

Processing metaphor: The role of conventionality, familiarity and dominance

Giulio Dulcinati¹, Diana Mazzarella¹, Nausicaa Pouscoulous¹ and Jennifer Rodd²

Abstract:

According to the ‘career of metaphor’ hypothesis (Bowdle & Gentner, 2005) metaphor interpretation shifts from a comparison process to a categorization process when the metaphor vehicle becomes conventional, and little importance is given to the conventionality of the topic-vehicle pairing (i.e., metaphor familiarity). In this study, we used a form-preference paradigm to investigate whether highly familiar and conventional metaphors are preferred in the categorization form. We manipulated conventionality and familiarity and we controlled for aptness. Results did not show clear effects of conventionality or familiarity while they confirmed previous findings that aptness is an important predictor of form preference. An additional post-hoc analysis on the role of dominance (i.e., relative frequency) of metaphorical versus literal uses of the vehicle in the form-preference task suggests that this factor plays a key role on this task.

Keywords: metaphor, language processing, conventionality, familiarity, metaphor dominance

1 Introduction

1.1 Comparison or Categorization?

The question of the cognitive underpinnings of metaphor understanding has been lively debated in the recent psycholinguistics literature (for a review see Glucksberg (2001, 2003)). Most accounts ascribe to one of the following views: (i) metaphors are comprehended as categorizations, (ii) metaphors are comprehended as comparisons.

Let us consider the following nominal metaphor³:

- (1) My lawyer is a shark.

Categorization models maintain that metaphors are understood by attributing the topic of the metaphor (i.e., *lawyer*) to the category denoted by the vehicle (i.e., *shark*). Metaphor interpretation involves the construction of a ‘metaphorical category’ for the vehicle term that differs from its lexically encoded concept. This metaphorical category denotes elements which share some relevant features with (a subset of) items that fall under the concept lexically encoded by the vehicle. For example the metaphorical category SHARK*⁴ may include things that are predatory and vicious (e.g., some sharks but also some lawyers). The class-inclusion account (Glucksberg & Keysar, 1990; Glucksberg & Haught, 2006) and the

¹ Research Department of Linguistics, UCL, 2 Wakefield Street, London WC1N 1PF

² Research Department of Experimental Psychology, UCL, 26 Bedford Way, London, WC1H 0AP, UK

³ Nominal metaphors are metaphors of the form ‘X is a Y’ where X is the topic and Y is the vehicle. There is no theoretically significant reason to distinguish nominal metaphors from other forms. In this paper we will only discuss this form of metaphor because the experiment described here and most of the other relevant studies discussed focus on this form.

⁴ We will follow the practise of representing metaphorical categories as starred concepts.

relevance-theoretic account (Wilson & Carston, 2007; Sperber & Wilson, 2008) can be considered representative for the categorization view of metaphor.

Comparison models maintain that metaphors are ‘implicit similes’ and they are understood by comparing the concepts encoded by the topic and the vehicle. The metaphor expressed by (1) is understood by comparing the lexically encoded concepts MY LAWYER and SHARK. An example of the comparison view is the *structure alignment* model proposed by Gentner and colleagues (Gentner & Clement, 1988; Gentner & Wolff, 1997), which is an application of the *structure-mapping theory* of analogy (Gentner, 1983). This model claims that metaphors (and similes) are understood by means of alignment and inference projection. First, the properties and the structural relations of the concepts of topic and vehicle are aligned. The structural alignment means that non-identical properties can be aligned if they play parallel roles in the two concepts (e.g., sharks and lawyers are predatory in different ways). Then, all the properties of the vehicle that are connected to the aligned structure are projected to the topic.

1.2 Comparison and Categorization

Bowdle and Gentner (2005) proposed a hybrid account of metaphor comprehension that combines the categorization and the comparison views: the career of metaphor hypothesis. This account claims that when a metaphor is novel it is understood as a comparison, when it becomes conventional it is understood as a categorization. The comparison-based process is carried out between the representation at the same level of abstraction (e.g., LAWYER and SHARK). The categorization-based process maps representations at different levels of abstraction (e.g., LAWYER and SHARK*, where the metaphorical category SHARK* is more abstract than LAWYER). Thus, conventionalisation is thought of resulting in a shift in mode of processing from comparison to categorisation.

Crucially, according to this account, whether a metaphor is novel or conventional is determined solely by the vehicle of the metaphor. A conventional metaphor is a metaphor with a conventional vehicle and, conversely, a novel metaphor is a metaphor with a novel vehicle. For example, “That lawyer is a shark” is as conventional as “That employer is a shark”, as long as the vehicle maintains the same metaphorical interpretation.

According to Bowdle and Gentner’s (2005) definition, a novel vehicle is a vehicle that does not evoke a metaphorical category in isolation and that might give way to various interpretations. For instance, consider example (2):

(2) His sister is a megaphone

The term ‘megaphone’ does not evoke a metaphoric category in isolation. Therefore, in order to understand the metaphor we need to compare (i.e., horizontally align) the two lexically encoded concepts, HIS SISTER and MEGAPHONE, and figure out in what way ‘his sister’ could be *like* a ‘megaphone’. Maybe she has a very loud voice or maybe she cannot keep secrets.

When a vehicle has been used many times with the same metaphorical interpretation, the vehicle becomes conventional. The metaphorical sense of a conventional vehicle is stable to the point that the word becomes ‘polysemous’, with the literal and metaphorical meanings being lexically associated to it. That is, both meanings are lexicalized and can be accessed in isolation. For example, consider the metaphor expressed by (1). Even in isolation the vehicle ‘shark’ elicits the metaphoric category of aggressive merciless creatures. While understanding sentence (1), the interpreter directly accesses the metaphoric category SHARK* and attributes the topic to this category. The interpreter vertically aligns the

concept LAWYER with the metaphorical category SHARK*, and all the properties of SHARK* are projected to the topic.

The shift in the mode of processing suggested by the career of metaphor hypothesis does not entail that a metaphor with a conventional vehicle can only be interpreted as a categorization. Indeed while Bowdle and Gentner (2005) claim that while comparison (i.e., horizontal alignment) is the only mode of interpretation available for *novel* metaphors, they suggest that upon encountering a *conventional* metaphor, the processes of comparison and categorization start simultaneously. There are several factors (e.g., context, salience) that may favour one process over the other. However, all things being equal, the categorization process is faster because of its smaller computational cost. Therefore, the competition between different modes of processing allows vehicles that are often used in their ‘conventional’ interpretation to be ‘reinterpreted’ as novel when the comparison-based process is initially favoured. This may happen in cases in which the conventional sense of the vehicle is not suitable for the context in which the metaphor is uttered. For example, a metaphor like “Your cat is a shark” may suggest consideration of similarities between the cat and a literal shark (e.g., biting).

1.3 The Form Preference Task

Bowdle and Gentner (2005) empirically investigated the effect of conventionality on metaphor processing through the form preference task. In the form preference task participants are presented with a list of figurative statements. Each statement is presented both as a nominal metaphor (categorical form) and as a simile (comparison form). For each statement they are asked to indicate the degree to which they prefer one form over the other. For example, participants were shown both the sentences “Faith is like an anchor” and “Faith is an anchor” and they were asked to indicate their preference for one or the other on a ten-point scale.

Bowdle and Gentner (2005) proposed that the form preference task could be used to investigate the processing of metaphor on the assumption that “form reflects function in figurative language” (Bowdle & Gentner, 2005, p. 200). On this assumption, preference for the comparison form should reflect that the figurative statement is processed as a comparison. Conversely, preference for the categorization form suggests that the statement is processed as a categorization.

Bowdle and Gentner (2005) pre-rated the conventionality of a set of stimuli with norming studies and then they measured the rate of preference for the categorical form for each figurative statement. They found that conventional metaphors were preferred in the categorical form (e.g., *This government is a parasite*) significantly more than novel metaphors (e.g., *Her generosity is a light*). These results seem to support the career of metaphor hypothesis.

The significance of their results has been undermined by one particular piece of criticism. Jones and Estes (2006) pointed out that their experiment manipulated conventionality without controlling for other factors that have been found to have an effect on metaphor processing. Jones and Estes argued that conventionality should not be manipulated without controlling for aptness (or vice versa) as these two variables are frequently correlated.

Aptness is a property of metaphor that has been defined in many different ways. The most widely accepted definition was given by Chiappe, Kennedy and Smykowski (2003, p. 97): “the extent to which the statement captures important features of the topic”. Aptness has been widely used in empirical research, especially in support of categorization models of metaphor, and it has also been found to predict categorical form preference in the form

preference task. Jones and Estes (2006) constructed a set of apt and a set of non-apt metaphors, each of these was equally divided between conventional and novel metaphors. They used this set of stimuli in a form preference task and they found that there is a significant difference in the degree of categorical form preference between apt and less apt metaphor. They used their results to argue that the effect of conventionality observed by Bowdle and Gentner (2005) in their experiment was probably due to the confounding effect of aptness. While this opens up a very interesting possibility, it is worth noting that, as Bowdle and Gentner (2008) have suggested, aptness is a ‘vague’ notion that tends to be correlated with several characteristics of metaphors and that, for this reason, it is arguably of little theoretical interest.

2 Metaphor Conventionality and Familiarity

2.1 Metaphor Familiarity

As we noted above, the technical definition of *conventionality* proposed by Bowdle and Gentner (2005) has a salient characteristic in that the *conventionality* of a metaphor depends only on the vehicle and not on the whole sentence. In fact, Bowdle and Gentner draw a distinction between *conventionality* and *familiarity*, which is a property of the whole sentence (the topic-vehicle pair). They claim that these two properties are distinct and that “conventional figurative expressions can be either familiar or unfamiliar, depending on [the topic]” (Bowdle & Gentner, 2005, p. 204). For example, “His anger is a fire” and “His panic is a fire” are equally conventional metaphors because they share the same vehicle, but while the former sounds familiar the latter sounds quite unfamiliar. The notion of familiarity, intuitive as it may be, has not received a precise and univocal definition in the literature and, as a consequence, its effect on metaphor processing has received relatively little attention. In what follows, we aim at investigating the role of familiarity in metaphor comprehension.

To begin with, let us discuss some theoretical considerations which point to an effect of familiarity in metaphor comprehension. From a theoretical point of view, although the career of metaphor hypothesis stresses the importance of the vehicle in determining whether the processing will be through comparison or categorization, it is possible to postulate a role for the topic-vehicle pair as well. According to Bowdle and Gentner (2005), novel metaphors are processed as comparison through a process of horizontal alignment. This is true for every novel metaphor, independently of the topic-vehicle pairing. However, it is possible to hypothesise that the repeated encounters of a novel vehicle with the same topic results in an alignment that is stable (i.e., ‘abstractable’) enough to allow a categorization processing. In other terms, the repeated activation of the same cluster of aligned properties could result in the creation of associative shortcuts. They would allow a particular metaphorical category to be immediately available in association with a specific topic-vehicle pair. This, however, would only be accessed for that particular topic-vehicle pairing and would not be ‘conventionalised’ or stored as an additional lexical meaning. In other words, a novel vehicle that appears in a familiar metaphor may activate an abstract category specific for that topic. This would allow processing the familiar-novel metaphor as a categorization. For example, the familiar (and novel) metaphor “Her faith is an anchor” may activate a metaphoric category for ‘anchor’ that is specific to the ‘faith-anchor’ pairing. This metaphoric category may not be available for unfamiliar metaphors which have ‘anchor’ as a vehicle (e.g., *Her pension is an anchor*).

The empirical investigation of the notion of familiarity requires it to be operationalized in a tractable way. The limited psycholinguistics literature on the effect of familiarity on metaphor comprehension has offered the following operational definitions.

Blasko and Connine (1993) adopted a subjective measure of familiarity, which they defined as “the perceived experience with the metaphor” (Blasko & Connine, 1993, p. 305). Familiarity is said to reflect direct experience with the metaphor (topic-vehicle pair). They found that figurative interpretations for familiar metaphors were as accessible as literal interpretations. Conversely they found that figurative interpretations of unfamiliar metaphors were less accessible than literal interpretations. It is worth questioning whether this familiarity measure really targets the topic-vehicle relation at issue. As Blasko and Connine suggest, the perceived experience with a particular metaphor may also be affected by previous experiences with a set of related metaphors (e.g., from the same conceptual domains).

Following a more promising route, Thibodeau and Durgin (2011) have suggested an alternative *objective* measure of metaphor familiarity. They found that subjective ratings of familiarity (obtained by asking participants to give ratings for a set of metaphors used in previous research (Jones & Estes, 2006)) were highly correlated to metaphor frequency (obtained using Google as a corpus). Because of this, they propose to regard metaphor frequency as an objective measure of familiarity. Interestingly, while they found that familiarity and aptness were significantly correlated, they did not find any significant correlation between conventionality and familiarity.

2.2 Our Study

Previous studies investigating the effect of conventionality on metaphor processing have found conflicting results (e.g., Bowdle & Gentner, 2005; Jones & Estes, 2006). These studies did not take into account all the relevant variables (e.g., conventionality and aptness) at the same time. This is an obvious problem as many of these variables seem to be correlated. Furthermore, apart from very few cases (Thibodeau & Durgin, 2011), empirical research on conventionality has largely ignored familiarity. The question, therefore, remains of how the effects of these conventionality and familiarity on metaphor processing stand with respect to each other. The present study aims at filling this gap in the literature.

We will focus on two distinct ways of operationalizing personal experience with a metaphor: experience with the vehicle across all its different topics (i.e., conventionality as proposed by Bowdle and Gentner (2005)) and experience with the specific topic-vehicle pairing (i.e., familiarity). In order to have a clearer understanding of how personal experience influences how metaphors are processed, these two variables (familiarity and conventionality) should be pitted against each other. Therefore, in this study we compare the effects of familiarity and conventionality in a form preference paradigm employed by Bowdle and Gentner in order to address some open questions: Is conventionality the sole factor determining whether a metaphor is processed as categorization rather than comparison? Or does familiarity also have an effect?

In our study we investigated three variables: (i) familiarity (ii) conventionality and (iii) aptness. We actively manipulated familiarity and conventionality in a set of stimuli. We also measured aptness in the stimuli and introduced it in the analysis in order to control for the potential confounding effect of this variable, which has been found to be correlated with both conventionality and familiarity (Jones & Estes, 2006; Thibodeau & Durgin, 2011). We measured conventionality and aptness with three rating studies in which we followed the procedures used by Jones and Estes (2006) and Glucksberg and Haught (2006). For familiarity we used metaphor frequency (i.e., objective familiarity) instead of subjective familiarity as a measure. We used Google search as a corpus to construct the familiarity score as there are several reasons to consider Google equivalent to traditional linguistic corpora

(Kilgarriff & Grefenstette, 2003). We used the scores of the three variables to divide the stimuli into four groups which we then used in the main experiment (form-preference task).

We predict that that familiarity will have an effect on the form preference task, consistently with the role of the topic-vehicle pairing proposed above. Familiar metaphors, whether conventional or novel, should favour a categorization-based process of interpretation. In addition, we expect both aptness and conventionality to have an effect. We expect conventionality to predict categorization-form preference, consistently with the results of Bowdle and Gentner (2005). Although we are not actively manipulating aptness, we expect apt metaphors to be preferred in the categorical form, consistently with the results of Jones and Estes (2006).

2.3 Methods

The process of constructing the stimuli involved an initial stimulus selection and three rating studies.⁵ Each phase is described below.

2.3.1 Initial Stimulus Selection. A list of 83 pairs of nominal metaphors was constructed. Each pair contained one familiar metaphor and one unfamiliar metaphor. The metaphors in each pair had the same vehicle but different topics (e.g., *Her beauty is a light* - *Her generosity is a light*). Operational definitions of familiar and unfamiliar metaphors (a, b) in conjunction with Google engine in the verbatim search mode were used to classify the metaphors.⁶

- a. A *familiar metaphor* is a metaphor that produces at least 100 results when searched verbatim and without a determiner for the topic (e.g., “company is a family” for the metaphor *This company is a family*) on the Google engine and at least 10 results when searched verbatim, without a determiner for the topic and followed by the conjunction ‘and’ (e.g., “company is a family and”).⁷
- b. An *unfamiliar metaphor* is a metaphor that when searched verbatim on the Google engine with no determiner for the topic (e.g., *librarian is a doll*) produces a maximum of five results in which it is used metaphorically. Results in which the metaphor is used literally as part of a larger expression (e.g., *That librarian is a doll collector*) are not included in the count.

One third of the metaphors had the topic preceded by the demonstrative ‘*this*’ or ‘*that*’ (e.g., *That child is a sponge*), one third by the possessive pronoun ‘*my*’, ‘*his*’ or ‘*her*’ (e.g., *Her faith is an anchor*) and in the last third the topic was not preceded by anything (e.g., *Alcohol*

⁵ All the experiments described here were approved by the appropriate ethics committee at University College London. The participants of all the experiments gave their consent to take part in this research after they were informed about the experimental procedure and given contact details of the researcher.

⁶ The number of occurrences of the figurative statements in the comparison form (e.g., “company is like a family”) was not considered in the operational definitions. However, there was no unfamiliar metaphor with more than 12 occurrences in the comparison form, and no familiar metaphor with 0 occurrences in the comparison form.

⁷ The requirement for familiar metaphors to produce at least 10 results when followed by the conjunction ‘and’ is motivated by the concern that some vehicles are very common as modifiers in nominal compounds (e.g., “company is a family run business”) and failing to ascertain that a metaphor appears in isolation could have led to a fallacious assessment of the metaphor conventionality.

is a crutch). The metaphor vehicles were always preceded by the indefinite article ‘a/an’ or by nothing (e.g., *Her beauty is a light*, *Her caress is honey*).

2.3.2 First Rating Study: Generation of Interpretations. This rating study was conducted in order to generate interpretations for the vehicle of each metaphor. These interpretations are needed in order to rate metaphors’ conventionality (second rating study). Eleven adult, monolingual English speakers were recruited as participants via personal contacts and they were sent the experiment as a Microsoft Word document via email. Participants did not receive any compensation for taking part in the experiment. Participants were instructed to read a list of 83 pairs of metaphors and to “write a single word that best captures the metaphorical meaning conveyed by both sentences in the pair” in the appropriate blank space and they were given four examples (*Sentence 1*: A ferry is a bridge. *Sentence 2*: A telephone is a bridge. *Interpretation*: Connecting). Metaphors in each pair shared the same vehicle. The metaphors in the list were in a random order⁸. After completing the task, participants were asked to email the filled document to the researcher as an email attachment. For each metaphor vehicle there were 11 interpretations, one for each participants. The most frequent interpretation for each vehicle was selected as the preferred interpretation for that particular vehicle. Similar interpretations (i.e., words) were grouped together when they were listed as synonyms in the Collins Thesaurus of English. For example, the modal interpretation of the vehicle ‘family’ was ‘welcoming’. These interpretations were used in the second rating study.

2.3.3 Second Rating Study: Conventionality. This rating study was conducted in order to measure the conventionality of metaphor. Forty-nine adult, monolingual English speakers not involved in the previous study were recruited via personal contacts and they were sent the experiment as a web link (Qualtrics website). Ten participants failed to complete the task. The analysis was carried out on the data from thirty-nine participants. Participants were instructed to read a list of 87 statements in the form “X is a family” and assess on a scale ranging from one to seven (1=novel, 7=conventional) how conventional it was to use each vehicle to convey its most frequent interpretation (e.g., “How conventional is it to use the concept FAMILY to convey the meaning WELCOMING?”⁹). The metaphors in the list were presented in a random order. The scores obtained in the experiment were used to calculate the average conventionality score of each metaphor.

2.3.4 Third Rating Study: Aptness. This rating study was conducted in order to measure the aptness of each metaphor. Sixty-six adult, monolingual English speakers not involved in the previous rating studies were recruited via personal contacts and they were sent the experiment as a web link. Participants were all native speakers of English. Six participants failed to complete the task (Qualtrics website). The analysis was carried out on data from sixty participants. Each participant was randomly assigned to one of two groups; they were instructed to read a list of metaphors and rate the aptness of each metaphor or “how well each statement communicates what the speaker intended to communicate” (following the wording of Glucksberg & Haught, 2006, p. 369) on a scale ranging from one to seven (1=not apt at all, 7=very apt). Each group of participants was shown only one of two lists of metaphors. Each list contained 83 metaphors in a random order and no metaphors with the same vehicle.

⁸ In this study and in the following experiments the random order of the stimuli was the same for all the participants (in each version).

⁹ The wording of the instructions was modelled on the procedure employed by Jones and Estes (2006).

The ratings of familiarity, conventionality and aptness obtained for each of the 166 metaphors were used to select the 80 stimuli for the main experiment. All metaphors with an average aptness rating below 3 were discarded to avoid risks of incomprehensibility. From the remaining metaphors a group of novel metaphors (average conventionality score between 3 and 5.20) and a group of highly-conventional metaphors (average conventionality score above 6.10) were selected. Each of these two groups was then divided into categories of familiar and unfamiliar metaphors according to the google-based familiarity score. Thus four experimental conditions of 20 metaphors each were obtained. There were significant differences in aptness across the four conditions ($F(3,76)= 9.08$ $p<0.001$). Means and standard deviations of aptness and conventionality for the stimuli in each experimental condition, alongside with an example from each category, can be observed in **Table 1**. Finally, the 80 metaphors were divided across two lists to avoid presenting metaphors with the same vehicle to the same participant (some but not all of the vehicles were present in two metaphors: one familiar and one unfamiliar). Thus there were two versions of the experiment each containing 10 metaphors from each condition (i.e., a total of 40) and no repeated vehicles.

Table 1

Conventionality and aptness scores of each experimental condition

Experimental conditions	Examples	Conventionality		Aptness	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Unfamiliar-novel	<i>Her generosity is a light</i>	4.38	0.83	4.17	0.73
Familiar-novel	<i>Her beauty is a light</i>	4.44	1.05	4.86	0.90
Unfamiliar-conventional	<i>Ice cream is a crutch</i>	6.37	0.19	4.38	0.87
Familiar-conventional	<i>Alcohol is a crutch</i>	6.36	0.18	5.38	0.72
Total		5.37	1.17	4.71	0.93

2.3.5 Participants One hundred and forty-five participants, who did not take part in any of the rating studies, were recruited via personal contacts and social media (Facebook) and they were directed to the online experiment with a web link. Participants were screened with initial questions and only adult monolingual native speakers of English initiated the task. Fifty-four participants were excluded because they did not complete or initiate (i.e., initial screening) the task. Nine participants were excluded because they did not follow the instructions correctly. This left 43 participants in one version and 39 participants in the other version of the experiment.

2.3.6 Procedure Participants were randomly assigned to one of the two versions of the experiment, which were hosted online (Qualtrics website). Each version contained 40 metaphor-simile pairs (e.g., *This company is a family*, *This company is like a family*) separated by a seven-point scale. The metaphors in each version were in a random order which was the same for all participants, and for half of the items (the same half for all participants) the order of the two forms on the scale (comparison, categorization) was inverted. Participants were instructed to indicate whether each metaphor sounded “more

natural” in one form or the other by marking their answer on the scale closer to the preferred form (wording adapted from Bowdle & Gentner, 2005, p. 201).

2.4 Results

The scores were converted to indicate categorical form preference. The score was on a seven-point scale in which ‘1’ indicated preference for the comparison form and ‘7’ indicated preference for the categorical form. One experimental item was excluded from the analysis because we realized that it had a literal interpretation also in its categorical form (*Age is a number*).

The means of the conditions were calculated by items and by subjects. For the analysis by items, a mean *categorical form preference score* was constructed for each metaphor and the data was analysed with a 2x2x2 ANOVA with categorical form preference as the dependent variable and Conventionality, Familiarity and Version as between factors. For the analysis by subjects the mean scores for each subject in each condition were calculated (Figure 1) and analysed with a 2x2x2 repeated measures ANOVA with Conventionality and Familiarity as within-subjects factors and Version as a between-subjects factor. Aptness was not included in this initial analysis.

The overall model of the ANOVA by items was not significant ($F(7,71)=1.17$, $p=.33$, partial $\eta^2=0.10$) and there was no significant effect of conventionality ($p=.06$), familiarity ($p=.39$), version ($p=.68$), or any of their interactions (all $ps>.05$). The repeated measures ANOVA by subjects showed a significant effect of conventionality ($F(1,80)=47.34$, $p<.001$, partial $\eta^2=0.37$), a significant effect of familiarity ($F(1,80)=30.92$, $p<.001$, partial $\eta^2=0.28$) and no significant effect of version ($F(1,80)=0.17$, $p=.90$). The analysis also revealed significant effects of three interactions: conventionality*familiarity ($F(1,80)=75.24$, $p<.001$, partial $\eta^2=0.48$); conventionality*version ($F(1,80)=5.18$, $p=.03$, partial $\eta^2=0.61$) and the three-way conventionality*familiarity*version interaction ($F(1,80)=21.65$, $p<.001$, partial $\eta^2=0.21$). The remaining interactions were non-significant ($p>.1$). While the variances in each group in the items analysis were very large, the variances in the analysis by subjects were much smaller, which made the analysis more powerful. This resulted in the variables having an effect in the analysis by subjects but not in the analysis by items. This suggests that the effects of the variables may not be reliable across items. Both the effect of familiarity and the effect of conventionality were in the expected direction. Familiar metaphors and Conventional metaphors were preferred in the categorical form significantly more than unfamiliar and novel metaphors respectively. However, it is evident from the chart (**Figure 1**) that one condition in particular is driving these effects: the condition of Conventional-Familiar metaphors. This also explains the relatively large effect size of the interaction between Conventionality and Familiarity.

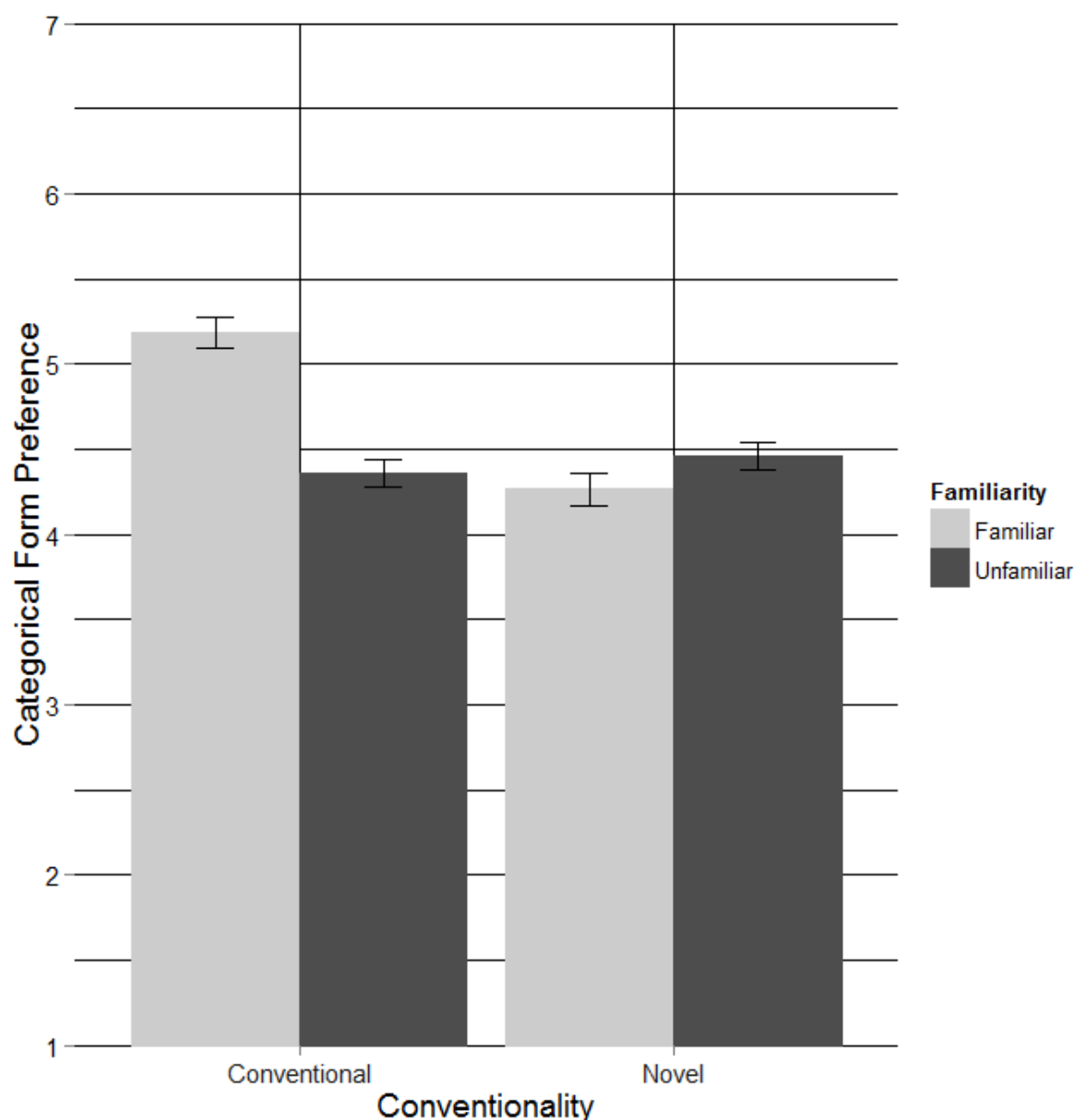


Figure 1

Mean subject scores for the four conditions from the Repeated Measure ANOVA by subjects (error bars: SE). Higher scores correspond to preference for the categorization form.

The data was analyzed with an ANCOVA (by items only) to take into account the effect of Aptness. The 2x2x2 ANCOVA had Conventionality, Familiarity and Version as factors and Aptness as a covariate. The overall model was significant ($F(8,70)=6.15$, $p<.001$, partial $\eta^2=0.41$). Aptness did have a significant effect ($F(1,7)=36.84$, $p<.001$, partial $\eta^2=0.34$) in the direction we expected (more apt metaphors were preferred in the categorization form). As with the main ANOVA there were no significant effects of Familiarity ($F(1,77)=3.82$, $p=.055$), Conventionality ($F(1,77)=0.76$, $p=.38$), Version or any of the interaction (all $ps>.1$). The estimated marginal means (displayed in **Table 2**) clearly show that when we factor out the effect of aptness, the effect of familiarity changes direction and unfamiliar metaphors are preferred in the categorization form more than familiar metaphors.

Table 2*Estimated marginal means for the four conditions*

Variables		Categorical preference	
		<i>M</i>	<i>SE</i>
Familiar	Conventional	4.62	0.23
	Novel	4.10	0.22
Unfamiliar	Conventional	4.76	0.22
	Novel	4.90	0.25
Grand Mean		4.60	0.11

3 Post-hoc Analysis on Metaphor Dominance

The results we found in our planned analyses were inconclusive. Interestingly, though, the variance of Categorical form preference across subjects was quite small: the standard deviation of scores given to each item was 1.65 scale points. This means that the majority of participants gave each item a score that was at most one point above or below the average score for that item. That is, our participants tended to agree on how to rate each item. However this apparent consistency was not explained by either of the two experimental variables.

While these variables pertain to the metaphorical use of the word, none of them take into account the use of its *literal meaning*. Interestingly, this might affect the processability of the metaphor itself. To explore this possibility, we investigated Metaphor Dominance (of the vehicle), that is, the proportion of metaphorical uses of each vehicle relative to its literal uses. This variable can also be described as the dominance of the metaphorical sense over the literal sense. We collected two new continuous variables for our data and we used them in a post-hoc analysis: Metaphor Dominance and Vehicle Frequency. The latter, which is a measure of word-frequency independent from considerations of word-use (metaphorical or literal), was introduced in order to assess whether Metaphor Dominance had an effect over and above the overall frequency. Recent studies have collected metaphor dominance scores in the following ways. Sardinha (2008) first annotated 432 metaphor vehicles from a Portuguese corpus of conference calls and then constructed a metaphor dominance score for each of this by counting how many out of 100 uses of these vehicles (in a different corpus) were metaphorical. Lai, Curran, and Menn (2007) collected metaphor dominance scores for nine verbs by inspecting 50 occurrences of each word in a corpus (BNC, 2007) and recording it as metaphorical or literal. Bethard, Lai, and Martin (2009) then used these metaphor dominance ratings in the development of a metaphor dominance estimator. They also advocated that psycholinguistic experiments on metaphor processing should control for metaphor dominance (in addition to word frequency).

Vehicle Frequency was collected by simply recording the frequency of each metaphor vehicle in the British National Corpus (BNC). We collected Metaphor Dominance by randomly sampling fifty uses of each vehicle in context in the BNC, recording the number of

these instances that used the metaphorical sense and then converting this number to a proportion. Homonyms with an unrelated sense were not related in the count of literal uses. Therefore, each of our stimuli was associated with a proportion that could range from zero (literal uses of the vehicle are much more frequent than its metaphorical uses) to 1 (metaphorical uses of the vehicle are much more frequent than its literal uses).

As shown by the correlation matrix below (**Table 3**), Metaphor Dominance is significantly correlated with categorical form preference, such that metaphors with higher metaphor dominance of the vehicle were preferred in the categorical form. Vehicle Frequency is significantly and negatively correlated with categorical preference and with Conventionality score (in this post-hoc analysis we use the original conventionality scores as a continuous variable instead of the binary variable we constructed from them for the main analysis), such that metaphors with a high frequency vehicle are novel and preferred in the simile form.

Table 3

Correlation coefficients (Pearson's r) between study variables

Measure	1	2	3	4	5	6
1. Categorical preference	--	.10	.26*	.58**	-.33**	.49**
2. Familiarity		--	.09	.41**	-.47	-.02
3. Conventionality score			--	.25*	-.31**	.19
4. Aptness				--	-.18	.27*
5. Frequency					--	-.17
6. Metaphor Dominance						--

*. $p < .05$. **. $p < .01$

In order to test whether Metaphor Dominance can do some explanatory work as a predictor of categorical preference once the other variables are taken into account we constructed two regression models. In Model 1 we regressed categorical preference on Aptness, Familiarity, Conventionality score and Vehicle Frequency. In Model 2 we regressed categorical preference on Aptness, Familiarity, Conventionality score, Vehicle Frequency and Metaphor Dominance. The predictors were entered hierarchically in our regression model in the order just mentioned¹⁰. Therefore, in Model 2 the effect of Metaphor Dominance was measured after the effect of the other predictors was taken into account. In **Table 4** we show these two models. We see only Aptness and Metaphor Dominance have a significant effect on Categorical form preference and that the proportion of variance explained (R^2) increases when we include Metaphor Dominance in the model. We also compared the two models with an F-ratio test and found that Model 2 is significantly better than Model 1 ($F(1,73) = 13.65$, $p < .001$).

¹⁰ The rationale for this ordering is the following: we first entered the variable used in the previous literature, Conventionality, then our original variable of interest, Familiarity, then our control variable, Aptness, then Vehicle Frequency, and lastly Metaphor Dominance because we want to measure its effect over an above any effect of word frequency.

Table 4
Hierarchical regression on categorical form preference

	R ²	Coefficient	SE	t	p
Model 1	0.40				
constant		0.82	0.76	1.07	.23
conventionality score		0.08	0.10	0.86	.39
familiarity		-0.45	0.24	-1.87	.07
Aptness		0.78	0.14	5.65	<.001
vehicle frequency		<0.001	<0.001	-1.41	.16
Model 2	0.50				
constant		1.08	0.71	1.53	.13
conventionality score		0.06	0.01	0.66	.51
familiarity		-0.32	0.22	-1.46	.15
aptness		0.65	0.13	4.97	<.001
vehicle frequency		<0.001	<0.001	-1.24	.21
metaphor dominance		0.04	0.009	3.69	.001

4 Discussion

In our main analysis we did not find evidence that Categorical form preference (how much a figurative statement is preferred as a metaphor rather than as a simile) is determined by Conventionality or Familiarity. Instead, we found that Aptness is a significant predictor of categorical form preference.

In the post-hoc analysis we found that categorical form preference is significantly and negatively correlated with Vehicle Frequency (i.e., metaphors with frequent words as a vehicles are preferred as similes) and significantly and positively correlated with Metaphor Dominance (metaphors with vehicles that are frequently used in a metaphorical sense are preferred in the categorical form). Furthermore, we found that once the variables from the main analysis are taken into account (Conventionality, Familiarity and Aptness), Vehicle Frequency does not have a significant effect whereas Metaphor Dominance is still a significant predictor of categorical form preference.

Our prediction that conventionality would have an effect was not confirmed. This is in contrast with Bowdle and Gentner (2005) as they found that conventional metaphors were preferred in the categorical form significantly more than novel metaphors. Similarly to Jones and Estes (2006) and Chiappe, Kennedy, and Smykowski (2003), who all observed an effect of aptness but not an effect of conventionality on form preference, we failed to replicate the results of Bowdle and Gentner.

We believe that the cause of such different results might be in the stimuli used. The most prominent difference is that the conventionality ratings for their group of novel metaphors ($M=2.32$) are lower than both the ratings of our novel metaphors ($M=4.41$) and those used by Jones and Estes (2006) ($M=3.42$).¹¹ This suggests two explanations for the difference in results. One explanation is that our novel metaphors were too conventional to show an effect of conventionality. The other, more speculative, explanation is that comprehensibility is playing a role in the experiment of Bowdle and Gentner (2005) but not in ours. In the phase of stimuli construction we eliminated many metaphors with very low conventionality because they were hard to comprehend. In fact, it has been argued that novel metaphors are also harder to comprehend (Giora, 1997). If low comprehensibility favours preference for the simile form, the effect of conventionality observed by Bowdle and Gentner could have been confounded with comprehensibility. While these explanations are speculative, we think that it is possible that we would have found an effect of conventionality with a different set of items.

Our prediction that familiarity would have an effect was also not confirmed. Again, this may depend on the set of stimuli we used. The construct of familiarity we were interested in was the subjective experience with a topic-vehicle pairing that Blasko and Connine (1993) describe. However, we choose to use an objective corpus-based proximate measure inspired by Thibodeau and Durgin (2011) as a proximate measure of this construct. We gave operational definitions of familiar and unfamiliar metaphors that were corpus based. These operational definitions may not accurately reflect the relevant theoretical construct, despite the correlation between subjective and objective measures of familiarity reported by Thibodeau and Durgin. It is an open issue whether a set of stimuli based on a different measure of familiarity could reveal an effect of this variable.

Our prediction that aptness would have an effect in the form preference task was confirmed. This result is not surprising as the effect of aptness on metaphor processing has been observed several times and in different paradigms (Chiappe et al., 2003; Jones & Estes, 2006). As previously discussed, aptness is a very vague and composite theoretical notion that is correlated with several other properties of metaphor (e.g., comprehensibility, relationality, etc.). For this reason, Gentner and Bowdle (2008) argue that in order for aptness to become a useful and explanatory variable for metaphor processing research it should be defined more strictly or broken down into its components.

As they stand, the results do not support the career of metaphor hypothesis and its claim that conventionality determines a shift in the processing mode of metaphorical utterances (from comparison to categorization). The mode of processing metaphorical utterances does not seem to depend on whether the metaphor vehicle has a metaphorical category associated to it (conventional metaphor) or it is frequently paired with a specific topic in a metaphorical statement (familiar metaphor).

Possibly the most surprising aspect of our results is that our analyses by items and by subjects yielded very different results. More specifically, there were patterns in the data that were significant in the subject analysis but not in the items analysis. This is because while there was agreement among subjects as to how to rate the stimuli, there was a lot of variation across the items that was not explained by the predictor variables included in our model.

In the post-hoc analysis we found that the more a word is used to convey a metaphorical rather than its literal meaning, the more the figurative statements where they are used as a vehicle are preferred in their categorical form. To draw an example from our data:

¹¹ All scores have been converted to a scale from 1 to 7 for an easier comparison with our ratings.

- (3) a. That talk show is like a war.
- b. Renovation is a struggle.

The term *war* is used metaphorically relatively rarely (Metaphor Dominance = 0.06) and the sentence in (3a) is preferred in the comparison form. The term *struggle* is used metaphorically relatively often (Metaphor Dominance = 0.88) and sentence in (3b) is preferred in the categorical form. That is, low metaphor dominance predicts preference for the comparison form, whereas high metaphor dominance predicts preference for the categorical form.

The fact that the dominance of the metaphorical sense of the vehicle, as opposed to the literal one, has an effect on this task suggests that this characteristic also affects metaphor processing. Whatever feature is associated to the metaphorical vehicle (degree of conventionality) and to the metaphorical topic-vehicle pair (degree of familiarity), the relation between the metaphorical and the literal use plays a key role in the processing of the figurative statement. This very interesting conclusion is not totally unexpected when considering the psycholinguistics literature on ambiguity.

The dominance of the metaphorical uses of a word indicates how frequently the word at issue is used metaphorically. This is also a proxy measure of how often we (and our participants) have processed the same word metaphorically, that is, either by retrieving a metaphorical category associated to it (conventional metaphor) or by constructing such a metaphorical category on the fly (non-conventional metaphor). Be there a conventionally associated metaphorical category or not, it is plausible to assume that the more a word is interpreted metaphorically (in proportion to its literal uses), the more its metaphorical interpretative route is facilitated. In other terms, a metaphorical category is more easily retrieved or constructed. We could thus explain our results along the following line: the easier the selection/construction of a metaphorical category, the stronger the preference for the categorical form. Conversely, when the metaphorical use is less frequent, that is, when the word is typically interpreted as conveying its literal meaning, the interpretation of the figurative statement might more heavily rely on such a meaning. For instance, the figurative statement might be interpreted via a process of comparison between the target and the vehicle. This would explain why low Metaphor Dominance is associated with prediction for the comparison form.

This finding is not totally unexpected when considering the psycholinguistics literature on lexical ambiguity, which has consistently shown that the dominance of a word's different (literal) interpretations plays a key role in how word meanings are represented and accessed during comprehension (see Rodd, Lopez Cutrin, Kirsch, Millar, and Davis, 2013). For example the word "pen" has a high-frequency literal meaning (a writing implement) and a low-frequency literal meaning (enclosure for animals). Numerous experiments have shown that when an ambiguous word occurs in a context in which both meanings are plausible (e.g., "The man thought that the pen..."), readers are biased to retrieve the more frequent meaning (e.g., Rayner & Duffy, 1986). More relevant to the current study are the strong effects of dominance that have been shown for sentences where the ambiguous word is presented in a constraining context (e.g., "The man enclosed the livestock with the pen"). Numerous studies have shown that even when the sentence context is only consistent with one interpretation, readers are relatively slow to retrieve the less frequent meaning compared to either the more frequent meaning or to an unambiguous baseline (e.g., Duffy, Morris, & Rayner, 1988), presumably because the reader must first access and reject the alternative dominant interpretation. It is therefore perhaps not surprising that direct access to a categorical meaning representation may be particularly challenging for metaphorical meanings that must compete with higher-frequency literal meanings (e.g., "war").

The results of the post-hoc analysis open up a very interesting direction for future research, introducing a new variable in the lively debated literature on metaphor understanding. The role of Metaphor Dominance in metaphor understanding has not only been underestimated in the psycholinguistics literature (Bethard et al., 2009) but it has been arguably neglected in theoretical linguistics. Our investigation call for experimental investigation on whether its effect extends beyond the form preference task and theoretical inquiry on how it may affect metaphor comprehension.

References

- Bethard, S., Lai, V. T., & Martin, J. (2009). Topic Model Analysis of Metaphor Frequency for Psycholinguistic Stimuli. *NAACL (North American Chapter of the Association for Computational Linguistics) Workshop on Computational Approaches to Linguistic Creativity*. Boulder, June 2009.
- Blasko, D. G., & Connine, C. M. (1993). Effects of familiarity and aptness on metaphor processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(2), 295-308.
- Bowdle, B. F., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, 112(1), 193-216.
- The British National Corpus, Version 3 (BNC XML Edition). (2007). Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: <http://www.natcorp.ox.ac.uk/>
- Chiappe, D., Kennedy, J. M., & Smykowski, T. (2003). Reversibility, aptness, and the conventionality of metaphors and similes. *Metaphor and Symbol*, 18(2), 85-105.
- Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language*, 27, 429-446.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7(2), 155-170.
- Gentner, D., & Bowdle, B. (2008). Metaphor as structure-mapping. In R. Gibbs (Ed.) *The Cambridge Handbook of Metaphor and Thought* (pp. 109-128). New York, NY: Cambridge University Press.
- Gentner, D., & Clement, C. (1989). Evidence for relational selectivity in the interpretation of analogy and metaphor. *The Psychology of Learning and Motivation: Advances in Research and Theory*, 22, 307-358.
- Gentner, D., & Wolff, P. (1997). Alignment in the processing of metaphor. *Journal of Memory and Language*, 37(3), 331-355.
- Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, 8, 183-206.
- Glucksberg, S. (2001). *Understanding figurative language: From metaphors to idioms*. Oxford University Press.
- Glucksberg, S. (2003). The psycholinguistics of metaphor. *Trends in cognitive sciences*, 7(2), 92-96.
- Glucksberg, S., & Haught, C. (2006). Can Florida become like the next Florida? When metaphoric comparisons fail. *Psychological Science*, 17(11), 935-938.
- Glucksberg, S., & Keysar, B. (1990) Understanding metaphorical comparisons: Beyond similarity. *Psychological Review*, 97(1), 3-18.
- Gregory, M. E., & Mergler, N. L. (1990). Metaphor comprehension: In search of literal truth, possible sense, and metaphoricity. *Metaphor and Symbol*, 5(3), 151-173.
- Jones, L. L., & Estes, Z. (2006). Roosters, robins, and alarm clocks: Aptness and conventionality in metaphor comprehension. *Journal of Memory and Language*, 55(1), 18-32.
- Kilgarriff, A., & Grefenstette, G. (2003). Introduction to the special issue on the web as corpus. *Computational Linguistics*, 29(3), 333-347.
- Lai, V. T., Curran, T., & Menn, L. (2007). *The comprehension of conventional and novel metaphors: An ERP study*. Poster presented at the 13th Annual Conference on Architectures and Mechanisms for Language Processing (AMLaP), August 2007. Turku, Finland.
- Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading – Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, 14, 191-201.
- Rodd, J. M., Lopez Cutrin B., Kirsch H., Millar A., & Davis M. H. (2013). Long-term priming of the meanings of ambiguous words. *Journal of Memory and Language*, 68 (2), 180-198.
- Sardinha, T. B. (2009). Metaphor probabilities in corpora. In M. S. Zanotto, L. Cameron, & M. do Couto Cavalcanti (Eds.), *Confronting Metaphor in Use* (pp.127-147). John Benjamins.
- Sperber, D., & Wilson, D. (2008). A deflationary account of metaphors. *The Cambridge handbook of metaphor and thought* (pp. 84-105). Cambridge: Cambridge University Press.
- Thibodeau, P. H., & Durgin, F. H. (2011). Metaphor aptness and conventionality: A processing fluency account. *Metaphor and Symbol*, 26(3), 206-226.

Wilson, D., & Carston, R. (2007). A unitary approach to lexical pragmatics: relevance, inference and ad hoc concepts. In N. Burton-Roberts (Ed.), *Pragmatics* (Palgrave Advances in Linguistics) (pp. 230-260). Basingstoke: Palgrave Macmillan.