	3	4	5
cross straw	1	2	3	4	5
stray crown	1	2	3	4	5
moist graph	1	2	3	4	5
vain flush	1	2	3	4	5
faint realm	1	2	3	4	5
chill crash	1	2	3	4	5
vague drift	1	2	3	4	5
grave juice	1	2	3	4	5
strict purse	1	2	3	4	5
stern slate	1	2	3	4	5
dull ridge	1	2	3	4	5
mild stove	1	2	3	4	5
safe paste	1	2	3	4	5
tight flood	1	2	3	4	5
trim brass	1	2	3	4	5
arch stream	1	2	3	4	5
warm nerve	1	2	3	4	5
worse swear	1	2	3	4	5
fond steep	1	2	3	4	5
prompt couch	1	2	3	4	5
weird grove	1	2	3	4	5

null prose	1	2	3	4	5
soft squad	1	2	3	4	5
cool forge	1	2	3	4	5
grand nurse	1	2	3	4	5
flash knock	1	2	3	4	5
swift sweep	1	2	3	4	5
false slide	1	2	3	4	5
slow toast	1	2	3	4	5
slight slope	1	2	3	4	5
huge garth	1	2	3	4	5
weak trick	1	2	3	4	5
lean chase	1	2	3	4	5
crude sauce	1	2	3	4	5
blond flock	1	2	3	4	5
nude stake	1	2	3	4	5
limp breed	1	2	3	4	5
greek stray	1	2	3	4	5
halt layer	1	2	3	4	5
glad fleet	1	2	3	4	5
vast roast	1	2	3	4	5
joint stare	1	2	3	4	5
brief clock	1	2	3	4	5

bleak brick	1	2	3	4	5
quit grasp	1	2	3	4	5
proud mouse	1	2	3	4	5
damp bunch	1	2	3	4	5
clean swore	1	2	3	4	5
brave grill	1	2	3	4	5
rank boost	1	2	3	4	5
drunk blank	1	2	3	4	5
keen slice	1	2	3	4	5
thick dodge	1	2	3	4	5
prone bloom	1	2	3	4	5
dumb theft	1	2	3	4	5
waste globe	1	2	3	4	5
quaint midst	1	2	3	4	5
sick saint	1	2	3	4	5
course cheek	1	2	3	4	5
bare shaft	1	2	3	4	5
steep spray	1	2	3	4	5
curt tooth	1	2	3	4	5
tall sheer	1	2	3	4	5
quick chill	1	2	3	4	5
self dough	1	2	3	4	5
prime shake	1	2	3	4	5

smart mound	1	2	3	4	5
stiff crawl	1	2	3	4	5
sharp stall	1	2	3	4	5
tough lodge	1	2	3	4	5
slim flame	1	2	3	4	5
loud troop	1	2	3	4	5
grim ghost	1	2	3	4	5
loose cream	1	2	3	4	5
rare climb	1	2	3	4	5
pale ditch	1	2	3	4	5
pure brace	1	2	3	4	5
sole grief	1	2	3	4	5
smooth shelf	1	2	3	4	5
ripe thumb	1	2	3	4	5
spare crest	1	2	3	4	5

Appendix C

The novel word-pairs used in experiment one, and their mean ratings (out of 5) for meaningfulness

Word-pair	Mean Rating	Word-pair	Mean Rating
sheer drain	1.27	slight slope	3.87
fleet Stole	2.27	huge garth	1.87
flat chuck	1.8	weak trick	2.93
neat gloom	1.4	lean chase	2.2
void tense	1.8	crude sauce	2.27
plain brand	2.67	blond flock	2.13
blind spark	1.87	nude stake	1.73
tense blast	2.47	limp breed	1.93
frank print	1.6	greek stray	1.73
sweet quote	2.4	halt layer	1.2
meek tract	1.13	glad fleet	1.47
salt creek	2.93	vast roast	2.07
worst blade	2.27	joint stare	2.67
deaf waist	1.33	brief clock	1.73
dear spade	1.27	bleak brick	2.07
swiss gauge	2.4	quit grasp	1.2
wise cliff	1.13	proud mouse	2.67
plus curse	1.27	damp bunch	2.6
blank boots	2.07	clean swore	1.8
cross straw	1.47	brave grill	1.27
stray crown	1.67	rank boost	2.33
moist graph	1.47	drunk blank	2.67
vain flush	1.53	keen slice	2
faint realm	1.47	thick dodge	1.4
chill crash	1.93	prone bloom	1.33
vague drift	2.87	dumb theft	3.67
grave juice	1.33	waste globe	2.33
strict purse	1.93	quaint midst	1.47
stern slate	2.13	sick saint	2.73
dull ridge	2.2	course cheek	1.53
mild stove	1.8	bare shaft	2.93
safe paste	2.4	steep spray	2.07
tight flood	1.33	curt tooth	1.2
trim brass	1.87	tall sheer	2.73
arch stream	2.53	quick chill	2.67
warm nerve	2.07	self dough	1.67
worse swear	3.2	prime shake	1.53
fond steep	1.4	smart mound	1.67
prompt couch	1.27	stiff crawl	2.4
weird grove	2.27	sharp stall	1.47
null prose	1.47	tough lodge	1.67

soft squad	1.8	slim flame	2.4
cool forge	2.13	loud troop	3.07
grand nurse	2.53	grim ghost	3.67
flash knock	1.67	loose cream	2.2
swift sweep	3.27	rare climb	3.13
false slide	1.47	pale ditch	1.53
slow toast	3.33	pure brace	1.47
ripe thumb	1.2	sole grief	2.47
spare crest	1.93	smooth shelf	2.8

Appendix D

Conventional metaphoric and literal word-pairs used in experiment one

25 Metaphoric Word-Pairs	25 Literal Word-Pairs
small mind	dark eyes
black plague	cheap flight
sweet heart	commuter train
defence line	tasty sandwich
clear thought	daily news
sour grapes	favourite dress
bad egg	curly hair
painful lesson	unhappy childhood
stubborn stain	normal family
warm reception	exciting film
dark secret	boring job
sharp tongue	buttery pastry
bubbly personality	ancient art
old flame	happy couple
lost cause	pretty girl
visual field	good mood
last resort	great idea
thin air	fresh milk
weak will	folk music
rock star	old shoes
high note	angry bull
raw talent	hot water
lion heart	messy room

Appendix E

Consent form for experiment one

TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI
VICTORIA
UNIVERSITY OF WELLINGTON

Gina Grimshaw, Ph.D.	Senior Lecturer	Gina.grimshaw@vuw.ac.nz
Jessica Stewart	Postgraduate	Jess.stewart@vuw.ac.nz

What is the purpose of this research?
- This research will allow us to examine people's perceptions of meaningfulness.

Who is conducting the research?
- We are a team of researchers in the School of Psychology. Dr. Grimshaw is supervising this project. This research has been approved by the University ethics committee.

What is involved if you agree to participate?
- If you agree to participate in this study you will be shown word pairs, and asked to decide whether they are meaningful or not.
- We anticipate that your total involvement will take no more than 15 minutes.
- During the research, you are free to withdraw, without any penalty, at any point before your data have been collected.

Privacy and Confidentiality
- We will keep your consent forms and data for at least five years after publication.
- You will never be identified in my research project or in any other presentation or publication. The information you provide will be coded by number only.
- In accordance with the requirements of some scientific journals and organizations, your coded data may be shared with other competent researchers.
- Your coded data may be used in other, related studies.
- A copy of the coded data will remain in the custody of Dr Gina Grimshaw.

What happens to the information that you provide?
- The data you provide may be used for one or more of the following purposes:
    - The overall findings may be submitted for publication in a scientific journal, or presented at scientific conferences.

- The overall findings may form part of a PhD thesis, Masters thesis, or Honours research project that will be submitted for assessment.

If you would like to know the results of this study, they will be available by approximately July via email (of course you'll have to provide an email address).

If you have any further questions regarding this study, please contact any one of us above.

Statement of consent

I have read the information about this research and any questions I wanted to ask have been answered to my satisfaction.

I agree to participate in this research. I understand that I can withdraw my consent at any time, without penalty, prior to the end of my participation.

Name: ___________________________________

Signature: ___________________________________

Date: ___________________________________

Student ID: ___________________________________

Email: ___________________________________ (if you would like a copy of the results).

Copy to:
[a] participant,
[b] researcher (initial both copies below)

Appendix F

Debrief form for experiment one

TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI
VICTORIA
UNIVERSITY OF WELLINGTON

Gina Grimshaw, Ph.D.	Senior Lecturer	Gina.grimshaw@vuw.ac.nz
Jessica Stewart	Postgraduate Student	Jess.stewart@vuw.ac.nz

Thank you for participating in this experiment. This study investigated how people perceive meaningfulness. When people see a phrase such as “lion-heart” they can interpret it literally (the actual heart of a lion) or metaphorically (somebody who is inherently brave and noble). When people see phrases in the context of a sentence (e.g. “he had a lion-heart” vs. “The vicious bear began to eat the lion-heart”) the appropriate meaning is usually easy for them to accept, as other meanings do not make sense.

In this task, participants saw loosely associated word-pairs. These word-pairs were not presented in the context of a sentence, and they could only make sense if interpreted metaphorically (e.g. “friendly tree”). While some participants only saw loosely associated word-pairs, others also saw conventional metaphoric word-pairs (e.g. “small mind”) mixed in with the loosely associated word-pairs. Everybody’s task was to decide whether each word-pair was meaningful or not. We were interested in seeing whether the participants who saw conventional metaphors mixed in with their loosely associated word-pairs would interpret the loosely associated word-pairs differently to the participants who only say the loosely associated word-pairs.

More specifically we were interested to see if people who saw literal word-pairs with the loosely associated word-pairs would be biased to think literally, therefore processing the critical word-pairs literally, and thus thinking that they didn’t make sense. We were also interested in seeing if people who saw metaphoric word-pairs with the loosely associated word-pairs would be biased to think metaphorically, therefore processing the critical word-pairs metaphorically, and thus thinking that they did make sense. If we do find that the way in which people process loosely associated word-pairs changes depending on experimental context, we wish to use these results to tease apart whether this occurs due to an actual change in their perception of meaningfulness, or simply due to them changing their criteria.

If we do find that experimental context can affect peoples’ perception of meaningfulness, then this knowledge would be useful in both the realms of poetry and advertising. It would also be useful in the experimental design of other experiments investigating meaningfulness. It would show researchers that they have to be cautious when selecting which types of phrases to include in their experiments.

Thanks again for your help with our study. If you have further questions please contact the lead investigator, Dr. Gina Grimshaw, at gina.grimshaw@vuw.ac.nz. If you would like a summary of our findings at the end of term, please leave your email address with the researcher today.

Appendix G

The novel word-pairs used in experiment two

Novel word-pairs	
fond steep	bare shaft
ocean upset	handed steam
varied rating	upwards assign
rugged walker	secure waiter
crude sauce	swing arrow
lousy temper	warm nerve
blond flock	steep spray
slim flame	mixed ballot
outer client	rare climb
cotton reject	bleak brick
greek stray	poetic ticket
curt tooth	rough elder
yearly carpet	infant decay
ironic statue	giant cavity
alike insert	expert outfit
ardent media	sweet quote
awake magnum	backed cowboy
trim brass	modest sector
crazy chaos	sudden potato
canvas usage	rank boost
gentle serum	absent picnic
nude stake	rapid knight
prone bloom	prime shake
worse swear	unpaid monkey
mighty rally	neat gloom
rotary burial	limp reed
lean chase	quick chill
faint realm	quit grasp
worst blade	closet ration
stern slate	devil debut
weak trick	shock mosque
flat chuck	swift sweep
brave grill	pious hello
unlike walnut	quaint midst
cool forge	sturdy marina
brush abuse	drunk blank
poison heater	prompt couch
harsh quest	colony assets
sacred cubism	proud mouse
stray crown	silver praise
sole grief	grand nurse
joint stare	dear spade
cubic tenure	waking suburb
slow toast	bigger morale
stolen glaze	alien ferry

soft squad	border cease
angry chorus	unwed sleeve
loaded collar	blind spark
damned button	drying folly
honey miser	solar convey
patent violin	damp bunch
finite fabric	latest orbit
dusty array	sharp stall
plus curse	rebel pulley
keen slice	stiff crawl
seldom breast	pale ditch
lever helium	sick saint
proof rescue	ending piazza
glad fleet	spare crest
uneasy anchor	saving foster
naval cigar	stupid fringe
awful forum	cosmic puzzle
false slide	hazard hunter
safe paste	savage saloon
wise cliff	moist graph
acute posse	ruling hunger
mature lemon	intent rider
lonely vocal	chill crash
muddy tackle	earthy hymen
allied toilet	witty resume
swiss gauge	tense blast
titled sport	loose cream
verbal debot	eager needle
arch stream	match patch
grimly razor	blank boots
dirty pollen	coarse bishop
grave juice	census resort
void tense	humble dental
pure brace	undue gossip
clean swore	whiskey rubber
vivid organ	vain flush
rival ghetto	fluid thread
sheer drain	tall sheer
unfair thesis	mutual sunset
blonde bacon	beaten atlas
ritual logic	cappy elite
tight flood	slight slope
smooth shelf	brief clock
worthy lunar	deaf waist
loud troop	course cheek

Appendix H

Conventional metaphors used in experiment two and people's averaged proportion of responses to questions about them.

Word-pair	Proportion of times people responded yes to the following questions:		
	Does this pair make sense?	Is it metaphoric	Is it a familiar word-pair
thin air	1.00	0.57	1.00
busy bee	1.00	1.00	1.00
wet blanket	1.00	1.00	1.00
cold blood	1.00	1.00	1.00
spiritual bond	1.00	0.71	1.00
idiot box	0.86	0.86	0.43
quick buck	1.00	0.43	1.00
happy camper	1.00	1.00	1.00
loose cannon	0.86	1.00	1.00
lost cause	1.00	0.57	1.00
shady character	1.00	1.00	1.00
grim contemplation	0.86	0.57	0.43
light conversation	1.00	0.57	0.86
third degree	1.00	1.00	1.00
deep despair	1.00	0.86	1.00
main dish	1.00	0.29	1.00
top dog	0.86	1.00	1.00
stable economy	1.00	0.29	1.00
bad egg	1.00	1.00	1.00
necessary evil	1.00	0.43	1.00
stormy expression	0.86	1.00	0.43
funny farm	0.86	0.86	0.86
secretive feeling	0.86	0.43	0.71
visual field	1.00	0.43	0.86
corrupt file	0.86	0.29	0.86
old flame	0.86	1.00	1.00
sour grapes	1.00	0.86	1.00
brain child	1.00	0.86	0.86
sweet heart	1.00	1.00	1.00
high hopes	1.00	0.86	1.00
vivid idea	0.86	0.57	0.86
painful lesson	1.00	0.57	1.00
defence line	1.00	0.86	1.00
small mind	1.00	1.00	1.00
bubbly personality	1.00	0.86	1.00
black plague	1.00	1.00	1.00
political position	1.00	0.57	1.00
warm reception	1.00	1.00	1.00
healthy relationship	1.00	0.43	1.00
colourful remark	0.86	1.00	0.71

last resort	1.00	0.43	1.00
dry retort	1.00	1.00	0.43
dark secret	1.00	1.00	0.86
free speech	1.00	0.57	1.00
stubborn stain	0.86	0.57	1.00
rock star	1.00	0.86	1.00
hot stuff	1.00	0.86	0.86
raw talent	1.00	0.71	1.00
clear thought	1.00	0.43	1.00
weak will	1.00	1.00	1.00
cat nap	1.00	0.86	1.00
knuckle sandwich	1.00	1.00	1.00
shotgun wedding	1.00	1.00	1.00
bread winner	1.00	0.71	1.00
sunset years	0.86	0.86	0.57
pass away	0.86	0.71	1.00
green thumb	0.71	0.86	0.86
ground rules	1.00	0.71	1.00
jungle fever	1.00	1.00	1.00

Appendix I

Consent form for experiment two

TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI
VICTORIA
UNIVERSITY OF WELLINGTON

Gina Grimshaw, Ph.D.	Senior Lecturer	Gina.grimshaw@vuw.ac.nz
Jessica Stewart	Postgraduate	Jess.stewart@vuw.ac.nz

What is the purpose of this research?

- This research will allow us to examine people's perceptions of meaningfulness.

Who is conducting the research?

- We are a team of researchers in the School of Psychology. Dr. Grimshaw is supervising this project. This research has been approved by the University ethics committee.

What is involved if you agree to participate?

- If you agree to participate in this study, then you will be fitted with an EEG cap. This cap has electrodes on it, which we will use to measure the electrical activity from your scalp. Don't worry- it is perfectly safe and non-invasive.
- Once you have the EEG cap on, we will insert some conductive gel through the electrodes, so that it sits between them and your scalp.
- You will then be run through a cognitive experiment, whilst your scalp electrical activity is being measured.
- During the research, you are free to withdraw, without any penalty, at any point before your data have been collected.

Privacy and Confidentiality

- We will keep your consent forms and data for at least five years after publication.
- You will never be identified in my research project or in any other presentation or publication. The information you provide will be coded by number only.
- In accordance with the requirements of some scientific journals and organizations, your coded data may be shared with other competent researchers.
- Your coded data may be used in other, related studies.
- A copy of the coded data will remain in the custody of Dr Gina Grimshaw.

What happens to the information that you provide?

- The data you provide may be used for one or more of the following purposes:
  - The overall findings may be submitted for publication in a scientific journal, or presented at scientific conferences.
  - The overall findings may form part of a PhD thesis, Masters thesis, or Honours research project that will be submitted for assessment.

If you would like to know the results of this study, they will be available by approximately July via email (of course you'll have to provide an email address).

If you have any further questions regarding this study, please contact any one of us above.

Statement of consent

I have read the information about this research and any questions I wanted to ask have been answered to my satisfaction.

I agree to participate in this research. I understand that I can withdraw my consent at any time, without penalty, prior to the end of my participation.

Name: ____________________________________

Signature: ____________________________________

Date: ____________________________________

Student ID: ____________________________________

Email: ____________________________________ (if you would like a copy of the results).

Copy to:

[a] participant,

[b] researcher (initial both copies below)

Appendix J

Debrief form for experiment two

TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI
VICTORIA
UNIVERSITY OF WELLINGTON

Gina Grimshaw, Ph.D.	Senior Lecturer	Gina.grimshaw@vuw.ac.nz
Jessica Stewart	Postgraduate Student	Jess.stewart@vuw.ac.nz

Thank you for participating in this experiment. In this experiment, we were interested in examining why people sometimes find meaning in something, that at other times they do not find meaning in. In other words, we are interested in determining what can influence somebody's decision to call something "meaningful".

A study we have already finished showed us that people who are only asked to judge loosely associated word-pairs (like "sweet quote") are more likely to judge these word-pairs as meaningful, than people who see both loosely associated word-pairs and conventional metaphoric word-pairs (like "small mind") are. However, our previous study was unable to determine why these results occurred. This could have been because:

1. People who only saw loosely associated word-pairs actually interpreted them metaphorically and thus saw meaning behind them (whereas the people who also saw conventional metaphoric word-pairs did not). Or
2. All participants felt like they had to say "meaningful" sometimes. Participants who saw both types of word-pairs could simply call the conventional metaphors "meaningful", but participants who only saw loosely associated word-pairs did not have this opportunity, and so had to call some loosely associated word-pairs "meaningful".

The aim of this experiment was to determine which of these two alternates, is the reason that people who see only loosely associated word-pairs, judge them as more meaningful than people who see both loosely associated word-pairs and conventional metaphoric word-pairs judge them.

The reason we measured your brainwaves while you completed this task, is that there is a certain change in brain waves that occurs around 400 milliseconds after seeing words, that is related to how easy it is for you to find them meaningful. We call this the N400 component. The harder it is for you to find meaning in a word-pair, the more negative the N400 becomes. We think that if alternative 1 is correct, then the N400s of the participants who see both types of word-pairs should be more negative than the N400s of participants who only see loosely associate word-pairs. This is because if alternative one is correct, then the participants who see both types of word-pairs are less likely to report meaning, and this must be because it is harder for them to find meaning (making the N400 more negative).

However, it alternative two is correct, then there should be no difference in the N400s of both groups of participants. This is because if alternative two is correct, then participants who only see loosely associated word-pairs, are no more likely to *find* them meaningful, but are more likely to *call* them meaningful (because they have to say something is meaningful, and they have nothing else they could call meaningful). Therefore, while participants who only see loosely associated word-pairs are *reporting* them to be more meaningful, they are not actually *finding* them to be so, and so their N400s should be just as negative as those of the participants who see both types of word-pairs.

The reason that we find all of this so interesting is that researchers in cognitive psychology, and in linguistics, often mix up the types of word-pairs that they present to participants in an experiment. However, as our first experiment showed, the type of word-pairs included in an experiment can influence a participant's response. Therefore, future researchers need to be aware of this during experimental design. In this experiment, we wanted to determine *why* the type of word-pairs included in an experiment can influence a participant's response, as this would be useful to future researchers by showing them what to avoid and why, when designing future experiments.

Thanks again for your help with our study. If you have further questions please contact the lead investigator, Dr. Gina Grimshaw, at gina.grimshaw@vuw.ac.nz. If you would like a summary of our findings at the end of term, please leave your email address with the researcher today.