Metaphor and hyperbole: Testing the Continuity Hypothesis

Paula Rubio-Fernández, University College London

Catherine Wearing, Wellesley College

Robyn Carston, University College London and

Centre for the Study of Mind in Nature, Oslo

Abstract

In standard Relevance Theory, hyperbole and metaphor are categorized together as loose uses of language, on a continuum with approximations, category extensions and other cases of loosening/broadening of meaning. Specifically, it is claimed that there are no interesting differences (in either interpretation or processing) between hyperbolic and metaphorical uses (Sperber and Wilson 2008). In recent work, we have set out to provide a more fine-grained articulation of the similarities and differences between hyperbolic and metaphorical uses and their relation to literal uses (Carston & Wearing 2011, forthcoming). We have defended the view that hyperbolic use involves a shift of magnitude along a dimension which is intrinsic to the encoded meaning of the hyperbole vehicle, while metaphor involves a multi-dimensional qualitative shift away from the encoded meaning of the metaphor vehicle. In this paper, we present four experiments designed to test the predictions of this analysis, using a variety of tasks (paraphrase elicitation, self-paced reading and sentence verification). The results of the study support the view that hyperbolic and metaphorical interpretations, despite their commonalities as loose uses of language, are significantly different.

1.   Introduction

While metaphor is usually treated as a distinctive, even unique, use of language, hyperbole has

generally been categorized, by those few theorists and experimentalists who have considered it,

as belonging with irony and meiosis (understatement) (Fogelin, 1988; Clark, 1996; Colston,

1997; Gibbs 2000).[1] On the relevance-theoretic account of non-literal language, by contrast, metaphor and hyperbole are not even considered "genuinely distinct categories, at least from a descriptive, psycholinguistic, or pragmatic point of view" (Sperber and Wilson 2008: 95).[2] Instead, Sperber and Wilson argue that metaphor and hyperbole fall along a continuum of loose uses, with no effective criteria for distinguishing between them and no interesting generalizations applying to one but not the other.

The present study challenges this strong 'continuity' view of metaphor and hyperbole and argues for a more modest position. A strongly unified account is appealing because considerable psycholinguistic evidence disconfirms models of non-literal language use (such as those based on Gricean pragmatics (Grice 1975)) which require a two-step interpretive process in which the literal interpretation is inevitably tried first and non-literal possibilities are considered only when the literal meaning clearly violates a pragmatic maxim (Gibbs 1994, chapter 3). However, the claim that there are *no genuine categories* of hyperbole or metaphor 'from descriptive, psycholinguistic, or pragmatic points of view' seems too strong. In this paper, we show that there are *at least* descriptive and psycholinguistic differences between metaphor and hyperbole worth

---

[1] The last three decades have seen a proliferation of experimental studies on metaphor interpretation (Glucksberg et al. 1982; Gildea & Glucksberg 1983; Gibbs & Gerrig 1989; Keysar 1989; Blasko & Connine 1993; Giora 1997; Kintsch 2000; Gernsbacher et al. 2001; Noveck et al. 2001; Coulson & van Petten 2002; Bowdle & Gentner 2005; Gibbs 2006; Almor et al. 2007; Hussey & Katz 2009; Wolff & Gentner 2011; Giora et al. 2012). By contrast, hyperbole remains relatively understudied. It is mentioned briefly in Gricean pragmatics as a case of flouting a conversational maxim, and in relevance-theoretic pragmatics as a kind of loose use, but neither of these theoretical accounts makes any testable predictions about the nature of hyperbole interpretation or the processes it recruits. There is some corpus-based empirical research (e.g., Kreuz et al. 1996; Kreuz et al. 1998; McCarthy & Carter 2004; Cano-Mora 2009; Claridge 2011), a few experimental studies that look at the role of hyperbole in the expression of irony (Kreuz & Roberts 1995; Colston & Keller 1998; Colston & O'Brien 2000; Filippova & Astington 2010) and one recent study comparing hyperbole and metaphor interpretation (Deamer et al. 2010).

[2] Irony is given an altogether different treatment within relevance theory (see, for instance, Wilson & Sperber 1992; 2012).

paying attention to.[3] In section 2, we briefly examine the claim that metaphor and hyperbole are not genuinely distinct from a descriptive point of view. In section 3, we report our experimental study, which challenges the claim that there are no psycholinguistic differences between the two. We conclude in section 4 by clarifying the resulting positions of metaphor and hyperbole. To anticipate: while a single pragmatic mechanism may account for the understanding of both of these non-literal uses of language, significant descriptive and psycholinguistic differences between them must be respected.

2.   Descriptive differences between hyperbolic and metaphorical uses

As just noted, the intuition that metaphor is a distinctive, even unique, use of language is widespread. In trying to characterize the special nature of metaphor, many theorists have talked of it as involving a mapping across distinct cognitive domains (e.g. Lakoff 1993; Clausner & Croft 1997; Fauconnier & Turner 1998; Kövescses 2002; Steen 2008; Tendahl & Gibbs 2008). This view can be briefly exemplified by so-called 'double-function' adjectives: compare 'cold/frosty/warm/sunny/bright/breezy/dry/harsh/rough' as predicated of the weather or of a person.[4] When the adjectives are used to describe a person, there is a projection from the domain of the physical to the domain of the psychological, specifically to human personality traits.

But no one has felt any inclination to make comparable claims about hyperbole; there is no talk in the literature of hyperbolic domain shifts. The standard view of hyperbole (to the

---

[3] We take descriptive differences to be the sort of thing that can be distinguished by pre-theoretical intuitions or the analysis of canonical examples, whereas psycholinguistic differences are elicited through testing under experimental conditions. Pragmatic differences are differences in the explanations or models provided by a pragmatic theory (such as Relevance Theory) for observed descriptive or psycholinguistic differences. We will take up the question of pragmatic differences in section 4.

[4] The metaphorical senses of many of these words have become conventionalised but that is orthogonal to the discussion here.

extent that there is one) is that it is merely an exaggerated expression of a literally applicable property. So when someone says 'I have trillions of essays to mark', although she is clearly exaggerating, still we take it that she has some largish number of essays to mark (more than she expected or wants); that is, in comprehending this use of the word 'trillions', our thought doesn't shift to a conceptual domain which is distinct from that of the literal meaning. Similarly for an utterance of 'Our morning jog was a marathon' when the run in question was in fact a slow two-mile circuit of the local park: interpretation of the utterance stays within the domain of physically demanding instances of running, of which actual marathons are a paradigm case.

In our view, a close inspection of clear examples shows that while hyperboles involve only a change in magnitude between the concept encoded and the concept expressed, metaphors involve a qualitative change, resulting in a difference in kind (Carston and Wearing 2011). As we see it, hyperbolic uses of language are straightforward loose uses: one of the essential (perhaps defining) features of the lexically encoded meaning consists of a point or interval on a scalar dimension and this point or interval clearly outstrips the reality described, so must be relaxed or broadened. For instance, 'trillions' is a lot higher on the number scale than the actual number of essays needing to be marked, so the concept expressed is understood as denoting a considerably broader interval or range (of large numbers) on that scale; a 'marathon' is considerably longer than the distance of the jog round the park, so the concept expressed is understood as denoting a broader range of tiring runs.

Metaphorical interpretations, by contrast, are rather more complex: the shift in meaning is never merely quantitative. Describing someone as 'cold' or 'breezy' requires entirely dropping the defining or logical dimensions of these words (that they pertain to physical temperature or weather), rather than merely moving along an encoded scalar dimension; similarly, talking of an

4

intellectual activity as a 'marathon' ('Writing my thesis was a marathon') requires dropping the defining feature of a run or race over a stretch of ground, rather than extending some scalar dimension, such as the distance or time associated with it. Furthermore, metaphorical uses typically involve focusing on some more peripheral features(s) of the information associated with the literal meaning (e.g. 'psychological strain' in the case of 'marathon') or on features that only 'emerge' through a process of drawing an analogy between aspects of the topic (say, a particular person) and aspects of the literal meaning of the metaphor vehicle (say, physical coldness or brightness).[5] For example, the psychological and social effects of interacting with a 'cold' person may be experienced as analogous to the effect that physically cold things have on us.

We analyse these descriptive differences between hyperboles and metaphors in greater detail in Carston & Wearing (2011, forthcoming). For the purposes of this paper, however, the following characterisation will suffice: hyperbolic interpretations are those that result from a straightforward loose use of a word, where what is loosened or broadened is the range on a scalar dimension which is intrinsic to the encoded meaning of the word; metaphoric interpretations are those for which there is both a broadening of meaning, because a defining (or logical) property of the word is dropped, and a narrowing, because properties peripherally associated with the word are central to the new expressed meaning.

We turn now to the issue of *psycholinguistic* differences between hyperbole and metaphor. On the basis of the qualitative/quantitative distinction just given, we can make the following testable predictions about differences in the interpretation of hyperboles and metaphors: (1) The interpretation of metaphors is much more likely than the interpretation of

---

[5] For discussion of the possible role of analogy within a relevance-theoretic account, see Wearing (2014).

5

hyperboles to include so-called 'emergent' properties, that is, properties which feature prominently in the pragmatically inferred ad hoc concept but are not associated with the lexical concept from which it was derived; e.g. the property of 'requiring several years of sustained effort of thinking and writing' for the metaphorical use of 'marathon' discussed above; (2) The interpretation of hyperboles, more than metaphors, involves the attribution of what we label 'intra-domain' properties, that is, properties that are not only essential to the pragmatically derived concept but are also central to the lexical concept from which it is derived; e.g. the property of 'physically demanding run' for the hyperbolic use of 'marathon' discussed above; (3) From the previous two predictions there follows the more general prediction that hyperbolic interpretations are semantically closer to the lexical vehicle from which they are derived than are metaphorical interpretations. The standard relevance-theoretic account, with its strong commitment to the continuity of metaphor and hyperbole, does not predict any of these differences.

3.   A study of the interpretation of nominal hyperbole and nominal metaphor

Our first two experiments (labeled 1A and 1B) investigated the interpretation of nominal hyperboles and metaphors using a paraphrase elicitation task.[6] A general first aim of these experiments was to check the comprehensibility of our materials. The second, more specific, aim was to carry out an initial test of our predictions, thus we focused on the interpretation of metaphors and hyperboles with respect to the two types of properties just outlined: 'intra-domain properties' and 'emergent properties'. The key difference between the two kinds of property is

---

[6] This task has been used successfully before as a means of revealing significant differences between the interpretations of two other intuitively related figurative uses, metaphors and their corresponding similes (see Glucksberg & Haught, 2006).

that, while both are attributed to the topic of the sentence, intra-domain properties are also applicable to the literal vehicle of the hyperbole/metaphor while emergent properties are not.[7] Given the different kinds of meaning shift that we claim arise in hyperbolic and metaphorical interpretation, we predicted that metaphors, more than hyperboles, would elicit emergent properties, while hyperboles, more than metaphors, would elicit intra-domain properties.

If the results confirm our predictions and participants do attribute different types of properties when paraphrasing metaphors from those they attributed when paraphrasing hyperboles, this would provide some initial support for our view that there are significant differences in the kinds of interpretations that people give to metaphors and hyperboles.

Experiment 1A

Method

Participants

20 undergraduates from Princeton University took part in the study. They were all native speakers of English and participated for monetary compensation.

Materials, design and procedure

We constructed 12 novel metaphors and 12 novel hyperboles of the form *X is a Y*. We tried to construct metaphors that were novel (e.g., 'Their baby boy is a *vacuum cleaner*'), while the hyperboles clearly conveyed a sense of exaggeration (e.g., 'In our family, everyone's birthday is

---

[7] Although it is ultimately an empirical question, we assume that intra-domain properties are associated with the vehicle concept in long term memory, while emergent properties are not. The issue of how emergent properties are generated is crucial for any theory of metaphor interpretation (for a relevance-theoretic approach, see Wilson & Carston, 2006; Vega-Moreno, 2007; for an account developed within 'blending' theory, see Grady et al., 2000).

a *national holiday*') [8]. Each figurative expression was preceded by a short context in order to facilitate comprehension (mean number of words per item: 24 for the hyperboles and 27 for the metaphors). We used 12 conventional idioms in American English as fillers. The idioms were also preceded by a short context (mean number of words per item: 20).

An online questionnaire was constructed with all the materials presented in a fixed random order. Participants were given the following instructions:

"We are interested in how learners of English come to understand figurative language. In this study we investigate how native speakers of English help learners understand certain expressions that might not have a direct translation in their mother tongue (e.g., 'to spill the beans').

You are going to be presented with 36 short passages. The last sentence of each passage is in bold (e.g., 'In the end, **Martha spilled the beans'**). Your task is to put the sentence in bold in your own words, as if you were explaining that expression to a learner of English who didn't understand it."

The questionnaire was posted on-line through the Paid Psychology Experiments website of Princeton University. A log-in system ensured that only Princeton undergraduates could participate. Participants' name and email address were recorded at the start of the questionnaire but their responses were recorded anonymously in real time. That is, the server did not provide a

---

[8] In order to make a controlled comparison between hyperbole and metaphor, it was necessary to construct nominal hyperboles using the same *X is a Y* format that is normally used in metaphor studies. However, the examples and corpora discussed in the literature make clear that hyperbolic uses of adjectives and other noun modifiers (e.g., 'trillions of essays', 'endless questions') are much more common than hyperbolic uses of nouns. It is for this reason that some nominal hyperboles might be perceived as hyperbolic metaphors by some readers. However, insofar as *all* cases of hyperbole are taken as continuous with metaphor on the relevance theoretic view, we do not take these observations to affect the point at issue.

record of which response corresponded with each participant and simply grouped the responses by item as they were entered. Statistical analyses were therefore performed only by items.

Coding of responses

The three authors first coded the responses individually. It was agreed that 'emergent properties' would include those properties which could not apply literally to the figurative vehicle. Reponses that contained words with both a literal as well as a fairly conventionalized figurative interpretation (e.g., 'cold', which can denote either physical coldness or emotional coldness) were coded as emergent if it was clear from the paraphrase that the intended meaning was not applicable to the literal vehicle concept (e.g., for the metaphor 'His embrace was an *icebox*', one such paraphrase was "His embrace seemed cold and no longer intimate"). Properties were coded as 'intra-domain' if they were equally applicable to the literal vehicle concept and to the pragmatically derived concept (e.g., the property of being 'a hugely important occasion' given for the hyperbole 'In our family, everyone's birthday is a *national holiday*'). In order to test that our hyperboles conveyed a greater sense of exaggeration than our metaphors, we also coded the 'intensifiers' used in the paraphrases of these two types of expression; i.e. clear expressions of emphasis, e.g. 'hugely'. Finally, it was agreed that when responses contained more than one property, each property should be coded individually if they were independent of each other[9].

In a second stage, the three codings of the responses were put together, resulting in 91% agreement on the classification of the properties. Those items for which there were coding differences were discussed until agreement was reached.

Comprehensibility

---

[9] This means that each response could contain more than one property, so the mean proportion of properties per satisfactory response could exceed 1.00 in each condition (see Figures 1 and 2).

Only 13 paraphrases revealed either a literal interpretation or an erroneous figurative interpretation (2.7% of data). These erroneous responses were given for a number of different items and were simply discarded from further analyses. Overall, the comprehensibility of the materials was very high, with 97% of paraphrases revealing a satisfactory interpretation of the items.

Results and discussion

The proportion of emergent properties, intra-domain properties and intensifiers was calculated for each item relative to the number of satisfactory responses. The mean proportion of properties of each type is shown in Figure 1.
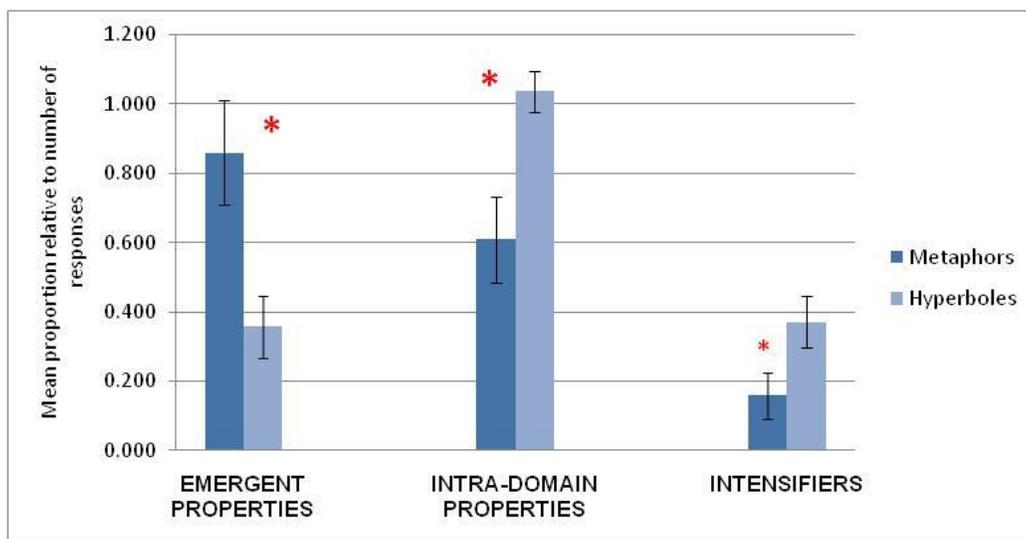
Figure 1: Mean proportion of properties of each type relative to the number of satisfactory responses in the Metaphor and Hyperbole conditions (large asterisk, $p<.05$; small asterisk, $p <.1$).

A 2x3 ANOVA was carried out on the proportion of properties of each type produced for each figure. The analysis revealed a significant main effect of Property type, $F(2, 66) = 16.059$, p

< .001. Crucially, the interaction between Property type and Figure type was significant, $F(2, 66)$ = 12.055, $p < .001$. Pair-wise analyses revealed a significant difference in the proportion of emergent properties produced for metaphors and hyperboles, $t(11) = 2.816$, $p < .018$. This difference was also significant for intra-domain properties, $t(11) = 3.635$, $p < .005$; and marginally significant for intensifiers, $t(11) = 1.839$, $p = .093$.

The results of Experiment 1A support our predictions: participants produced more emergent properties when paraphrasing metaphors than hyperboles, while the opposite pattern was observed for intra-domain properties. Also, the proportion of intensifiers was higher for hyperboles than for metaphors, which confirms that our hyperboles conveyed a stronger sense of exaggeration than the metaphors. These results suggest that there are indeed differences in the kind of interpretation that people give to metaphors and hyperboles and that these differences are consistent with our distinct analyses of hyperbole and metaphor, as outlined in section 2.

However, someone might object to this conclusion on the grounds that we used different metaphorical and hyperbolic vehicles in this experiment. In other words, it is possible that the higher proportion of emergent properties observed in the metaphor condition was not a reflection of differences in interpretation between hyperbole and metaphor, but instead a reflection of some other difference between the specific words chosen as metaphorical and hyperbolic vehicles. In order to rule out this possibility, we ran a second on-line questionnaire using pairs of metaphors and hyperboles that employed the same vehicle word. This manipulation also allowed for a more stringent test of the proportion of intensifiers elicited for each type of expression.

Experiment 1B

Method

Participants

32 undergraduates from Princeton University took part in the study for monetary compensation. They were all native speakers of English.

Materials, design and procedure

6 of the hyperboles from Experiment 1A were modified and used again in Experiment 1B. These nominal hyperboles were selected because the vehicle could also be used in a nominal metaphor. 6 new hyperboles were added to the original ones. 12 new metaphors were constructed using the same vehicles as the hyperboles. That is, paired hyperboles and metaphors had the same vehicle, although their topics were different (e.g., 'Our morning jog is a *marathon*' vs. 'Writing a thesis is a *marathon*'). As in Experiment 1A, all hyperboles and metaphors were of the form *X is a Y* and were preceded by a short context in order to facilitate comprehension (mean number of words per item: 25 for the hyperboles and 29 for the metaphors). The same 12 idioms that were used as fillers in Experiment 1A were used again in Experiment 1B.

The materials were presented in a fixed quasi-random order that ensured that the same vehicle appeared only once in each half of the task, counterbalanced by figure type. The procedure was the same as in Experiment 1A.

Coding of responses

The procedure for the coding of the responses was the same as in Experiment 1A. The initial agreement among the three coders was 87%, although all responses were discussed until agreement was reached. The final classification of the properties was sent to a linguist who was informed of the classification criteria but remained ignorant about the experimental hypothesis. This fourth, 'blind' coder agreed with the final classification of the properties in 96% of cases; again, all responses were discussed until agreement was reached.

Comprehensibility

5.9% of the paraphrases were discarded because they either revealed a literal interpretation of the expression or an erroneous figurative interpretation. As in Experiment 1A, these erroneous responses were given for a number of different items and were simply discarded from further analyses. Overall, the comprehensibility of the materials was very high, with 94% of responses revealing a satisfactory interpretation.

Results and discussion

The proportion of emergent properties, intra-domain properties and intensifiers was calculated for each item relative to the number of satisfactory responses. The mean proportion of properties of each type is shown in Figure 2.
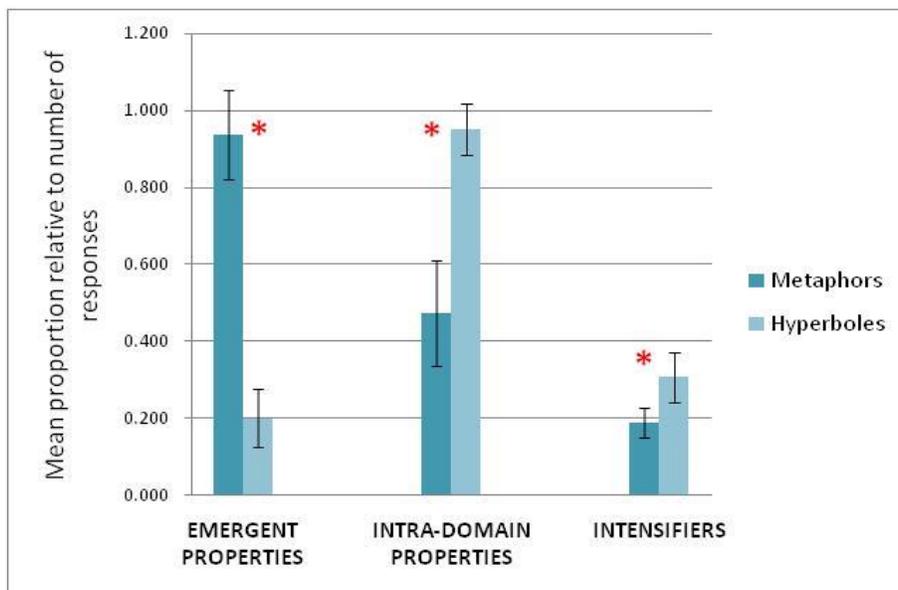
Figure 2: Mean proportion of properties of each type relative to the number of satisfactory responses in the Metaphor and Hyperbole conditions (asterisk, $p < .05$).

A 2x3 ANOVA was carried out on the proportion of properties of each type produced for each figure. The analysis revealed a significant main effect of Property type, $F(2, 66) = 14.252$, $p < .001$. Crucially, the interaction between Figure type and Property type was significant, $F(2, 66) = 24.437$, $p < .001$. Pair-wise analyses comparing metaphors and hyperboles revealed a significant difference in the proportion of emergent properties, $t(11) = 6.286$, $p < .001$; and intra-domain properties, $t(11) = 3.873$, $p < .004$. Importantly, the difference in the proportion of intensifiers was also significant, $t(11) = 2.425$, $p < .035$.

The results of Experiment 1B confirm those observed in Experiment 1A and hence our predictions: participants produced significantly more emergent properties when paraphrasing metaphors than hyperboles, while the opposite pattern was found for intra-domain properties. Also, a higher proportion of intensifiers was elicited by hyperboles than metaphors. We can therefore conclude that there are clear differences in the interpretation of hyperboles and metaphors regarding the type of properties that people attribute to the topic when interpreting these two types of figurative use.

Experiment 2

Our next experiment aimed to further investigate the differences between hyperbole and metaphor interpretation found in Experiments 1A and 1B by using a more naturalistic task; namely, self-paced reading. We predicted that the hyperboles in our materials would be processed more quickly than the metaphors for two reasons. First, the semantic relation between the hyperbole topics and vehicles would result in priming of the vehicles relative to the metaphor condition. Note that the topic-vehicle relation in the hyperbole condition is not an artifact of our materials but a consequence of having to construct nominal hyperboles of the form *X is a Y*,

where Y is an 'exaggeration' of X (e.g., 'Our morning jog was a *marathon*', 'Their lake is an *ocean*', 'The quartet was an *orchestra*')[10]. Second and more interestingly, the differential availability of the properties accessed in hyperbolic and metaphorical interpretations (with the former including more intra-domain properties and the latter including more emergent properties) should result in nominal hyperboles being processed more quickly than nominal metaphors.

In order to confirm that the predicted differences between the hyperboles and the metaphors in our materials are due to differences in both topic-vehicle priming and property accessibility, we also included a literal condition where the topic and the vehicle of the nominal expression were not semantically related (e.g., 'The landmark was an *ocean.*'). We predict that, relative to the literal baseline, the nominal hyperboles will be processed more quickly given the topic-vehicle relation that characterizes such uses. Regarding the metaphor and literal conditions, even though there should be no topic-vehicle priming in either condition, we predict that the metaphors will be processed more slowly than the literal controls given the lower accessibility of emergent properties relative to intra-domain properties. In sum, we predict that our nominal hyperboles will be processed more quickly than the literal controls in our materials, which will in turn be processed more quickly than our nominal metaphors. This pattern is not predicted by the relevance-theoretic account.

Method

Participants

---

[10] More specifically, in instances of nominal hyperbole, X and Y share a salient scalar property and Y is an order of magnitude greater than X on that scale.

69 undergraduates from the University of Kent took part in the experiment for monetary compensation. They were all native speakers of English.

Materials, design and procedure

The same 12 pairs of nominal hyperboles and metaphors that were used in Experiment 1B were used again in Experiment 2, plus 3 new pairs. 15 literal sentences were constructed for the baseline condition. The literal sentences included the critical vehicles but these were not semantically associated to the topic/subject of the sentence. In all cases, care was taken that the preceding contexts did not contain any associates of the vehicle word, with the exception of the last word in each context, which was the topic of the critical sentence in each item and was therefore associated to the vehicle in the hyperbole condition (see below for an example). This was done in order to keep the level of priming constant across conditions. An extra sentence was added to each item, with the first word/phrase of the new sentence being the same across each hyperbole/metaphor/literal triplet. This was done in order to measure any possible spill-over effects following from the critical words. The mean number of words per item was 39 in the hyperbole condition, 42 in the metaphor condition and 49 in the literal condition.

Below is an example of the experimental materials (with the critical sentence in bold and forward slashes marking the three segments that were compared in the three conditions):

Hyperbole. My husband and I decided we should do something about our low level of fitness. Just a couple of weeks ago we decided to take up jogging. **Our morning jog is / a marathon**. / Hopefully, / it won't feel like that for long.

Metaphor.   Mary was intelligent and hard-working but it had taken her five years to write her PhD thesis. **Writing a thesis is / a marathon**. / Hopefully, / her efforts will soon pay off.

Literal.   I've been going to the gym every evening for almost three months. I can really tell I'm getting fitter, but I'm still far from my target. **My target is / a marathon**. / Hopefully, / I'll be ready in another six months.

Each passage was divided into 12 segments (mean number of words per segment, including the experimental segments: 4). The vehicles in the experimental items were always presented in isolation for greater accuracy (e.g., 'a marathon'; mean number of words per critical segment: 2). Reading times were also collected for the segment previous to the critical one (e.g., 'Our morning jog is'; mean number of words per previous-to-critical segment: 3) and for the segment following the critical one (e.g., 'Hopefully'; mean number of words per next-after-critical segment: 3).

While the critical and next-after-critical segments were matched across the three conditions, the segment previous to the critical one was different in each condition. However, the previous-to-critical segments did not differ in number of words across conditions (mean number of words per segment: 3 in all three conditions). Also, the main word in each previous-to-critical segment did not differ in terms of frequency (mean frequency of the main word: 54257 in the hyperbole condition (SD 60303); 51247 in the metaphor condition (SD 86659), and 43337 in the literal condition (SD 58164); the three t-tests revealing p values > .656). In any case, the purpose of comparing reading times for the previous-to-critical and next-after-critical segments across the

three conditions was to see whether any differences observed in the critical segments may be paralleled by differences in the previous or following segments.

The experimental materials were distributed in 3 lists so that each vehicle appeared only once per list and each list contained 5 items per condition. 24 filler items, similar to the experimental materials but not including figurative language, were added to each list.

Participants were allocated to one of the 3 lists in an even, random manner. They were asked to read the passages at a normal pace, making sure they understood the content, and were told that they would have to perform a short memory test at the end of the task in order to make sure that they had paid adequate attention to the passages.

Pre-test of the materials

The materials used in Experiment 2 were pre-tested for familiarity in a questionnaire that was administered to 14 undergraduate students at University College London. All participants took part for monetary compensation. The questionnaire included the complete set of materials (15 hyperboles and 15 metaphors) in a fixed random order. Participants were asked to rate the critical nominal expression in each item (presented in bold in the context) in terms of familiarity using a 1-7 scale ranging from 'Completely unfamiliar (*Never heard it before*)' to 'Highly familiar (*Conventional expression*)'.

2 participants were discarded because English was not their dominant language and their ratings were, on average, 2 points above the overall mean. The perceived familiarity of the items was relatively low (mean: 2.7 in the Hyperbole condition (SD .875) and 2.8 in the Metaphor condition (SD .884)), not revealing a significant difference between the two conditions ($p=.792$). Looking at the analyses per item, mean ratings ranged from 1.6 to 4.4 in the Hyperbole condition

and 1.5 to 4.6 in the Metaphor condition. Overall, 41% of responses were 1s (or the lowest possible rating) while 4% were 7s (or the highest possible rating).

The results of the pre-test confirm, together with the results of Experiment 1B, that the nominal hyperboles and metaphors used in our study were relatively novel and highly comprehensible.

Results and discussion of the self-paced reading task

All participants performed satisfactorily in the memory test, so no data were discarded. Response times that were over 2.5 standard deviations above the mean were discarded and replaced by the corrected mean for that condition. Mean reading times for the critical segments in each condition are displayed in Figure 3.
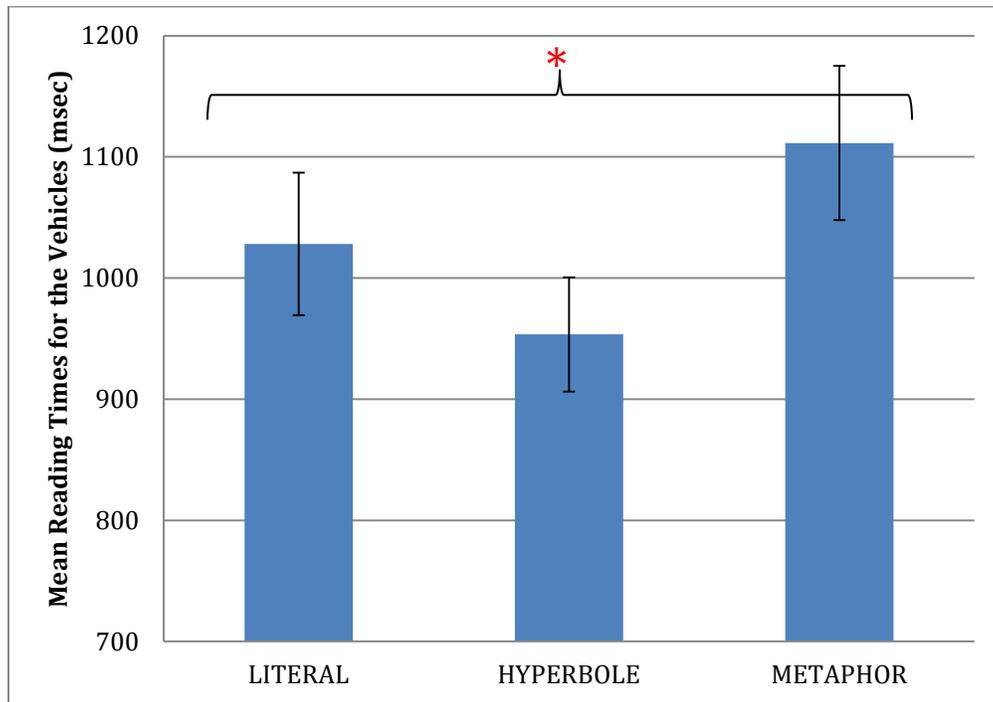
Figure 3: Mean reading times for the vehicles (in milliseconds) in the Literal, Hyperbole and Metaphor conditions (asterisk, $p < .05$).

Mean RTs were entered into a one-way ANOVA. The results revealed a main effect of Sentence type per subjects, $F_1(2, 136) = 8.559$, $p < .001$; and per items, $F_2(2, 28) = 10.455$, $p < .001$. Pair-wise analyses revealed a significant difference between the Hyperbole and Literal conditions, $t_1(68) = 2.264$, $p < .028$; $t_2(14) = 3.010$, $p < .010$. The difference between the Metaphor and Literal conditions was also significant, $t_1(68) = 2.159$, $p < .035$; $t_2(14) = 2.195$, $p < .047$. Thirdly, the difference between the Hyperbole and Metaphor conditions was also significant, $t_1(68) = 3.724$, $p < .001$; $t_2(14) = 4.059$, $p < .002$.

The reading times for the segments previous to the critical ones were comparable in the three conditions (mean reading times for the previous-to-critical segment: 921ms in the Literal condition (SD 403); 895ms in the Hyperbole condition (SD 338), and 925ms in the Metaphor condition (SD 373); all three paired comparisons revealing p values >.228). Therefore, the pattern of results observed for the vehicles was not anticipated in the previous segment.

No spillover effects were found in any of the conditions, with the segment following the vehicles being read at a comparable speed (mean reading times for the next-after-critical segments: 909ms in the Literal condition (SD 373); 908ms in the Hyperbole condition (SD 337), and 924ms in the Metaphor condition (SD 357); all three paired comparisons revealing p values >.420). The lack of spillover effects is not surprising given the relatively long reading times observed for the vehicles. This pattern of results can be interpreted as evidence that participants had already fully interpreted the whole nominal expression when they pressed the key for the target word.

The results of the self-paced reading task support our predictions. First, the hyperboles were processed more quickly than the literal controls. This we anticipated in view of the semantic topic-vehicle relation that characterizes nominal hyperbole, which would result in priming of the hyperbole vehicles relative to the literal vehicles in our materials. Second, the hyperboles were also processed more quickly than the metaphors. This we understand as the result of both priming of the hyperbolic vehicles relative to the metaphorical vehicles, and differences in the accessibility of the predicated properties (with more intra-domain properties being accessed in the interpretation of the hyperboles, while more emergent properties were accessed in understanding the metaphors – see Experiments 1A and 1B). Finally, the literal controls were processed more quickly than the metaphors, which we explain as a result of differences in property accessibility alone (given that there was no semantic relation between topic and vehicle in either type of sentence and so no difference in priming was expected). The results of Experiment 2 therefore support our view that hyperbolic and metaphorical interpretations exhibit differences at the psycholinguistic level. We'll return in Section 4 to the implications of this result for the processing of hyperbole.

Experiment 3

Our last experiment aimed to test for a difference in the relation of the two kinds of figurative interpretations, hyperbolic and metaphorical, to the literal encoded meaning of the figurative vehicle. The difference predicted by our theoretical analysis, as described in section 2, is that hyperbolic interpretations are semantically closer to the literal encoded meaning than metaphorical interpretations are. This is because hyperbolic interpretations involve a simple quantitative shift along one dimension of the literal meaning, while metaphorical meanings

involve a qualitative shift in which a defining property of the literal meaning is dropped and peripheral properties are promoted (what some theorists talk of as a 'domain shift').

In this experiment, we used the materials from Experiment 1B in a sentence verification task in which participants had to indicate whether a sentence was literally true or literally false in the given context. We took our cue here from an early metaphor study using the sentence verification task, in which Glucksberg et al. (1982) showed that people took longer to reject (that is, judge as not literally true) sentences which have a (true) metaphorical meaning (e.g., 'Some roads are snakes') than those which are simply nonsensical (e.g., 'Some roads are desks'). This difference was explained as the result of interference by the metaphorical meaning during the verification task.[11] Our study replicates this use of the sentence verification task but also extends it to sentences with a (true) hyperbolic meaning and compares the effects of the two kinds of non-literal interpretation on participants' task performance.

Our first objective was to check whether sentences with a readily available (true) hyperbolic interpretation would (like their metaphorical counterparts) have an interference effect on participants' judgment of the literal meaning as not true. The figurative meaning of both the metaphors and the hyperboles is made highly accessible in the contexts given in our materials, and moreover, our target sentences did not have an appropriate literal meaning which could compete with the figurative one in the context (see below). We therefore predict that, like metaphors, hyperboles will reveal an interference effect relative to nonsense sentences.

However, our main interest in using the sentence verification task with both metaphors and hyperboles was that it could provide us with a novel way of comparing the interpretation of

---

[11] This was further interpreted as evidence that metaphorical meanings are accessed automatically when available, even when the specified task requires focusing solely on the literal meaning of the expression (Glucksberg et al., 1982; see also Keysar, 1989).

these two types of figurative use. Our characterization of metaphor and hyperbole interpretations focuses on the meaning shift between the lexical concept encoded by the vehicle (e.g., MARATHON) and the ad hoc concept communicated by its hyperbolic or metaphorical use (e.g., MARATHON* in 'Our morning jog is a *marathon*', paraphraseable as something like "long, difficult run"; MARATHON** in 'Writing a thesis is a *marathon*', roughly paraphraseable as "long, strenuous activities of either physical or psychological sorts"[12]). As the paraphrases clearly reflect, the derived (true) meaning in the hyperbolic case is much closer to that of the (false) literal meaning than is the derived (true) meaning in the metaphorical case. We predicted that this difference should have an effect on the time it would take to perform the task of falsifying the literal meaning in each case, with the hyperbolic interpretation causing greater interference than the metaphorical interpretation by virtue of its greater degree of semantic similarity to the literal meaning of the figurative vehicle.

Method

Participants

72 undergraduates from Princeton University took part in the experiment for monetary compensation. They were all native speakers of English.

Materials, design and procedure

The same 12 pairs of nominal hyperboles and metaphors that were used in Experiment 1B and Experiment 2 were used again in Experiment 3. 12 nonsense sentences of the form *X is a Y* were constructed by scrambling the vehicle words of 6 of the hyperbole sentences and 6 of the

---

[12] The paraphrases of pragmatically inferred (ad hoc) concepts expressed by metaphorical uses are inevitably rough and inadequate because metaphor is usually employed when there just is no available means of literally verbalising the concept or idea the speaker wants to communicate.

metaphor sentences. We used the nonsense sentences as our baseline since the correct response to all experimental items had to be the same (i.e., 'Not literally true'). Again, care was taken to keep the level of priming constant and down to a minimum across conditions. Below is an example of the experimental materials:

Hyperbole. My husband and I decided we should do something about our low level of fitness. Just a couple of weeks ago we decided to take up jogging. **Our morning jog is a marathon**.

Metaphor. Mary was intelligent and hard-working but it had taken her five years to write her PhD thesis. **Writing a thesis is a marathon**.

Nonsense. Since he retired, George only does things out of habit and refuses to try anything new. He definitely has the wrong attitude. **His attitude is a marathon**.

The experimental materials were distributed in 3 lists so that each vehicle appeared only once per list and each list contained 4 items per condition. 24 filler items, similar to the experimental materials but not including figurative language were added to each list. The correct response was 'Literally true' for 18 of the fillers. Overall, the correct response was 'Not literally true' for half of the items in each list and 'Literally true' for the other half.

Participants were allocated to one of the 3 lists in an even, random manner. They were told that our study investigated how people interpret different writing styles. They were going to be presented with a series of passages and they had to indicate whether the last sentence of each passage was literally true in the context. To do this, they had to press the designated YES or NO key (i.e., 'Yes, literally true' or 'No, not literally true') using the index and middle fingers of their dominant hand.

Each item was presented in two parts. First, the short context that preceded each critical sentence was presented in isolation and participants could read it at their own pace. Then they had to press the space bar and the final sentence was displayed in bold in the centre of the screen. Participants were asked to verify the final sentence as quickly and accurately as possible.

Results and discussion

8 participants were eliminated from the statistical analyses because subsequent checking showed that they had taken part in the on-line questionnaire from Experiment 1B and were therefore familiar with the materials. Incorrect responses and response times that were over 2.5 standard deviations above the mean were discarded and replaced by the corrected mean for that condition. Mean (correct) response times for each condition are displayed in Figure 4.
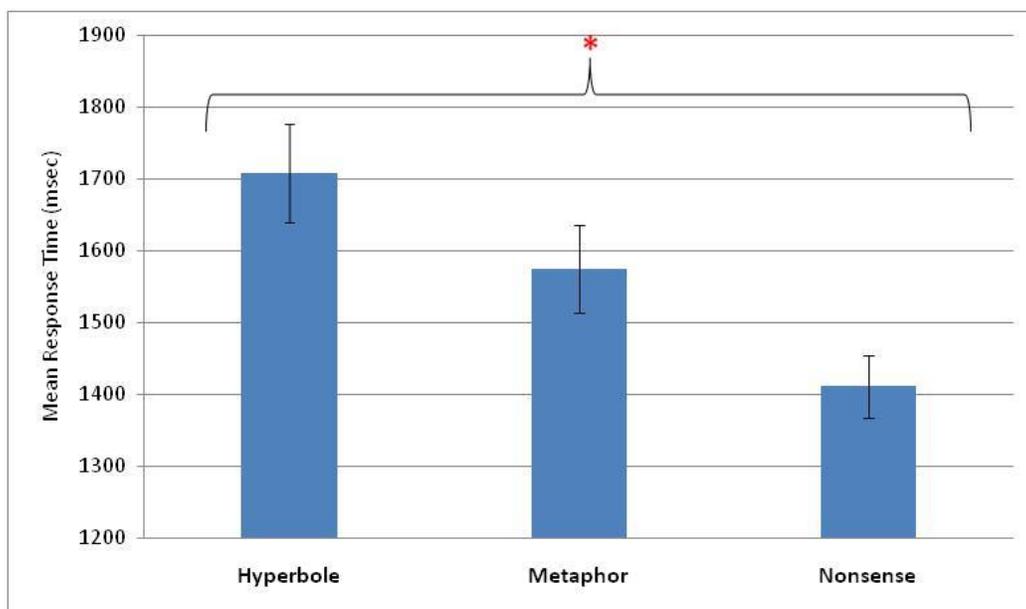
Figure 4: Mean sentence-verification time (in milliseconds) for the Hyperbole, Metaphor and Nonsense conditions. The correct response was 'No' (not literally true) in the three experimental conditions (asterisk, $p < .05$).

Mean RTs were entered into a one-way ANOVA. The results revealed a main effect of Sentence type per subjects, $F_1(2, 128) = 19.449$, $p < .001$; and per items, $F_2(2, 22) = 6.376$, $p < .008$. Pair-wise analyses revealed a significant difference between the Hyperbole and Nonsense conditions both by participants, $t_1(64) = 6.243$, $p < .001$; and by items, $t_2(11) = 4.280$, $p < .002$. The difference between the Metaphor and Nonsense conditions was significant by participants, $t_1(64) = 3.430$, $p < .002$; and marginally significant by items, $t_2(11) = 1.951$, $p < .078$. Thirdly, the difference between the Hyperbole and Metaphor conditions was only significant by participants, $t_1(64) = 2.788$, $p < .008$; $t_2(11) = 1.432$, $p = .180$.

We explain the weaker results observed in the analyses per items as an effect of the relatively low number of items (12 per condition). To further test for significance, we used a linear mixed-effects model in R (R Development Core Team 2012), positing a fixed effect of Condition and random effects of Item and Participant, with full random slopes. Model comparison revealed that the effect of Condition is highly significant ($p < 0.001$). In order to explore this effect further, we restricted our attention to the Hyperbole and Metaphor conditions and applied the same procedure. Model comparison disclosed a significant effect of condition ($p = 0.012$).

The pattern of error rates parallels that of response times: the proportion of incorrect responses in the Literal condition was .02, while it was .10 in the Hyperbole condition, and .14 in the Metaphor condition. A chi-square test with Yate's correction revealed a reliable difference between the Literal and Hyperbole conditions, χ^2 (1, N = 576) = 15.558, $p < .001$, and between the Literal and the Metaphor conditions, χ^2 (1, N = 576) = 28.942, $p < .001$. However, the difference between the two critical conditions did not reach significance, χ^2 (1, N = 576) = 2.371,

p = .124. The difference observed between the control and the critical conditions is explained by the fact that the hyperboles and the metaphors, but not the nonsense sentences, had a figurative meaning that was true in the context. That the critical conditions did not differ from one another (unlike the verification times) may be related to the overall low proportion of erroneous responses (mean .09).

The overall pattern of results of Experiment 3 supports our hypotheses. Both hyperboles and metaphors took significantly longer to be judged as not literally true than nonsense sentences, thus extending the original findings by Glucksberg et al. (1982). More importantly for our investigation, there was also a significant difference between the two critical conditions, with participants taking longer to falsify hyperboles than metaphors. This pattern of results is exactly as predicted by our theoretical analysis in terms of degree of semantic closeness between the literal and the figurative meaning of hyperboles and metaphors. In sum, the results of the sentence-verification task suggest that there are clear interpretive differences between hyperbole and metaphor.

4.   General discussion

Our study aimed to investigate the strong continuity thesis defended by Sperber and Wilson (2008), specifically concerning the interpretation of hyperboles and metaphors. In our view, there are important descriptive and psycholinguistic differences between the interpretation of hyperbolic and metaphorical uses of language, even while both recruit the same general mechanisms of concept adjustment for their interpretation (Carston and Wearing 2011, forthcoming).

Experiments 1A and 1B provided support for our position, showing that people access representations of different types of properties when they interpret hyperboles and metaphors. We argued above that the type of properties that interpreters understand as attributed to the topic of nominal hyperboles and metaphors reflects the kind of meaning shift – quantitative and qualitative, respectively – that is achieved by these distinct non-literal uses, with hyperbolic interpretations therefore containing relatively more intra-domain properties and metaphorical interpretations containing relatively more emergent properties. The experimental results revealed that these predictions were borne out by the paraphrases the participants produced.[13]

Experiment 2 provided further support for our view, suggesting that differences in the accessibility of intra-domain and emergent properties, as well as differences in topic-vehicle relation, had an effect on the speed with which the nominal hyperboles and nominal metaphors were processed. Thus, the nominal hyperboles in our materials were read more quickly than both the literal controls and the nominal metaphors. The difference between the hyperboles and the literal controls can be explained as an effect of semantic priming, given the topic-vehicle relation that characterizes nominal hyperbole and the lack of such a relation in our literal condition. However, the critical difference between the nominal hyperboles and the nominal metaphors is the result of *both* priming and differences in property accessibility (with intra-domain properties being more readily accessible than emergent properties). That our results reflect differences not only in topic-vehicle relations but also in property accessibility is evident from the difference that we observed between the literal and metaphor conditions. The faster processing of the literal

---

[13] The results of Experiment 1a/1b are broadly compatible with several other accounts of metaphor, such as conceptual metaphor theory (Lakoff & Johnson 1980, Gibbs 1994) and conceptual blending theory (Fauconnier and Turner 2002). However, it is less clear what other views would predict about hyperbole because (as we noted in the introduction) hyperbole is significantly less discussed in the literature. There is some exploration of similarities between hyperbole and irony (Colston & O'Brien 2000), but as Colston and Gibbs (2002) caution, it is difficult to extrapolate from one figure to another. We are thus reluctant to speculate in advance of more detailed discussions.

controls relative to the nominal metaphors could not result from priming of the literal vehicles as there was no topic-vehicle relation in either condition. We can therefore conclude that there are interesting psycholinguistic differences between nominal hyperbole and metaphor.

Experiment 3 also confirmed our claim that there are important differences in the nature of the interpretations given to hyperbolic uses and to metaphorical uses, with participants taking longer to judge hyperboles than metaphors as literally false. The first point to make about the results of this experiment is that they show that hyperboles as well as metaphors take longer to falsify than nonsense statements, thus corroborating and extending earlier studies that demonstrated this interference effect for metaphors (Glucksberg et al., 1982; Keysar, 1989).

However, the more important result of Experiment 3 is that it took participants significantly longer to reject the literal interpretation in cases of hyperbole than in cases of metaphor. In our view, the key to this difference lies in the relationship between the figurative and literal interpretations in each case. We maintain that hyperbolic interpretations are the result of relaxing a (single) scalar property of the literal vehicle. Focusing on the relation between the resulting interpretation and the literal interpretation, the difference in content between them consists in this single dimension of adjustment; other key (associated) features of the vehicle concept remain untouched/still applicable. Metaphorical interpretations, by contrast, cannot be produced in this unidimensional/quantitative way; at the very least, something beyond a merely quantitative adjustment has to be made, and this (qualitative) adjustment has considerable follow-on effects with respect to other (associated) features of the vehicle concept. In short, there is a significantly greater degree of 'overlap' or semantic similarity between the literal meaning of the vehicle and its figurative interpretation in hyperbolic uses than in metaphorical ones. It is this greater similarity between the (true) hyperbolic interpretation and the (false) literal meaning than

between the (true) metaphorical interpretation and the (false) literal meaning that accounts for the greater time taken to judge the literal meaning as false in the hyperbole case than in the metaphor case.

It is important to note that, while the difference in topic-vehicle relation between nominal hyperboles and metaphors has an effect on the processing of these two types of figurative expression (as shown by the results of the self-paced reading task in Experiment 2), it cannot explain the results of the sentence verification task in Experiment 3. That is, the fact that the hyperbole vehicles would have been primed by the topics and so their figurative interpretations would have been available earlier in processing does not explain the greater interference that these interpretations exerted on the sentence verification task. If anything, the earlier availability of the figurative interpretation of the hyperboles should have allowed participants to switch their attention more rapidly to the literal interpretation and reject it as false. In our view, it is the semantic closeness of the figurative and literal interpretations of the hyperboles (prompting competing responses: true/ false) that explains the greater interference of the hyperboles than the more semantically distant metaphors.

To sum up, it is clear that there are differences between hyperbole and metaphor of both descriptive and psycholinguistic significance. This shows that the strong continuity thesis is *too* strong. But what of the remaining claim, that there are no significant *pragmatic* differences between hyperbole and metaphor? As Sperber and Wilson emphasize, an important element of the relevance-theoretic account is the claim that a single inferential procedure underlies the interpretation of both metaphorical and hyperbolic cases (2008: 95). It should be clear that this weaker continuity thesis is untouched by our results – the sorts of interpretive differences we have found in the present study are compatible with there being a single underlying processing

mechanism operating in both hyperbolic and metaphorical cases. What our results contradict is the strong claim that there are no fine-grained differences in the way that the mechanism applies to cases of hyperbole and metaphor. Instead, we have shown that the sorts of interpretations the hearer generates by means of that mechanism are distinctly different in the two cases. Similarly, her ability to judge the status of the literal interpretation (its truth/falsity) varies with the sort of figurative interpretation she has constructed. The continuity of hyperbole and metaphor, then, should be understood as a claim within a pragmatic theory about the mechanism underwriting the interpretive process across a range of cases. As such, it is compatible with the differences that we have found, and more generally, with the widespread pre-theoretical intuition that hyperbole and metaphor are strikingly different figures.

Acknowledgements

REFERENCES

Almor, A., Arunachalam, S. & Strickland, B. 2007. When the 'creampuff' beat the boxer: Working memory, cost, and function in reading metaphoric reference. *Metaphor and Symbol, 22,* 169-193.

Blasko, D. & Connine, C. 1993. Effects of familiarity and aptness on metaphor processing, *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*, 295-308.

Bowdle, B. & Gentner, D. 2005. The career of metaphor. *Psychological Review, 112*, 193-216.

Cano-Mora, L. 2009. All or nothing: A semantic analysis of hyperbole. *Revista de Lingüística y Lenguas Aplicadas, 4*, 25-35.

Carston, R. 1996. Enrichment and loosening: Complementary processes in deriving the proposition expressed? *UCL Working Papers in Linguistics*, 8, 61-88. Reprinted in 1997 in *Linguistische Berichte*, *8,* 103-127.

Carston, R. 2002. *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.

Carston, R. & Wearing, C. 2011. Metaphor, hyperbole and simile: A pragmatic approach. *Language and Cognition, 3*, 283–312.

Carston, R. & Wearing, C. forthcoming. Hyperbolic language and its relation to metaphor and irony. .

Claridge, C. 2011. *Hyperbole in English: A Corpus-Based Study of Exaggeration*. Cambridge University Press.

Clark, H. 1997. *Using Language*. Cambridge University Press.

Clausner, T. & Croft, W. 1997. The productivity and schematicity of metaphor. *Cognitive Science,* 21, 247-82.

Colston, H. L. 1997. "I've never seen anything like it": Overstatement, understatement and irony. *Metaphor & Symbol,* 12, 43-58.

Colston, H.L., & Gibbs, R.W. 2002. Are irony and metaphor understood differently?. *Metaphor and Symbol.* 17.1:57-80.

Colston, H. L., & Keller, S. B. 1998. You'll never believe this: Irony and hyperbole in expressing surprise. *Journal of Psycholinguistic Research*, *27*, 499-513.

Colston, H. L. & O'Brien, J. 2000. Contrast of kind versus contrast of magnitude: The pragmatic accomplishments of irony and hyperbole. *Discourse Processes*, *30*, 179-199.

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

Deamer, F., Pouscoulous, N. & Breheny, R. 2010. A contrastive look at metaphor and hyperbole. *UCL Working Papers in Linguistics, 22,* 1-15.

Fauconnier, G. & Turner, M. 1998. Conceptual integration networks. *Cognitive Science, 21,* 133-187.

Fauconnier, G. & Turner, M. 2002. *The Way We Think*. Basic Books.

Filippova, E. & Astington, J.W. 2010. Children's understanding of social-cognitive and social-communicative aspects of discourse irony. *Child Development,* 81 (3), 913-928.

Fogelin, R. 1988. *Figuratively Speaking*. Yale University Press.

Gernsbacher, M. A., Keysar, B., Robertson, R. W. & Werner, N. K. 2001. The role of suppression and enhancement in understanding metaphors. *Journal of Memory and Language*, *45*, 433-450.

Gibbs, R. W. 1994. *The Poetics of Mind: Figurative Thought, Language and Understanding*. Cambridge University Press.

Gibbs, R. W. 2000. Irony in talk among friends. *Metaphor and Symbol* 15: 5-27. Reprinted in R. W. Gibbs and H. Colston (eds.) 2007. *Irony in Language and Thought*, 339-360. Lawrence Erlbaum Associates.

Gibbs, R. W. 2006. Metaphor interpretation as embodied simulation. *Mind and Language* 21: 434-458.

Gibbs, R. W. & Gerrig, R. J. 1989. How context makes metaphor comprehension seem "special". *Metaphor and Symbolic Activity, 4,* 145-158.

Gildea, P. & Glucksberg, S. 1983. On understanding metaphor: The role of context. *Journal of Verbal Learning and Verbal Behaviour*, *22*, 577-590.

Giora, R. 1997. Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics, 7,* 183–206.

Giora, R., Gazal, O. & Goldstein, I. 2012. Salience and context: Interpretation of metaphorical and literal language by young adults diagnosed with Asperger's Syndrome. *Metaphor and Symbol*, *27*, 22–54.

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

Glucksberg, S. & Haught, C. 2006. On the relation between metaphor and simile: When comparison fails. *Mind and Language, 21*, 360-378.

Grady, J., Oakley, T. & Coulson, S. 2000. Blending and metaphor. In G. Steen & R. Gibbs (eds.) *Metaphor in Cognitive Linguistics,* 101-124. Philadelphia: John Benjamins.

Grice, H. P. & Strawson, P. F. 1956. In defense of a dogma. *The Philosophical Review*, 65.2, 141-58.

Hussey, K. & Katz, A. N. 2009. Perception of the use of metaphor by an interlocutor in discourse. *Metaphor and Symbol, 24*, 203-236.

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

Kintsch, W. 2000. Metaphor comprehension: A computational theory. *Psychonomic Bulletin & Review, 7*, 257-266.

Kövescses, Z. 2002. *Metaphor: A Practical Introduction.* Oxford University Press.

Kreuz, R. J., Kassler, M. A. & Coppenrath, L. 1998. The use of hyperbole in discourse: Cognitive and social facets. In S. R. Fussell & R. J. Kreuz (eds.), *Social and cognitive approaches to interpersonal communication*. Mahwah, NJ: Lawrence Erlbaum Associates.

Kreuz, R. J. & Roberts, R. M. 1995. Two cues for verbal irony: Hyperbole and the ironic tone of voice. *Metaphor and Symbolic Activity, 10*, 21-31.

Kreuz, R. J., Roberts, R. M., Johnson, B. K. & Bertus, E. L. 1996. Figurative language occurrence and co-occurrence in contemporary literature. In R. J. Kreuz & M. S. MacNealy (Eds.), *Empirical Approaches to Literature and Aesthetics*. Norwood, NJ: Ablex Publishing Corp.

Lakoff, G. & Johnson, M. 1980. *Metaphors We Live By*. University of Chicago Press.

Lakoff, G. 1993. The contemporary theory of metaphor. In: A. Ortony (ed.) *Metaphor and Thought*, 202-51. Cambridge University Press.

McCarthy, M. & Carter, R. 2004. 'There's millions of them': Hyperbole in everyday conversation. *Journal of Pragmatics, 36*, 149–184.

Noveck, I. A., Bianco, M. & Castry, A. 2001. The costs and benefits of metaphor. *Metaphor and Symbol*, *16*, 109-121.

Searle, J. 1980. The background of meaning. In: J. Searle, F. Keifer & M. Bierwisch (eds.) *Speech Act Theory and Pragmatics*, 221-232. Dordrecht: Reidel.

Sperber, D. & Wilson, D. 1986. Loose talk. *Proceedings of the Aristotelian Society* 86, 153-171. Reprinted in S. Davis (ed.) 1991, *Pragmatics: A Reader*, 540-549. Oxford: Oxford University Press.

Sperber, D. & Wilson, D. 1986/95. *Relevance: Communication and Cognition.* London and New York: Blackwell. (Second edition with new postface 1995.)

Sperber, D. & Wilson, D. 2008. A deflationary account of metaphors. In: R. Gibbs (ed.), *The Cambridge Handbook of Metaphor and Thought*. Cambridge: Cambridge University Press.

Steen, G. 2008. The paradox of metaphor: Why we need a three-dimensional model for metaphor. *Metaphor and Symbol 23 (4),* 213-241.

Steen, G.J., Dorst, A.G., Herrmann, J.B., Kaal, A.A., Krennmayr, T., Pasma, T. 2010. *A Method for Linguistic Metaphor Identification: From MIP to MIPVU*. Amsterdam/Philadelphia: John Benjamins.

Tendahl, M. & Gibbs, R.W. 2008. Complementary perspectives on metaphor: Cognitive linguistics and Relevance Theory. *Journal of Pragmatics, 40,* 1823-1864.

Vega-Moreno, R. 2007. *Creativity and convention: The pragmatics of everyday figurative speech.* Amsterdam: John Benjamins.

Wearing, C. 2014. Interpreting novel metaphors. *International Review of Pragmatics. 6,* 78-102.

Wilson, D. 2003. Relevance Theory and lexical pragmatics. *Italian Journal of Linguistics/ Rivista di Linguistica*, *15,* 273-291.

Wilson, D. & Carston, R. 2006. Metaphor, relevance and the 'emergent property' issue. *Mind and Language, 21*, 404-433

Wilson, D. & Carston, R. 2007. A unitary approach to lexical pragmatics: Relevance, inference and ad hoc concepts. In: N. Burton-Roberts (ed.). *Pragmatics*, 230-259. Basingstoke and New York: Palgrave Macmillan.

Wilson, D. & Sperber, D. 1992. On verbal irony. *Lingua* 87, 53-76.

Wilson, D. & Sperber, D. 2012. Explaining irony. In: D. Wilson & D. Sperber *Meaning and Relevance*, 123-145. Cambridge: Cambridge University Press.

Wolff, P. & Gentner, D. 2011. Structure-mapping in metaphor comprehension. *Cognitive Science, 35*, 1456–1488.