

John Benjamins Publishing Company



This is a contribution from *Metaphor and Communication*.

Edited by Elisabetta Gola and Francesca Ervas.

© 2016. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Tables of Contents, abstracts and guidelines are available at www.benjamins.com

Metaphor and simile

Categorizing and comparing categorization and comparison

John Barnden

University of Birmingham

Do understanders use comparison or categorization in understanding simile and/or metaphor? I claim that the intense debate about this has been misguided, partly because of the obscurity of the distinction between comparison and categorization, and partly because of a frequent, misleading characterization of comparison theory as casting (a type of) metaphor as elliptical simile. Instead, it is more fruitful to consider a theoretical space defined by various processing dimensions that cut across comparison and categorization. I propose three dimensions here: Target/Source Contribution Disparity, Target/Source Mediator-Preservation Degree, and Target/Source Mediator-Carefulness. Experimental results adduced in the debate are not necessarily resolving whether categorization or comparison is being used but rather where the understanding process is positioned on the dimensions. Comparison and categorization can both supply whatever positioning is required. In addition, where a particular act of metaphor/simile understanding is placed on the dimensions is not purely a matter of linguistic form and general cognitive considerations, but can also be importantly affected by the communicative context and matters such as familiarity that depend on communication history.

Keywords: categorization, comparison, simile, metaphor

1. Introduction

In recent years there has been a debate about whether understanders use processes of *comparison* on the one hand or *categorization* on the other for the purpose of understanding simile and/or metaphor (see, e.g., Bowdle & Gentner 2005; Glucksberg 2001, 2008). In this article I will call this debate ‘the C/C debate’ or just ‘the debate’ for short. It has centred on similes of the form ‘T is/are like S’

and corresponding metaphors of the form ‘T is/are S,’ which we will call *be-form* metaphors in this chapter. T is the *target* term and S is the *source* term. (They will always be noun phrases in our discussion.) One purpose of this article is to claim that the debate has been severely corrupted by three main factors:

1. neglect of some slipperiness in the notions of ‘categorization’ and ‘comparison,’ and hence in the distinction between them;
2. unwarranted assumptions about categorization and comparison;
3. an ellipsis-based mischaracterization of the general notion of comparison theory: basically, mischaracterizing comparison theory as necessarily saying that a *be-form* metaphor is an elliptical form of the corresponding simile.

A second purpose is to go on from the analysis of those factors to make a particular conjecture about the C/C debate. Put in its starkest form this conjecture is that what is actually revealed by experiments involved in the debate is not whether a process of comparison or categorization is involved in understanding, but, rather, other processing differences: namely, differences between positions on various underlying dimensions. These dimensions include the following three:

1. *Target/Source Contribution Disparity*: qualitative and quantitative differences in the way that the source and target sides (S and T) in cases of simile or metaphor contribute to the understanding process.
2. *Target/Source Mediator-Preservation Degree*: roughly speaking, the extent to which the relationship between source and target remains activated in the understander’s mind along with the information gleaned about the target item itself, rather than merely serving as a stepping stone towards that information.
3. *Target/Source Mediator-Carefulness*: the amount of care taken by the understanding process to establish the relationship between source and target.

Movement along these dimensions is possible both within a given theory of simile/metaphor understanding and between theories. Within a theory: a given theory can propose that the degree of contribution disparity, mediator preservation or mediator carefulness depends on the particular T and S involved in the simile or metaphor, on the form of the utterance (e.g., simile form versus *be-form* metaphor), and/or on contextual factors including the communicative purpose of the utterance and surrounding ones. Between theories: for a given T and S, a given utterance form and a given context, different theories can propose different degrees of contribution disparity, mediator preservation or mediator carefulness.

Crucially, I will suggest that it is different positions along the above dimensions that can account for at least some salient experimental results in the debate, irrespective of whether comparison or categorization is involved. The dimensions

cut across comparison and categorization: i.e., both comparison and categorization accounts can fall on many, and perhaps any, positions on each of the dimensions, depending on the details of the account. Thus, in seeking to explain experimental results, the distinction between categorization and comparison is relatively powerless, compared to distinctions between positions on the above dimensions. Instead of trying to account for results directly in terms of comparison and categorization, it would be more fruitful to account for them in terms of the above dimensions, and then, if one wishes, try to frame detailed comparison or categorization theories (or hybrid theories) that provide the desirable positioning along the dimensions. But the suspicion is that it will always be possible to find *some* comparison theory and *some* categorization theory to explain the result if one tries hard enough. My observations add to the points by Kennedy & Chiappe (1999) and Chiappe & Kennedy (2001) that the ultimate issue within the Debate may not be comparison versus categorization.

Section 2 outlines the nature of the C/C debate, and points out that the distinction between categorization and comparison is murky and ill-defined. Section 3 addresses a different type of obfuscation that has occurred in the debate, concerned with the notion that metaphor is elliptical simile. Section 4 discusses the three dimensions listed above, indicating briefly how they relate to experimental effects studied in the debate, and showing that this relationship may be a more fruitful thing to study than the relationship of the comparison/categorization distinction to the experiments. Section 5 sketches how the proposed dimensions provide a useful analysis of comparison and categorization processes that have been proposed in a particular cognitive modeling framework (Utsumi 2011). Section 6 concludes, and in particular summarizes how the article engages with the communicative dimension highlighted in this volume's Introduction.

The ideas in this article came indirectly out of an artificial intelligence project on metaphor understanding (see, e.g., Barnden 2001, 2008, 2015).

2. The Comparison/Categorization debate

The C/C debate is basically about how similes such as 'Businesses are like dictatorships' and corresponding be-form metaphors such as 'Businesses are dictatorships' are interpreted by ordinary understanders. The main battle is about be-form metaphor rather than simile, and concerns the question of whether categorization or comparison is the mental processing mechanism involved in understanding be-form metaphor (or particular subclasses of it, such as conventional cases or novel cases). For simile, on the other hand, there appears to be general agreement that comparison is the mechanism in play.

Many psychological experiments have been done to provide evidence for one account or another (see, e.g., Bowdle & Gentner 2005; Chiappe & Kennedy 2001; Chiappe, Kennedy & Smykowski 2003; Gokcesu 2009; Jones & Estes 2006; Glucksberg 2008; Glucksberg & Haught 2006; Johnson 1996; Pierce & Chiappe 2009; Utsumi 2007; Xu 2010). The experimental landscape is complex, both because of the number of different experimental measures used and because there are several major theory variants, including hybrids of theories. The measures and the ways of forming hybrid theories are outlined below.

It does not matter for the purposes of this chapter whether or not simile should be viewed as a form of metaphor. The contrast at issue in this chapter can be framed either as a contrast between two different forms of metaphor (be-form form and simile form) or as a contrast between one figure of speech – (be-form) metaphor – and another figure, namely simile. Notice that the C/C debate does overwhelmingly address just be-form metaphor as opposed to other syntactic manifestations of metaphor, although debaters do sometimes address the latter (e.g., Glucksberg 2001, 48–49).

Also, in common with most discussion within the debate I will not address any possible systematic differences between variants of simile form or between variants of be-form. For instance, I will assume that, as far as the debate is concerned, the three sentences ‘A business is a dictatorship’, ‘Mike’s business is a dictatorship’ and ‘Businesses are dictatorships’ are understood using the same main processes as each other; and the same goes for the simile versions of those be-form metaphors.

I now summarize the nature of categorization and comparison accounts. According to a categorization account of the understanding of ‘T is/are S’, a suitable superordinate category S^* of the category literally meant by S is accessed/constructed such that the literal S category is included in S^* , and the sentence is understood as asserting that the item or category meant by T is included in S^* . For example, for ‘Businesses are dictatorships’, the understander might take *dictatorship*^{*} to be the category S^* of organizations/communities that are managed non-consensually and punitively by one person. This is superordinate to (i.e., a super-category of) the category of literal, political dictatorships denoted by S; and the metaphor is taken by the understander as asserting that businesses are within S^* as opposed to being within the literal S category.

From now on I will use the symbol S for the literal S category as well as for the expression S itself, to save on notation. Similarly for T, though in this case T may be an individual entity such as a particular person’s business rather than a category such as businesses in general.

Various names other than ‘categorization account’ have been used for this type of theory, both by their proponents and others. Alternative names are ‘class-inclusion account’ and ‘interactive property-attribution account’ (see Glucksberg

2001 for the last). We can also approximately include Relevance Theory accounts (Sperber & Wilson 2008) because of their appeal to ‘broadening’ of S to a broader category S*, but some RT accounts also involve ‘narrowing’ of S* to get a final category S** in which to place T (Carston 2011; Sperber & Wilson 2008).

In comparison accounts of ‘T is/are S’, the concepts literally meant by the expressions T and S are compared. I will use the symbols T and S for those concepts as well as for the expressions. As a result of the comparison, particular similarities (and possibly differences) are focussed upon. It is typically proposed that this is done through structural analogy finding, as in Structure Matching Theory (SMT: Gentner 1983) or in ACME (Holyoak & Thagard 1989). It is also typical to propose that some of the excess information about S – i.e., information about S that is not found by the matching process to correspond to something about T – is under certain conditions provisionally stipulated to apply to T (perhaps in modified form). These stipulations are for instance the ‘candidate inferences’ provided by the Structure Matching Engine (SME: Falkenhainer, Forbus & Gentner 1989) based on SMT.

Also, similarities found between T and S may instigate the mental construction of a common abstraction S* covering both T and S. Indeed, an important, hybrid, variant of the comparison approach that is salient in the debate is the *Career of Metaphor* theory (Bowdle & Gentner 2005), CoM. This says that as the use of expression S as a metaphor source item becomes entrenched, one or more remembered common abstractions S* arising from such uses of S (possibly with different Ts) become standardly associated with S as superordinate categories/concepts, and become conventional metaphorical meanings of S. These can then be used much as S* categories in a categorization account are, even with new Ts.

Thus, in essence, the *Career of Metaphor* account supplements comparison with categorization. But, dually, categorization theory has elsewhere been supplemented with comparison. For example, Glucksberg (2008) has stated that it is possible that, under some conditions, comparison between the literal T and S in a be-form metaphor is done. It is plausible to suggest that this happens when no S* relevant to T is yet available or readily abstractable from S. So, both some comparison theorists and some categorization theorists now allow that either process might happen in be-form metaphor interpretation: the real argument is then about the conditions under which they occur (Glucksberg 2008).

Surprisingly, however, the mental processes of comparison and categorization have been poorly distinguished within the debate. Comparison theorists, notably those adhering to a particular detailed theory of analogy-finding such as SMT or ACME, often do a good job of describing comparison. However, typically neither categorization nor comparison theorists are specific about what categorization or categories are. Usually, in the debate, categorization is just presented as being obviously different from comparison.

And, there is a major problem in the fact that under certain important views of categorization, much of it just boils down to comparison anyway. Bowdle & Gentner (2005) pointed out, within the debate literature itself, the importance of prototype-based and exemplar-based psychological theories of categorization. For presentation or discussion of such theories, see Connolly *et al.* (2007), Hampton (2007), Rein *et al.* (2010), Rosch (1975), Verbeemen *et al.* (2007) and Wills *et al.* (2006). Such theories inherently involve comparison, because, for instance, to determine whether a particular item is in the category it must be compared to the prototype of the category or to exemplars of it; and indeed we can add that the process of comparison here could be much as it is in comparison theories of simile/metaphor. However, categorization theorists have largely not taken on board the disruptive consequences of these points.

Finally, experimental researchers within the debate have studied what I will call 'effects' on experimental participants such as:

1. Form preference: whether participants show a preference for be-form-metaphor form over simile form or vice versa, for a particular T and S.
2. Relative ease and/or speed of understanding of the two forms, for a given T and S.
3. Relative accuracy of understanding of the two forms, for a given T and S.
4. Differences between the features of a given T and/or S that are accessed by participants, according to which form is used.
5. The relative diversity of interpretations that participants produce for the two forms.
6. What happens to the above effects when the metaphor or simile is 'reversed' (i.e., T and S are interchanged).

Some of these will play a role below.

3. Ellipsis-based mischaracterization of Comparison theory

The difficulty we have noted of adequately describing comparison and categorization and distinguishing between them is itself enough to make one want to look elsewhere in seeking to account for experimental results adduced in the C/C debate. Before I consider what that elsewhere might be, I turn to another way in which the debate has been obfuscated.

A commonly held view – which is, however, sometimes only tacitly or vaguely conveyed – is that a comparison theory necessarily treats be-form metaphors as if they were the corresponding similes. Under such a view of comparison theories, a theory treats 'Businesses are dictatorships' as if it were the corresponding simile

'Businesses are like dictatorships' (or at least this happens under suitable conditions, such as 'dictatorship' not having a relevant entrenched metaphorical meaning). That is, intuitively, the comparison theory treats the metaphor as an elliptical form of the simile. Miller (1979, 381) says that 'understanding [a be-form] metaphor requires the reinsertion of 'is like...'. Glucksberg (2001) says that according to comparison theory metaphors are essentially 'implicit similes'. Chiappe, Kennedy and Smykowski (2003) assume that comparison theory treats metaphor as elliptical simile. Fogelin (2011) puts forward an elliptical-simile view of metaphor, though he is cautious in that he feels that instead of thinking explicitly of metaphor as elliptical simile it is clearer and more productive to claim that a be-form metaphor means figuratively what the corresponding simile means figuratively.

Now, an ellipsis-based view of comparison theory naturally leads to assuming that *when a comparison process is involved in understanding 'T is/are S,' the process is the same as the comparison process involved in understanding 'T is/are like S.'* This assumption has been used in the C/C debate as a weapon against comparison theories of be-form metaphor. This is because some experimental results show that a be-form metaphor can have *different* effects from the corresponding simile (with 'effect' being as at the end of Section 2): so, the argument goes, the metaphor cannot involve comparison.

However, the debate has appeared to overlook a simple but crucial observation, *viz:* the mere fact that two different linguistic forms are both understood through comparison in no way implies that the two comparison processes are the same or even particularly similar to each other. Thus, it is perfectly open to a comparison theory to propose that a simile and its corresponding be-form metaphor lead to markedly different forms of comparison in the understander's mind. I call such a theory a *Different-Comparison* theory. It is notable that the debate is sparse on discussion of this possibility. What seems to have happened is that because people (on both sides of the debate) have intuitively viewed comparison theories as saying that be-form metaphor is elliptical simile, they have been tacitly led to the erroneous conclusion that in a comparison theory the two forms must obviously involve the *same* comparison process. I will call this type of theory a *Same-Comparison* theory.

It is not my task to support or attack either of these two forms of comparison theory. Rather, the point is that ellipsis-based presumptions about comparison theory have led to certain unfortunate consequences in the C/C debate.

The first unfortunate consequence is that, through neglect of the Different-Comparison possibility, psychological experiments have been given unwarranted interpretations as favouring categorization accounts or as favouring comparison accounts. For example, some experiments have indicated that somewhat different sets of features of T and/or S are mentally activated by a be-form metaphor

as opposed to its corresponding simile. As an illustration, when the source term is 'a shark' as in 'My dog is [like] a shark,' a simile is more likely than metaphor to activate lower-level (e.g., more physical) features of sharks such as swimming, instead of activating, or as well as activating, higher-level (e.g., more abstract) features such as aggressiveness. This result has been used as evidence that the be-form metaphor is not understood by comparison (Glucksberg 2001). But, of course, the result in no way provides such evidence unless it is shown that if a comparison process were to be used for the be-form metaphor it would be the same process as for the simile; or unless it is shown that, even if different comparison processes are involved, their particular differences are not adequate to explain the particular feature-activation differences found. So, clearly, while the experimental results are indeed a challenge to Same-Comparison theories, they are much less of a challenge to Different-Comparison theories.

As another example, some experiments have unearthed form preferences (see the list of 'effects' above). These have been adduced as evidence that people prefer categorization to comparison in understanding be-form metaphor under certain conditions (see, e.g., Chiappe, Kennedy, & Smykowski 2003; Glucksberg & Haught 2006; Jones & Estes 2006). But, it could be that under those conditions people prefer a particular comparison process to another, rather than preferring a categorization process to a comparison process.

There are further unfortunate consequences of ellipsis-based assumptions about comparison theory. Even under a Same-Comparison theory, an elliptical view of metaphor cannot *seriously* mean that understanders actually convert the metaphor into the corresponding simile, i.e., convert between *surface* forms in their minds. For one thing, considerable linguistic processing, involving at least syntactic analysis and possibly more, might well be needed to see where to place the word 'like' into the sequence of words making up the surface form of a be-form metaphor in order to convert it into the corresponding simile. This can easily be seen if the T and S terms are themselves syntactically complex – for example, the metaphor might be 'The lawyer who is dealing with David's divorce is in my opinion a marauding shark, or worse.' Finding where to insert 'like' here is a syntactically complex task. Thus, in reality, the intended equivalence between the be-form metaphor and its corresponding simile must surely be at a deeper level of representation/processing of the utterances. Despite this point itself being rather obvious, the following simple consequences of it appear not actually to have been seen in the debate.

Under a Same-Comparison theory, there is by definition a comparison process $\text{Comp}(T,S)$ such that both 'T is/are like S' and 'T is/are S' (when taken as possibly being a be-form metaphor rather than a literal statement) lead by certain utterance-processing routes to $\text{Comp}(T,S)$. Let us call these routes the 'pre-processing' routes. Now, because both routes both lead to $\text{Comp}(T,S)$, they must first converge

either at $\text{Comp}(T,S)$ itself or at some earlier common representation or process. And, importantly, this point must be *deeper than surface form* if we are to avoid supposing that the understander converts between surface forms. (See Figure 1.) So, the portions of the pre-processing routes up to the convergence point are necessarily non-empty and necessarily different. Moreover, there may be different ‘side-processing’ hanging off those route portions (see the Figure): i.e., the side-processing does not of itself form part of the processing leading up to the convergence point, but is just additional processing arising in some way from the original utterance, or from some interim processing result later in the pre-processing route. For example, one major possibility for side-processing in the be-form metaphor case is that the understander might be investigating, also, whether a categorization of T within the literal S category is possible or useful.

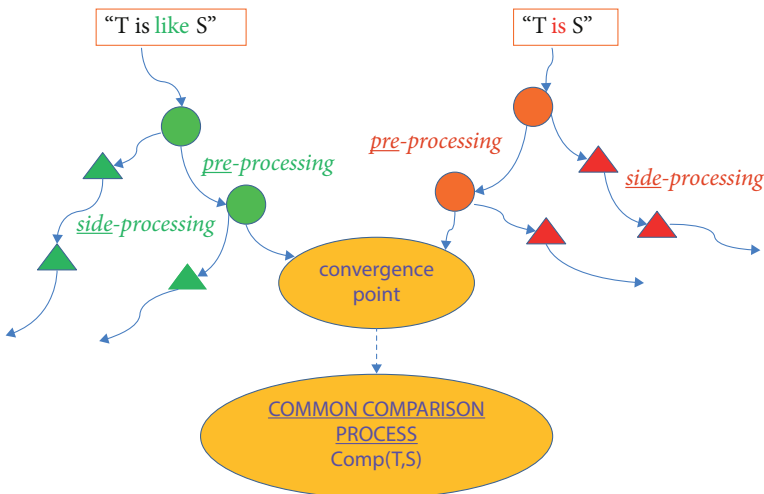


Figure 1. Processing in a Same-Comparison theory

Now, the debate has been largely blind to the question of where the convergence point of the two pre-processing routes is, because of naively thinking of the metaphor being processed *via the simile form* – i.e. *as if* the convergence point were the simile form itself. So, the point about the simile and its corresponding metaphor leading to necessarily different pre-processing, and possibly different side-processing also, has been missed. But these differences could in principle seriously compromise the interpretation of experiments that allegedly pit comparison against categorization theory. For example, processing time differences could be affected by differences in the pre- and side-processing. So processing time differences between a simile and its corresponding metaphor do not show that they are not both using comparison (or not both using categorization).

Equally, the question of exactly where the convergence point of the two processing routes in a Same-Comparison theory is – is it at Comp(T,S) itself, or earlier? – and the question how that location could effect the interpretation of experiments, have not been considered.

4. Towards an alternative battlefield

Partly as a result of the above analyses, I suggest that there are various dimensions of potential variation between and within accounts of the understanding of simile and metaphor that should be considered – at least as an addition to, but quite possibly a replacement for, the categorization/comparison distinction (C/C distinction). Amongst the new dimensions are *Target/Source Contribution Disparity*, *Target/Source Mediator-Preservation Degree*, and *Target/Source Mediator-Carefulness*. These new dimensions are, I suggest, more sharply definable than the C/C distinction is, though they themselves remain to be sharpened further, in future work. Also, I suggest that they are often more directly and clearly related to the experimental results than the C/C distinction is. This is both because the dimensions intrinsically cut across the C/C distinction and because there are reasons to think that the experimental effects can often be explained by where an account lies on the dimensions, *irrespective* of whether it uses comparison or categorization. The suspicion is that for a particular point in the space of possibilities that is defined by the dimensions, either a suitably framed comparison theory or a suitably framed categorization theory could lie at that point.

4.1 Target/Source contribution disparity

In either a simile ‘T is/are like S’ or its corresponding be-form metaphor ‘T is/are S’, the understanding process is obviously affected both by T and by S. The question in this subsection is about the particular ways in which T and S contribute to parts of the process of understanding, according to comparison and categorization theories. I will focus on what is involved in working out what I will call the ‘mediator.’ The mediator in the case of comparison is the similarity or analogy that is found, and in the case of categorization it is the superordinate category found. In this subsection I will not be addressing the remaining parts of the understanding process, notably what happens as a result of finding the mediator, such as proposing new conjectures about T.

The intuition beyond the word ‘mediator’ is that it mediates between the relevant literal meaning of T and the relevant literal meaning of S. But be careful to note that the word ‘mediator’ does not embody an assumption that the

literal meaning of S is actually accessed in understanding: it may or may not be, according to the particular theory at hand and to the particular circumstances. For instance, in the case of an expression S that is frequently used in a particular metaphorical way, it may be that only a superordinate category S* or other sort of entrenched metaphorical meaning is accessed from memory when understanding a metaphor *T is S*, the literal meaning being bypassed entirely. But I still call S* a mediator because it is superordinate to the literal meanings of both S and T.

Our specific question about the contributions of T and S to the mediator-finding will be about the possible quantitative and qualitative similarities or disparities between those contributions. That is, do T and S affect the process to roughly the same degree, or is the process markedly more dependent on S, say? Do T and S affect the process in qualitatively the same way or in markedly different ways? It will be important to keep in mind that the nature of the contribution of T or S to mediator-finding in a particular case could depend (under a particular theory) on the particular T and S involved, on the grammatical form (simile or be-form), and on contextual factors.

Now, there is a highly salient contrast between existing comparison and categorization accounts as regards the contributions of T and S to mediator-finding. Comparison theories have assumed much more equality between T and S than categorization theories have. Notably, in accounts based on SMT or something like it, and therefore in the Career of Metaphor theory as a salient example, the finding of a partial structured analogy between the T concept and the literal S concept is completely symmetrical: the same mappings (in reversed form) would be found if T and S were interchanged. In this sense, T and S contribute equally to the mediator-finding (i.e., the computation of the mappings), both quantitatively and qualitatively. Wolff & Gentner (2011) use this symmetry to account for their experimental findings that forward and reversed versions of highly directional metaphors are rated equally comprehensible at an early stage of processing, but unequally later. The later inequality arises from the strong asymmetry in the stage of their model that *follows* mediator-finding, namely the stage of positing candidate inferences about the target, based on unmapped parts of the source. In contrast, in existing categorization theories, the finding of the mediator, i.e. the superordinate category S*, uses S and T very differently: it basically proceeds by modification of the literal S category or by accessing some existing superordinate of S, under some *guidance* from T and context. For example, Glucksberg (2001) proposes that 'topic dimensions' arising from T and context guide the S* access/construction. T has a less detailed and more filter-like influence than S has. If T and S were interchanged, a very different superordinate might be found. This consequence could indeed, as has been claimed, help to explain what happens when metaphors or similes are reversed (Glucksberg 2001). A comparison theory

whose comparison process is symmetrical needs to appeal to other aspects of understanding to explain reversal effects, aspects such as the transfer of unmapped information from S to T (cf. the candidate inferences mentioned above).

But this contrast between comparison and categorization theories is not a necessary one. It is just a feature of particular theories that happen to have been proposed. There's no reason in principle why an analogy-finding process cannot treat T and S very differently. For example, analogy-finding might have more tolerance for leaving out information on one side than on the other, given that usually only a partial analogy will be found in any case. As another suggestion, what specific analogy is found may depend on whether the process is guided by looking first at target features or looking first at source features (or looking at them in parallel). These two possibilities for asymmetry in analogy-finding are just in-principle ones, amongst many others, and I do not claim that there is evidence that they apply: the point is rather that an experimental effect (such as a change in what T and S features are activated occurring when a simile or be-form metaphor is reversed) does not show that comparison is not the process being used, but at most that *a symmetrical form of comparison* is not the process being used – and it may not even show that.

Conversely, there's no reason in principle why a categorization process that finds a common superordinate for two categories T and S, such as businesses and dictatorships, cannot be completely symmetrical. There is nothing about categorization in itself that dictates that T and S should be treated disparately. It is merely a special assumption of existing categorization theories that they are. It could be this assumed disparity, rather than categorization as such, that is responsible for certain experimental results. This would apply most obviously to the results of experiments about changes in meaning or understandability when be-form metaphors are reversed (see, e.g., Campbell & Katz 2006; Chiappe, Kennedy & Smykowski 2003; Glucksberg 2001). And similar effects could potentially be predicted by an asymmetrical comparison theory.

People's form preferences could also be affected by the extent/type of T/S contribution disparity involved in a comparison or categorization process. For example, suppose both simile and be-form metaphor are handled by comparison, but by using different versions of comparison. Suppose one of the comparison processes treats T and S equally, but the other treats them very disparately in some way. Let us call these the 'equal' and 'disparate' comparison processes respectively. The disparate process will be differently appropriate according to which way round a simile or be-form metaphor is, whereas the equal version will be less affected by this. Hence, if, say, the disparate process is used by the understander for be-form metaphor but the equal version for simile, then his/her preferences will be more affected by reversals in the metaphor case than in the simile case.

In sum, the extent and type of Target/Source Contribution Disparity is itself a central feature of a theory of simile and/or metaphor, and can contribute to the effects studied in experiments in the C/C debate. It can contribute in ways independent of whether the underlying process is one of comparison or categorization.

4.2 Target/Source mediator-preservation degree

Does the main information that an understander gleans from a T/S simile or be-form metaphor consist only of some information about T in its own right (e.g. that businesses are non-consensually run, in the businesses/dictatorship case), or does it also include the mediator itself, e.g. the analogy found between businesses and dictatorships, or a superordinate category S^* covering them both? To put it another way, is the mediator *merely* a stepping-stone towards information about T in its own right, or is it itself preserved as part of the meaning of the utterance?

We can put the issue a little more precisely and usefully as follows. When understanders have understood a T/S simile or be-form metaphor, they presumably have derived some information about T from the utterance, and/or have cognitively highlighted some already known aspects of T. I will call either of these cases the ‘comment’ on T that they have derived from the utterance via the mediator. For some period (which may or may not extend for an appreciable time beyond the time of understanding the utterance) this T-comment will be active in the understander’s mind. The question now is whether the understanding process preserves the mediator itself in the sense of making it keep a comparably high level of activity during that period, or whether it allows the mediator activation to drop away once the comment on T is determined.

I thus identify the extent of ‘[target/source] mediator preservation’ as another dimension of variation between and within theories of simile and be-form metaphor (and indeed other types of metaphor). Main issues about the dimension include that of

- i. whether simile and be-form metaphor generally differ as regards mediator-preservation, or under what conditions they differ if ever;
- ii. whether similes differ amongst themselves on the extent of mediator-preservation;
- iii. similarly, whether be-form metaphors do; and
- iv. whether comparison or categorization accounts have anything helpfully *different* from each other to say about mediator-preservation (in simile and/or be-form metaphor).

Now, if the answers to (i), (ii) and (iii) were No – i.e., the extent of mediator-preservation were constant both between simile and be-form metaphor and between different similes/metaphors – then there would be little virtue in considering the dimension. But there are reasons to think that the three answers are all Yes, as I will show in the following. As for (iv), I will suggest a No answer in the course of the discussion.

Taking (i) first, simile arguably can be expected in most discourse contexts to have a stronger mediator-preservation tendency than (be-form) metaphor has, for a given T and S. More precisely, there is good reason to think that simile normally preserves the mediator, whereas there is no particular reason to think that be-form metaphor normally does so. This conjecture arises from considering what the simile and corresponding metaphor *apparently mean*. This apparently-meaning is related to something that has been a major factor in the C/C debate, with various theorists putting weight on the idea that a be-form metaphor appears by its very form to assert a categorization. Some authors then appeal to a principle of ‘grammatical concordance’ (the name used by Bowdle & Gentner 2005) to suggest that the *apparent* assertion of a categorization should be taken seriously to suggest that a process of categorization *is actually* used in understanding.

My own view, however, is as follows. A simile apparently means that T and S, taken literally, are alike, in some contextually relevant sense that the understander is meant to uncover. The corresponding metaphor apparently means that T is categorized under S, taken literally. But, crucially, *it is only in the simile case that reality matches appearance*: the understander does indeed normally take the simile to mean that T and literal S are *in fact* alike in certain respects. But, by contrast, in the case of the corresponding metaphor the understander does not, of course, normally take the utterance to mean that T is *in fact* categorized under the literal S.¹ Certainly, under a categorization account, a crucial intuition is that the understander does take the utterance to claim that T is to be categorized under *some* category, namely the chosen superordinate S*. So there is a real categorization – but still it is not the particular categorization apparently stated by the sentence. And of course under a comparison account the utterance is not taken actually to assert a categorization at all.

Thus, in the case of be-form metaphor, under either a categorization account or a comparison account the understander takes the utterance actually to mean something markedly different from what it apparently means – so the mediator, which is a contribution to determining the actual meaning, is part of a *departure*

1. Insofar as the understander might take the utterance to assert a categorization under the literal S, as in the special case of a “twice-true” metaphor, that categorization would be a separate interpretation, not an aspect of the metaphorical interpretation.

from the apparent meaning. But in the case of the corresponding simile the understander takes the utterance actually to mean what it apparently means, and the mediator merely *elaborates* the idea that the target and source are alike, an idea that is actually meant by the utterance.

We see from this that the mediator plays very different roles in be-form metaphor and in simile. Since a simile really does mean that there is a likeness between source and target, and the mediator portrays in detail what this likeness is (as far as the understander can see), there seems to be no good reason for the understander not to take the mediator to be part of the meaning. On the other hand, in the case of the corresponding metaphor, there is no particular reason for the understander to take the mediator itself to be part of the meaning as opposed to being merely a way of getting to the meaning – the derived comment about T.

Note carefully that this argument does not show that the mediator cannot be or should not be part of the meaning in the metaphor case, but merely that there is (normally) no pressure for it to be so. However, I argue elsewhere (Barnden 2010) that the mediator can be part of the meaning in some cases, and indeed that sometimes it is very difficult to see how the understander could intuitively understand the utterance without including the mediator in the meaning. So the claims above are about defaults: the default in the simile case is plausibly that the mediator is part of the meaning, the default in the be-form metaphor case is plausibly that it is not.

Our argumentation has been about meaning rather than directly about the extent to which the cognitive activation of the mediator is preserved. But it at least suggests that with a simile there will be more of a tendency for the activation of the mediator to be maintained than in the corresponding metaphor.

A further word about the statement above that a simile means that the literal T and S are in fact alike. The metaphor/simile literature is full of claims that simile only states a figurative comparison, as opposed to a literal one such as a comparison between apples and pears (see, e.g., Glucksberg 2001). But this can only be taken as a comment on the particular sort of likeness that is at issue: in either case a likeness is in fact being claimed. In simile, *whatever* likeness is at issue and irrespective of whether this makes the comparison ‘figurative’ or ‘literal’, there is *literally* a likeness, according to the utterance. The utterance is literal insofar as it claims S and T are *alike*: the use of ‘alike’ here is literal, not figurative!

As for questions (ii) and (iii), I will take them together and consider novel T/S pairings versus familiar ones, irrespective of whether T and S are in a simile or in a be-form metaphor. The more familiar the pairing, the less reason the understander has to suppose that the utterance is pointing out a comparison or categorization for its own sake rather than just using the comparison or categorization to comment on T; hence the less useful it is to keep the mediator (comparison result or

superordinate category) in mind along with the comment on T. Going in the other direction, the more novel the pairing, the more natural it is to suppose that the utterance is not only commenting on T but also pointing out an unfamiliar linkage between T and the literal S.

As for question (iv), the comparison/categorization distinction appears tangential to the mediator-preservation dimension and therefore misses an important issue. There does not appear to be anything about comparison or categorization that would differently affect the extent to which a mediator is preserved. Yet the extent of mediator-preservation could strongly influence some of the types of effect studied in experiments within the C/C debate. But if mediator-preservation degree is itself affecting experimental results, the difference between comparison and categorization as such is beside the point. There are several respects in which mediator-preservation might contribute to the experimental effects, but I will concentrate on some that are bound up with the third dimension to be considered, in the next subsection. Other, independent, contributions will be set out elsewhere.

4.3 Mediator carefulness

The dimension discussed here is the extent to which the understanding process is careful in working out the mediator. Carefulness could cover the amount of effort put in, the exactitude versus approximateness of the processing, the extent to which different alternatives are tried out, the fullness of an analogy found in a comparison process, the basis on which the process determines that it has adequately succeeded, and so forth.

It is reasonable to suppose that the degree of mediator-preservation correlates positively with the degree of mediator-carefulness. In particular, if the mediator is preserved because it is taken to be part of the meaning as opposed to being just a stepping-stone to meaning, then the understander can be expected to be more careful in working it out. Equally, the more care that goes into working out the mediator, the more it can be expected to stay activated along with the T-comment.

Therefore, in the case of a metaphor in which there is little or no inclination to include the mediator in the meaning, as Section 4.2 hinted will often be the case, the understanding process is more easily satisfiable just by finding any mediator that works well enough: i.e., is good enough for the discourse purposes at hand concerning finding some relevant comment on T.

Thus, we would expect, as a broad generalization, that vague, loose, quickly-derived mediators are more acceptable in be-form metaphor than in simile. This could help to explain why be-form metaphor sometimes involves less attention than simile does to the literal S (Glucksberg 2001), in that less careful processing

is likely to concentrate on more generally applicable features than on investigating peculiarities of the literal S.

Higher mediator-carefulness in understanding a simile or be-form metaphor could lead to a new attempt to find a mediator between T and S, even if a known one is already available – e.g., even if there is a known S^* within which T can readily be put. This could easily help to explain why a simile can lead to different interpretations from its corresponding metaphor, even for a conventional T/S pairing (Glucksberg & Haught 2006). (Recall here from Section 3 the claim that simile tends to appeal to lower-level features – such as swimming in the ‘My dog is [like] a shark’ example – more than metaphor does, with metaphor tending to focus on higher-level features such as aggressiveness). That is, to the extent that the simile encourages mediator preservation and therefore mediator carefulness, it will encourage a fuller attempt to find a mediator.

Given such links to experimental effects studied in the C/C debate, I propose that a central feature of simile/metaphor theories should be the particular way that the amount of care taken over finding a mediator is influenced by the particular T and S, by the form of the utterance (simile form versus be-form) and by context. But again, the dimension cuts across the comparison/categorization distinction. Either type of process can be arbitrarily careless or careful. So, the C/C distinction misses the point.

5. An example of the application of the dimensions

Here I illustrate the usefulness of our three dimensions by showing how they illuminate the nature of a particular cognitive-modelling framework for simile/metaphor understanding. The framework is that of Utsumi (2011). It includes precise, simple, algorithms *Compa* and *Categ* that he proposes as a way of modelling the comparison and categorization approaches to metaphor respectively, as a contribution to the C/C debate. Owing to space limitations I can only summarize my observations here. I spell them out in detail in Barnden (2012).

As Barnden (2012) makes clear, the two algorithms are in essence very similar, differing only in the precise details of how they select features that are considered to be related to target and source and how they use those features to construct a metaphorical meaning.

In the framework within which the algorithms operate, utterance meanings and concepts known to the system are represented as high-dimensional numerical vectors. So the target and source concepts T and S are represented as such vectors. Also, a collection of vectors can be regarded as providing a new concept, by averaging their vectors to form a new vector. There is a notion of closeness of concepts,

based on the angle between their vectors. The *Categ* and *Compa* algorithms proceed as follows. The processing in each case can be regarded as first finding a mediator and then applying the mediator to construct a metaphorical meaning.

Categ finds the m known concepts closest to S (m being a parameter set by the researcher). Amongst those concepts, it finds the k that are closest to T (k being another parameter). The average of these k concepts can be taken as encoding the superordinate category S^* in the categorization approach. That category is the mediator in our terms. Notice that the finding of the mediator is highly *asymmetrical* with respect to S and T – the result would normally be very different if they were interchanged. The k vectors are then used to provide a metaphorical meaning by averaging them together with the S and T vectors. But note that this can be reframed as finding a suitably weighted average of the mediator vector and the S and T vectors. This application of the mediator to find a metaphorical meaning is completely *symmetrical* in S and T – it would not matter if they were interchanged. (This symmetry is a questionable feature of the algorithm, but we will take it as it stands.)

Compa, on the other hand, finds (at least) k ‘common neighbours’ of S and T . It does this by taking larger and larger ‘neighbourhoods’ of vectors close to S and T – starting with the closest vectors to S and moving outwards, similarly for T – until the intersection of the S neighbourhood and the T neighbourhood has at least k vectors. This process is completely *symmetrical* in S and T . The vectors found are deemed to be the common neighbours, and are regarded as features that S and T have in common. In our terms, either the vectors collectively or their average can be regarded as the mediator. A metaphorical meaning is obtained by taking the average of the vectors with T (but not S). This application of the mediator is highly *asymmetrical* in S and T .

Despite Utsumi’s own view of the algorithms, it is a highly subjective and disputable act to construe one of them as doing categorization and the other as doing comparison. Why shouldn’t the k vectors found by *Compa* be taken to define a superordinate category, just as the k vectors found by *Categ* are? Conversely, why shouldn’t the k vectors found by *Categ* be regarded as encoding the result of a comparison of S and T , just as the k vectors found by *Compa* are? The alleged contrast depends largely on construing the features that *Compa* selects as being common neighbours of target and source but refraining from adopting this construal in the case of *Categ*. However, Barnden (2012) throws considerable doubt on the validity of this divergence of construal. What can be established much more clearly and objectively than whether the algorithms achieve comparison or categorization is the positioning of the two algorithms on the three dimensions above. In brief, *Compa* has zero contribution disparity in the calculation of the mediator (because of the symmetry noted above) but a high degree of disparity in

the way the mediator is applied to construct the metaphorical meaning (because of the asymmetry noted above). *Categ* is the other way round: it has a high degree of disparity in mediator calculation but no disparity in the use of the mediator to construct a metaphorical meaning. *Compa* is considerably more careful in finding the mediator than *Categ* is, partly through *Compa*'s use of a relatively elaborate, incremental process of finding common neighbours of target and source. On the other hand, both algorithms have a moderately high degree of mediator-preservation. This is because the metaphorical meaning in each case is in fact a distorted version of the mediator.

A further observation is that many detailed variations of the *Compa* and *Categ* algorithms could be proposed that would vary their amounts of contribution disparity, mediator preservation and carefulness in many different ways and to different extents. In particular, mixtures of or compromises between the algorithms can easily be imagined, and would make it even more difficult to usefully and uncontentionally label a resulting algorithm as doing either comparison or categorization. Indeed, the algorithms actually lie on a rich continuum of possible algorithms. For example, *Compa* averages the mediator vectors together with the vector describing the target, whereas *Categ* averages the mediator vectors together with *both* the target vector and the vector describing the source. But in principle there could be any weighting of the target and source in doing such averaging.

This point contributes to casting doubt on the usefulness of the notions of comparison and categorization for distinguishing between algorithms on the continuum. By contrast, the amounts and types of disparity, mediator preservation and carefulness of any algorithm on the continuum would remain precisely describable, and the dimensions would therefore be better placed to act as independent variables in experimental cognitive modeling based on the continuum.

6. Conclusion

The comparison/categorization battle, as currently framed, is misguided. There is an alternative to the battle, or at least there is a more pressing, immediately relevant battle: one to be waged within a theoretical space defined by various dimensions that cut across comparison and categorization. This article has proposed three particular dimensions: Target/Source Contribution Disparity, Target/Source Mediator Preservation Degree, and Target/Source Mediator Carefulness. But there could be others. The nature of the three dimensions discussed requires further work – the claim is not that they are yet as sharply defined as they need to be, but rather that they are the start of a better basis for systematic investigation of the relevant understanding processes than the notions of comparison and categorization are.

Moreover, I do not claim that theories should not be framed in terms of comparison and/or categorization. Rather, if they are, then they need to be more careful about the detailed nature of the proposed comparison and categorization. And the theories may most fruitfully be able to account for experimental results not in a direct way but instead indirectly: via the way in which the specific proposed comparison and categorization processes range along the dimensions under the influence of utterance form, the particular source and target, and context.

The influences just mentioned are in part a function of the particular communicative context and background as opposed to stable linguistic features of the simile or metaphor at hand or cognitive matters private to the understander. Notably, the degree of Mediator Preservation can be expected to be influenced by the understander's perception of whether the speaker herself is intending to point out a target/source similarity or merely to say something about the target. This matter is in part to do with how novel the understander takes the particular target/source pairing to be, which is itself a matter of a peculiarly communicative matter (the history of the use of that target and source in past communications the understander has been involved in). But there could be other types of clue in the specific current discourse as to what the speaker's intentions are. Also, independently of these influences, the degree of Mediator Carefulness can be expected to be affected by how carefully the understander perceives the speaker herself to be thinking about the mediator, notably through the extent of detail she gives as to why she has paired the particular target and source.

Acknowledgments

The research in this article was supported in part by Research Project Grant F/00 094/BE from the Leverhulme Trust in the UK. It derived from work supported by grant EP/C538943/1 from the Engineering and Physical Sciences Research Council in the UK. I have benefitted from discussion with Mihaela Popa, Robyn Carston and Dedre Gentner.

References

- Barnden, J. A. (2001). Uncertainty and conflict handling in the ATT-Meta context-based system for metaphorical reasoning. In V. Akman, P. Bouquet, R. Thomason, & R. A. Young (Eds.), *Modeling and using context: Third international and interdisciplinary conference (CONTEXT 2001)*. Lecture Notes in Artificial Intelligence, Vol. 2116, (pp. 15–29). Berlin: Springer.
- Barnden, J. A. (2008). Metaphor and artificial intelligence: Why they matter to each other. In R. W. Gibbs, Jr. (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 311–338). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511816802.020

- Barnden, J. A. (2010). Metaphor and metonymy: Making their connections more slippery. *Cognitive Linguistics*, 21(1), 1–34. DOI: 10.1515/cogl.2010.001
- Barnden, J. A. (2012). Metaphor and simile: Reframing the comparison *versus* categorization debate. *Cognitive Science Research Paper CSRP-12-01*, School of Computer Science, University of Birmingham, UK, March 2012.
- Barnden, J. A. (2015). Open-ended elaborations in creative metaphor. In Besold, T. R., Schorlemmer, M. & Smaill, A. (Eds), *Computational Creativity Research: Towards Creative Machines* (pp. 217–242). Berlin: Atlantis Press, Springer.
- Bowlde, B. F., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, 112(1), 193–216. DOI: 10.1037/0033-295X.112.1.193
- Campbell, J. D., & Katz, A. N. (2006). On reversing the topics and vehicles of metaphor. *Metaphor and Symbol*, 21(1), 1–22. DOI: 10.1207/s15327868ms2101_1
- Carston, R. (2011). Metaphor, hyperbole and simile: Two routes to metaphor understanding. Talk at Conference on *Metaphor and Communication*, Cagliari, Italy, 12–14 May 2011.
- Chiappe, D. L., & Kennedy, J. M. (2001). Literal bases for metaphor and simile. *Metaphor and Symbol*, 16(3–4), 249–276. DOI: 10.1080/10926488.2001.9678897
- Chiappe, D., Kennedy, J. M., & Smykowski, T. (2003). Reversibility, aptness, and the conventionality of metaphors and similes. *Metaphor and Symbol*, 18(2), 85–105. DOI: 10.1207/S15327868MS1802_2
- Connolly, A. C., Fodor, J. A., Gleitman, L. R., & Gleitman, H. (2007). Why stereotypes don't even make good defaults. *Cognition*, 103, 1–22. DOI: 10.1016/j.cognition.2006.02.005
- Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The Structure-Mapping Engine: algorithm and examples. *Artificial Intelligence*, 41(1), 1–63. DOI: 10.1016/0004-3702(89)90077-5
- Fogelin, R. J. (2011). *Figuratively speaking*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199739998.001.0001
- Gentner, D. (1983). Structure-mapping: a theoretical framework for analogy. *Cognitive Science*, 7(2), 95–119. DOI: 10.1207/s15516709cog0702_3
- Glucksberg, S. (2001). *Understanding figurative language: From metaphors to idioms*. New York: Oxford University Press. DOI: 10.1093/acprof:oso/9780195111095.001.0001
- Glucksberg, S. (2008). How metaphors create categories – quickly. In R. W. Gibbs, Jr. (Ed.), *The Cambridge Handbook of Metaphor and Thought* (pp. 67–83). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511816802.006
- Glucksberg, S., & Haught, C. (2006). On the relation between metaphor and simile: When comparison fails. *Mind and Language*, 21(3), 360–378. DOI: 10.1111/j.1468-0017.2006.00282.x
- Gokcesu, B. S. (2009). Comparison, categorization, and metaphor comprehension. *Procs. Conference of the Cognitive Science Society*, 567–572.
- Hampton, J. A. (2007). Typicality, graded membership, and vagueness. *Cognitive Science*, 31, 355–383. DOI: 10.1080/15326900701326402
- Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, 13(3), 295–355. DOI: 10.1207/s15516709cog1303_1
- Johnson, A. T. (1996). Comprehension of metaphors and similes: a reaction time study. *Metaphor and Symbol*, 11(2), 145–159. DOI: 10.1207/s15327868ms1102_3
- Jones, L. L., & Estes, Z. (2006). Roosters, robins, and alarm clocks: Aptness and conventionality in metaphor comprehension. *Journal of Memory and Language*, 55, 18–32. DOI: 10.1016/j.jml.2006.02.004
- Kennedy, J. M., & Chiappe, D. L. (1999). What makes a metaphor stronger than a simile? *Metaphor and Symbol*, 14(1), 63–69. DOI: 10.1207/s15327868ms1401_7

- Miller, G. A. (1979). Images and models, similes and metaphors. In A. Ortony (Ed.), *Metaphor and thought* (pp. 357–400). Cambridge: Cambridge University Press.
- Pierce, R. S., & Chiappe, D. L. (2009). The roles of aptness, conventionality, and working memory in the production of metaphors and similes. *Metaphor and Symbol*, 24(1), 1–19.
DOI: 10.1080/10926480802568422
- Rein, J. R., Goldwater, M. B., & Markman, A. B. (2010). What is typical about the typicality effect in category-based induction? *Memory & Cognition*, 38(3), 377–388.
DOI: 10.3758/MC.38.3.377
- Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104, 192–232. DOI: 10.1037/0096-3445.104.3.192
- Sperber, D., & Wilson, D. (2008). A deflationary account of metaphor. In R. W. Gibbs, Jr. (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 84–105). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511816802.007
- Utsumi, A. (2007). Interpretive diversity explains metaphor-simile distinction. *Metaphor and Symbol*, 22(4), 291–312. DOI: 10.1080/10926480701528071
- Utsumi, A. (2011). Computational exploration of metaphor comprehension processes using a semantic space model. *Cognitive Science*, 35, 251–296.
DOI: 10.1111/j.1551-6709.2010.01144.x
- Verbeemen, T., Vanpaemel, W., Pattyn, S., Storms, G., & Verguts, T. (2007). Beyond exemplars and prototypes as memory representations of natural concepts: A clustering approach. *Journal of Memory and Language*, 56, 537–554. DOI: 10.1016/j.jml.2006.09.006
- Wills, A. J., Noury, M., Moberly, N. J., & Newport, M. (2006). Formation of category representations. *Memory & Cognition*, 34(1), 17–27. DOI: 10.3758/BF03193383
- Wolff, P., & Gentner, D. (2011). Structure-mapping in metaphor comprehension. *Cognitive Science*, 35, 1456–1488. DOI: 10.1111/j.1551-6709.2011.01194.x
- Xu, X. (2010). The relationship between the content and form of metaphorical statements. *Journal of Psycholinguistic Research*, 39, 165–178. DOI: 10.1007/s10936-009-9131-5