Metaphor and Simile:
Reframing the Comparison *versus* Categorization Debate

John Barnden
School of Computer Science
University of Birmingham, UK

(+44)(0)121-414-3816   j.a.barnden@cs.bham.ac.uk

Abstract

Do understanders use comparison or categorization in understanding simile or metaphor? This article claims that the intense debate about this has been misdirected, for various reasons including inattention to the different possible forms of categorization and comparison, and to the murkiness of the comparison/categorization distinction. Instead of using that distinction as a crucial divider between metaphor/simile theories, it is more revealing to consider a theoretical space defined by various processing dimensions that cut across comparison and categorization: Target/Source Contribution Disparity, Target/Source Mediacy-Preservation, and Target/Source Mediator-Carefulness (where mediators and mediacy concern the relationship between the two concepts juxtaposed in a metaphor or simile). Experimental results adduced in the debate do not show whether categorization or comparison is being used but rather where the understanding process is positioned on those dimensions (and perhaps others). Comparison and categorization can both supply any needed positioning.

NOTES to Editors and Reviewers:

(1) This article has 16,600 words all-inclusive. I believe it merits being in the Extended article category (target length 18,000 words), possibly with added discussion of unclear points at the discretion of editors and reviewers, but I would be happy to have it instead as a Regular article (target length 12,000 words) after suitable shortening. If the decision is for the latter then I would be grateful for any suggestions about what content could be sacrificed.

(2) This article is a greatly extended and elaborated version of a book chapter (Barnden, forthcoming). That chapter, which has only about 7,400 words, is in essence an extended abstract of some of the current article. In particular, this article contains various additional major components (notably sections 5.4 and 6), ideas, evidence and arguments not reflected at all in the chapter, and refines some of the ideas there.

1 Introduction

There has been intense debate about whether people use processes of *comparison* as opposed to *categorization* (or *class-inclusion*) in understanding simile and/or metaphor. See, for example, Bowdle & Gentner (2005), Chiappe & Kennedy (2001), Glucksberg (2001, 2008, 2011), Glucksberg & Haught (2006a,b), Jones & Estes (2006), and Utsumi (2007, 2011). This comparison/categorization debate (henceforth the [C/C] Debate) is basically about how similes such as "Businesses are like dictatorships" and corresponding metaphors such as "Businesses are dictatorships" are interpreted by ordinary understanders. The first term ("businesses" here) is the *target* (T) term and the second is the *source* (S) term.

This article points out oversimplifications and gaps that severely compromise the Debate and suggest the need to reframe it in other terms. The problems arise from unwarranted assumptions about categorization and comparison, about there being a clear distinction between them, and about how they can contribute to metaphor or simile understanding.

A second purpose is to support a particular conjecture about the Debate: that what has been revealed by psychological experiments within the Debate is not whether comparison or categorization is involved in understanding, but rather some differences between positions on various other processing dimensions that cut across the C/C distinction. These dimensions include:

- Target/Source Contribution-Disparity: roughly, the extent of difference between how source and target contribute to the understanding process.

- Target/Source Mediacy-Preservation: roughly, the extent to which the source/target relationship is itself important information for the understander rather than merely a stepping stone to illuminating the target.

- Target/Source Mediator-Carefulness: the amount of care (of a certain sort) taken to establish the source/target relationship.

Movement along the dimensions is possible both within and between particular theories of simile/metaphor

understanding. Within a theory, the understanding process could vary along the dimensions depending on the particular target and source, the sentence form (simile form or metaphor form as above), and context. Between theories: for a given target and source, a given form and a given context, different theories can propose different positions on the dimensions.

Positioning on the dimensions can account for many salient experimental results in the Debate, irrespective of whether comparison or categorization is involved. Hence, instead of trying to explain results directly in terms of comparison and categorization, it would be more fruitful to explain them in terms of the above dimensions. Particular comparison- and/or categorization-based theories might then be devised to provide the requisite positioning along the dimensions. But the suspicion is that it will always be possible to find *some* comparison theory and *some* categorization theory that provide it. Our observations add to the argument made 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 Debate, and introduces some terminology and assumptions. Section 3 discusses the murkiness of the comparison/categorization distinction. Section 4 discusses an overlooked branch of comparison theory. Section 5 discusses the dimensions listed above, relating them to experimental findings. Section 6 illustrates various themes of the article in the specific case of Utsumi's (2011) understanding model. Section 7 concludes.

This article is an extended and refined treatment of issues covered in part in Barnden (forthcoming). The issues arose indirectly from the development of a computer-implemented AI model (ATT-Meta) of reasoning needed in metaphor understanding (Barnden 2001, 2008, Barnden, Helmreich, Iverson & Stein 1996, Lee & Barnden 2001; see also Barnden, Glasbey, Lee & Wallington 2004). The model can be viewed as including types of categorization and comparison, abetted by complex inferencing. However, the issues and observations in this article are independent of that model, which will not be discussed.

2 The Comparison/Categorization Debate

The main battle is about T-is/are-S metaphors, rather than about simile, and probes whether categorization or comparison is the mental process involved in understanding such metaphors. For simile there is more agreement that comparison is used. While a key earlier contribution, Glucksberg & Keysar (1990), claimed that similes are processed by categorization, Glucksberg (2008) allows this possibility only for apt similes.

I will call metaphors of the above form *be-form metaphors*, although they are typically labelled as nominal or copular. The Debate can be framed either as between two different forms of metaphor (be-form and simile form) or as between (be-form) metaphor and another figure, namely simile.

The S (source) term is almost always a phrase of the form "a/an <common noun>," such as "a shark," or a generic plural ("sharks"). As usual within the Debate (though see Gibb & Wales 1990, Xu 2010) we will not address meaning differences between syntactic variants of the T and S terms. Thus we will assume for simplicity that the metaphors "A business is a dictatorship," "Mike's business is a dictatorship" and "Businesses are dictatorships" are understood using the same main processes; similarly for the corresponding similes. Also, the Debate generally pretends there is only one simile corresponding to a given be-form metaphor and vice versa, though this is an oversimplification (cf. comments in Carston & Wearing 2011, Davidson 1978).

In the Debate, one broad camp leans more towards categorization for the understanding of be-form metaphor (and, at times, simile) and another leans more towards comparison for be-form metaphor (and proposes comparison for simile). "Leans" rather than "adheres," because some categorization theorists (e.g., Glucksberg 2001: 53–54, Glucksberg 2008) have allowed comparison sometimes to be the more appropriate process for metaphor and sometimes to be the process achieving categorization; and some comparison theorists have proposed that for suitably conventional metaphors categorization is the appropriate default process—cf. the Career of Metaphor theory (Bowdle & Gentner 2005). Glucksberg (2008) says that neither an all-out comparison theory nor an all-out categorization theory can be correct—comparison and categorization are used under different circumstances that remain to be elucidated. Somewhat related to this is the common acceptance that be-form metaphor and simile are often interchangeable (see, e.g., Carston & Wearing 2011 and Nowottny 1965 for evidence), the choice of form by an author often being Debate-irrelevant preferences or

influences such as stylistic factors or lexicosyntactic context. Thus, the Debate is not predicated on be-form metaphor and simile necessarily having markedly different effects, but only sometimes or often doing so.

In a categorization account of a metaphor "T is/are S," the understander accesses or constructs a suitable superordinate category S* of the category that is the S term's literal meaning (or, more precisely, its most-basically directly encoded meaning). So the literal S category is within S*, and may more strongly be required to be a good exemplifying subcategory of S*. The understander takes the utterance to assert that the T item or category is within S* (as member or subcategory respectively), not within the literal S category. For "Businesses are dictatorships", the understander might take *dictatorship** (S*) to be the category of organizations/communities managed non-consensually and punitively by one person. An appropriate S* category may already exist in the understander's mind linked to the S term or literal category, or may exist in his/her mind though not linked to S, or may need to be constructed during understanding. The choice or construction of an S* is often taken to be influenced by T, e.g., through the use of *topic dimensions* (Glucksberg, 2001), discussed below.

Categorization theory is motivated in part by the following rhetorical question: Why not take [be-form] metaphors to be exactly what they look like, namely categorization statements? (See Glucksberg 2001: 44 for a version of this.) The idea that be-form metaphors should be taken actually to convey categorizations, at least as a first try at understanding the metaphor, is one half of the notion of Grammatical Concordance (Bowdle & Gentner 2005). The other half is the idea of taking similes to be exactly what they look like, namely statements conveying presence of likeness (whose specifics are to be worked out by the understander). Grammatical Concordance is most thoroughly applied in the Distinct Statements theory of (Chiappe & Kennedy 2001, Chiappe, Kennedy & Smykowski 2003), where (be-form) metaphors always function like categorization claims, similes like similarity claims.

Various names have been used for categorization-based accounts: "class-inclusion" accounts, "dual-reference" accounts, and "interactive property-attribution" accounts. As for the last name (used by Glucksberg 2001), what is usually important about categories in the Debate is the properties they involve, not that they themselves be mentally-reified entities. However, most Debaters do talk as if the categories are reified.

In common with, e.g., O'Donoghue (2009), Sperber & Wilson (2008) and Utsumi (2011), I include Rele-

vance Theory (RT) accounts (Sperber & Wilson 2008) under categorization, as they appeal to "broadening" of the literal S category. But there can also be subsequent "narrowing" to get a final category for housing T (see especially Carston & Wearing 2011; also the approach is augmented there with a non-categorization-based branch).

In comparison accounts of be-form metaphor, concepts literally meant by the terms T and S are compared, uncovering similarities. Most typically, structural-analogy finding is proposed, notably as in Structure Matching Theory (SMT: Gentner 1983) or ACME (Holyoak & Thagard 1989). Aspects of the literal S concept that are not involved in the analogy are under certain conditions provisionally stipulated to apply (perhaps in modified form) to T. An example is the *candidate inferences* provided by the Structure Matching Engine (SME: Falkenhainer, Forbus & Gentner 1989).

Similarities found may instigate the mental construction of a common abstraction S* covering the literal T and S concepts. Indeed, CoM (the Career of Metaphor theory: Bowdle & Gentner 2005), a salient, hybrid, variant of the comparison approach, says that as the use of term S as a metaphor source becomes entrenched (conventionalized), one or more remembered common abstractions S* arising from such uses of S become standardly associated with S as superordinate categories/concepts. They become conventional metaphorical meanings of the term S, and can be used as known S* categories in a categorization account are. However, an understander can still go back to the literal S and compare it to the current T.

Thus, CoM supplements comparison with categorization. But, dually, categorization has been supplemented with comparison. For example, Glucksberg (2008) says that be-form metaphor may sometimes be understood by comparison. Presumably this is a possibility when no S* relevant to T is yet available or readily abstractable from the literal S. Utsumi's (2011) hybrid account proposes that an understander first attacks a be-form metaphor with categorization but under certain circumstances switches to comparison. The model will be important below because it contains detailed categorization and comparison algorithms.

The place of conceptual metaphor theory (CMT, Lakoff & Johnson 2003) in the Debate is curiously minor. Certainly, conceptual metaphor theorists do not say that the source-domain/target-domain mappings they propose that people already know are necessarily used online in understanding. They can merely be theoretical constructs motivating a particular understanding of the target domain, though perhaps developed

from online use earlier in ontogeny or history. But CMT-related proposals where mappings are available online (Gibbs 2011) could be regarded as comparison accounts where the analogy between the two sides is already known, not newly worked out. Indeed, in reality we need accounts where mixes of already-known and newly-found mappings can be used online. This is possible in Forbus, Ferguson & Gentner's (1994) I-SME. Also, Bowdle & Gentner (2005) and Gentner & Bowdle (2008) mention the need to be able to add incrementally to existing analogies. However, it is the working-out and online use of mappings that form the usual nub of comparison theory in the Debate.

Many psychological experiments have been done in support of one or another comparison, categorization or hybrid account. Amongst the effects on subjects and characteristics of metaphors and similes that have been studied in experiments, we will focus on the following:

- Form preference: which form—be-form over simile form—the subject prefers, depending on the nature of T and S.

- Relative ease/speed of understanding of the two forms.

- What sorts of feature of the literal T and/or S concepts, and/or additional features, are mentally accessed.

- The relative diversity of interpretations that subjects produce for the two forms.

- What happens to such effects when the T and S terms are interchanged.

- (Vehicle/source) conventionality: How conventional it is for the source [= vehicle] term to be used metaphorically.

- Familiarity (or: conventionality of whole metaphors or similes): How familiar subjects are with a particular T/S pairing being used in a metaphor or simile.

But there are other important effects and factors studied, notably perceived aptness of a metaphor or simile.

For brevity, a mental object directly accessed by the understander from the T or S term will be called a "concept" below, although labels such as "conceptual structure," "mental representation" or "category"

7

might often be better. The choice is itself a point of difference between theories. Again for brevity, we will use "S" variously for the term S itself or for the literal S concept, and "T" for the T term or concept.

3   Comparison *versus* Categorization (or Not)

There is considerable difficulty in distinguishing comparison and categorization in the first place, quite apart from using the difference to account for experiments.

3.1   Initial Considerations

Theoretically, to find or stipulate that two things are within a common category (i.e., to co-categorize them) *is* is just one way of finding or stipulating a similarity (comparison) between them. And if the category involves some necessary features for membership, the two things are found/stipulated to be similar in that both have those features. Conversely, comparison involves determining what properties two things share (and, possibly, do not share). Under "properties" we include deep, non-obvious ones such as some difficult-to-discern, highly abstract, underlying relational structure. But they are both then *ipso facto* in the category of things with those properties.

This theoretical inter-translatability of comparison and co-categorization is recognized in the Debate. Glucksberg (2001: 37–38) talks of literal and metaphorical comparisons putting the two things compared into a common category. Conversely, he says [*ibid,* p.40] "Because shark and salmon can belong to a common category, they are similar to each other," and in (Glucksberg 2011) he says that when a simile source term has an appropriate superordinate category, the comparison can occur by categorizing the target into the superordinate category rather than by feature matching. Gentner & Bowdle (2008) talk of the "common abstraction" (common superordinate concept/category) that results from doing any ("horizontal") comparison between a target and literal source.

Given the *theoretical* inter-translatability, Debaters must surely have in mind detailed differences at the *mental-processing* level. But even here, categorization and comparison have sometimes been equated, though usually assumed to be different. Chiappe *et al.* (2003) allow categorization to be performed by

comparison. Bowdle & Gentner (2005) see the fit of something into an encompassing category as a special, "vertical" sort of mental comparison. The creation of a new superordinate category S* may need to rest on comparison of S and T (Glucksberg 2008: 73). Chiappe and Kennedy (2001) say there is no motivation for invoking an ad hoc superordinate class at all: all the work is done by finding the common features of target and source, i.e. comparing them.

But a more general issue is that, surprisingly, both sides have generally been extremely vague about what categorization is, or categories themselves are, at the mental level. (Pierce & Chiappe 2009 make a similar point.) Usually Debaters just present categorization as being obviously different from comparison, and make assumptions about categorization that are often tacit or left unjustified. Often, categories seem to be baldly assumed to involve sets of necessary and jointly sufficient properties (e.g., this is assumed in Chiappe & Kennedy 2001) despite the importance of prototype and exemplar-based categorization theories, discussed below.

One salient exception to that unspecificity about the nature of categories is the detailed, mechanistic account of predication of Kintsch (2000, 2001, 2008; see also Kintsch & Mangalath 2011), claimed to account for aspects of metaphor understanding. Roughly speaking, the numerical vectors at the model's core represent word meanings, but can also be viewed as representing categories, and Kintsch (2008) suggests that the approach is closely related to Glucksberg's categorization model. Utsumi (2011) uses a precise algorithm borrowed from Kintsch (2001) and describes it as performing categorization. Thomas & Mareschal (2001) present a specific connectionist model that they cast as implementing categorization.

An exception to the widespread uncritical acceptance of a categorization/comparison distinction is the following under-appreciated point made by Bowdle & Gentner (2005)

> "Many theories of categorization assume that items are categorized by means of comparison, either to abstracted prototypes ... or to actual exemplars[.] Thus there is no reason to believe that the processes involved in categorization are different in kind from those involved in comparison."

(For presentation and discussion of prototype and exemplar theories, see for instance Hampton 2007, Rein,

Goldwater & Markman 2010, Rosch 1975, Verbeemen *et al.* 2007 and Voorspoels, Vanpaemel & Storms 2008.) That statement by Bowdle and Gentner is correct, but arguably too weak: Voorspoels, Storms & Vanpaemel (2011) claim that prototype and exemplar models are in fact "[t]he two most dominant computational theories of category representation. ... [E]xemplar representations provide the best description of human characterization." If they are right, much that is supposed in the Debate is questionable or needs major adjustment and clarification.

Now, there are further points to be made about prototypes and exemplars; there are additional points about categorization that need to be mentioned; and there is one additional, related claim by Bowdle & Gentner (2005) that we will argue *against*. The following subsections address these matters.

3.2 Complications Arising from Exemplars and Prototypes

In exemplar-based categorization, an item is deemed to be in a category, or to be so to some extent, according to its degrees of similarity to existing exemplars of the category. The similarity determination requires some form of comparison Also, categorization theories have been proposed that mix the exemplar and prototype ideas (Verbeemen *et al.* 2007): there can be a spectrum of mental representations, with the more specific ones representing exemplars and the least specific ones acting as prototypes, with intermediate ones acting as prototypes of subclasses.

Consider the metaphor example "My job is a jail" commonly used in the Debate. If a categorization account uses exemplar-based categorization, an interesting possibility arises that further narrows the gap between comparison and categorization. We assume that the understander, at least under some discourse conditions, tries to check the validity of the job/jail* categorization, rather than just stipulating the job to be a new exemplar of jail*. Thus, the job is compared to at least some exemplars of jail*. Now, *one* sort of exemplar of jail* is *physical*-jail exemplars. Thus, the comparisons actually done may also include comparisons of the job to physical-jail exemplars, or perhaps to an intermediate prototype for the physical-jails sub-class of jail*. Thus, not only does the categorization process rests on *some* sort of comparison, it might include (amongst others) the very same detailed comparisons as in a comparison theory of the sentence (assuming that that theory also takes the physical-jail concept to be represented by a prototype or exemplars).

What is also often unclear or deficiently supported, in a categorization theory of metaphor, is what the *results* of putting something T in a superordinate source-derived category S* are. The issue is sensitive to the particular type of categorization theory. This is where an additional claim by Bowdle & Gentner (2005) is defective:

> "The primary distinction between the two [comparison, categorization] may lie in the kind
> and degree of inference projection. Although comparison processing entails the projection
> of inferences, the inference process is highly selective; ... In contrast, categorization involves
> complete inheritance: Every property true of [the category into which target is being put] should
> be projected to the target."

This seems, ironically, to revert to a traditional notion of categories involving necessary properties. But, in a prototype-based categorization theory of metaphor, not all properties in the S* prototype would need to have equal, or any, likelihood of being selected for projection. For example, those most commonly instantiated in instances of the S* prototype, or those most strongly held by the current literal S concept, might reasonably be given priority. In an exemplar-based theory, there are more complications yet, because departures between T and a specific exemplar of S* may be more marked than departures between T and an S* prototype, there may be many relevant exemplars, and the exemplars may have conflicting properties. Moreover, inference projection from them must be selective, as in a comparison account. These matters, including selectivity and property-adaptation in inference projection, have been extensively studied in the AI area of case-based reasoning (Kolodner 1983), cases being essentially the same as exemplars.

Thus, Gentner and Bowdle's attempt to grant a comparison/categorization distinction on the basis of inference projection is undermined.

3.3   Process *versus* Product

Perhaps categorization should be viewed as primarily being about the product of understanding rather than the process, the reverse being broadly true for comparison. Then there need be no important battle. A comparison process could have a categorization product, for instance a common abstraction derived from

target and source in the process of comparison. Indeed, this seems precisely to be the point when, for example, Chiappe *et al.* (2003), in arguing that categorization is used for metaphor, say that comparison could be the mechanism used to achieve it. However, the Debate as a whole does not actually proceed on a process-as-opposed-to-product basis, but on a process-against-process basis. The rest of the article will therefore honor this spirit if only ultimately to exorcise it.

3.4   More Trouble for the C/C Process Distinction

In Glucksberg's categorization theory of metaphor the finding of the superordinate category S is in general guided by *topic dimensions* (Glucksberg 2001). These are important dimensions along which context suggests that the target T is being discussed. So, in "I'm very restricted in what I can do – my job is a jail" [my example] the first clause could supply the dimension of restrictiveness as one along which to find a useful superordinate category jail. Now, Glucksberg (2001: 53–54) accepts that this use of topic dimensions

> "does not preclude a comparison process in which information available in the metaphor vehicle [i.e., source] is assessed vis-à-vis information available in the metaphor topic [i.e., target]. The claim is that the properties per se of the topic and vehicle are not the appropriate inputs to the comparison process. Instead, vehicle properties on the one hand and topic dimensions on the other are the relevant inputs for comparison, analogous to the slots and fillers of head nouns and modifiers in conceptual combinations."

But this just says that the the argument is over precisely what form of comparison goes on, rather than over some distinction between categorization and comparison.

In some theories of categorization, categories are entirely implicit—that is, they are merely in the eye of the theoretical observer. Notably, prototype-based and exemplar-based theories can be like this, though are not necessarily so. If categories are merely implicit, one supposedly key feature of categorization theory collapses into something generally obvious and widely assumed. *Any* theory of metaphor must surely account for words having acquired highly entrenched metaphorical meanings that are superordinate with respect to their literal meanings. The noun "attack" plausibly has an entrenched general sense that subsumes both an

attack as in physical fighting and (some) types of attack in verbal criticism, sport, games, music, etc. (See Vervaeke & Kennedy 1996 for some discussion of "attack.") Plausibly, such entrenched general senses are directly activated on encountering the words, and can be cast as properties or categories: say, the property of being an action with a certain type of intent towards something, or the category of such actions. Put this way, the categorization-theory claim that words can have dual reference, one reference being to superordinate categories, is no more—in the particular case of conventional metaphor—than the statement that a conventional metaphorical sense can be directly accessed from a word. But metaphor researchers in general accept this.

Equally, CoM's proposal that categorization is used for conventional metaphor is not interesting in itself, and must *always and obviously* have been something that needed to be considered for addition to *any* comparison theory. What is interesting about the proposal is *the particular ways* in which superordinate categories or common abstractions arise during conventionalization, and of course CoM's way is one interesting hypothesis.

4 Additional Oversights

Both sides of the Debate have largely ignored a possibility within the space of comparison theories, as follows.

There is a common, largely unargued, and often tacit assumption that if a be-form metaphor is processed by comparison, then the comparison process is the same as that used for the corresponding simile. Sometimes the assumption is explicit—for example, Glucksberg & Haught (2006a: 361) say: "Comparison theories, no matter what their form, thus rely on the critical assumption that metaphors and similes are essentially equivalent to one another." But the hypothesis that comparison is used for both (be-form) metaphor and simile does not imply sameness of comparison process. The comparison processes used could be markedly different. This makes more precise Tirrell's (1991) observation that even if metaphor and simile both prompt mental comparisons, they might have different "analyses."

The neglect of the different-comparison possibility seems to result from fixation on the particular, traditional

13

class of comparison theories, ones under which be-form metaphor is *elliptical simile*. The view is often traced back to Aristotle's *Rhetoric* (though see Fogelin 2011: 28ff for a caution). Some modern instances are as follows. Miller (1979: 381) says that "understanding [a be-form] metaphor requires the reinsertion of 'is like' ...". Levinson (1983) says "metaphors are similes with suppressed or deleted predications of similarity." Glucksberg (2001) says that metaphors are essentially "implicit similes," and Glucksberg & Haught (2006b) say: "In the comparison view, metaphors are understood in terms of their corresponding similes." Chiappe, Kennedy & Smykowski (2003) assume that comparison theory treats metaphor as elliptical simile. Fogelin (2011) more carefully says that metaphors and similes are both forms of figurative comparison, and a be-form metaphor means figuratively what the corresponding simile means figuratively.

Given a metaphor-as-elliptical-simile view, it is an obvious corollary that the same comparison process is involved for be-form metaphor and simile. But ellipsis-based views are merely a tradition-enshrined special case of what comparison theory in general is about or can be about.

Interestingly, while the comparison aspect of CoM does *not* explicitly rely on ellipsis, CoM does not consider the different-comparison possibility. This is because the underlying structure-matching theory attempts to address all types of structural alignment, whether involved in metaphor, simile, explicitly analogical statements or literal comparisons, in an essentially uniform way, although with varying emphasis on attributes *versus* relational predicates. But the CoM comparison mechanism is just one amongst many possibilities.

I do not specifically claim that be-form metaphor and simile involve different comparison processes. Rather, the sheer possibility of significant difference completely vitiates a certain type of argument for categorization theory. Some experiments have indicated that different features of T and/or S can be mentally stimulated by a be-form metaphor as opposed to its corresponding simile. Notably, a simile is more likely than metaphor to activate low-level features of *shark* such as *can-swim*, and metaphor is more likely to stimulate emergent features (Bowdle & Gentner 2005, Glucksberg & Haught 2006b). Such results have been used as evidence that the be-form metaphors are not understood by comparison (Glucksberg & Haught 2006a,b, Glucksberg 2008). But the argument tacitly assumes that a be-form metaphor and corresponding simile would require the same comparison process, or that any differences are insignificant. But there is no reason why one particular comparison process could not, say, contain a preference for high-level over low-level features.

5 Towards an Alternative Battlefield

There are various dimensions of potential variation between and within processing accounts of simile and metaphor understanding that should be considered as an addition to, and possibly as a replacement for, the categorization/comparison distinction. These dimensions are more sharply definable than that distinction is, and are more directly and clearly related to the experimental findings. The dimensions cut across the C/C distinction: for a particular point in the space they define, a suitably framed comparison theory or categorization theory can in principle lie there. The dimensions we will discuss are *Target/Source Contribution-Disparity*, *Target/Source Mediacy-Preservation,* and *Target/Source Mediator-Carefulness*.

5.1 Dimension 1: Target/Source Contribution-Disparity

This is about the possibly different types and extents of contribution that the T and S terms make to the process of understanding a simile or be-form metaphor. We will mainly consider just one part of the process, namely the determination of the *mediator*. The mediator in the case of comparison is the similarity or analogy (system of correspondences) that is found, and in categorization it is the superordinate category found. We use "mediator" rather than the commonly-used term "ground" in order not to engage with past notions of ground.

The mediator mediates between the T concept and literal S concept. But the notion does not assume that the latter is necessarily accessed during understanding. For instance, in the case of a conventional term S, possibly only a remembered superordinate category S* is accessed directly, the literal S concept being ignored. But we still call S* a mediator as it is still superordinate to the T and literal S concepts. Relatedly, "determining" the mediator may range from just directly accessing an existing S* to elaborate category-discovery/creation or analogy-calculation.

The notion of salience imbalance (Ortony 1979) counts as one specialized form of contribution disparity proposed to hold in metaphor understanding, because a difference is claimed as to what sorts of properties the source and target contribute. Disparity more generally is mentioned under another name by Wolff & Gentner (2011), who effectively call a process without disparity "role-neutral" and one with disparity "role-

15

specific," alluding to the target and source roles in metaphor.

Long-term research questions include: Do the S and T terms contribute to the mediator-finding process to roughly the same degree, or is the process markedly more dependent on one of them? Are the T and S contributions qualitatively similar or markedly different? How do the extents and types of contribution depend on the particular T and S involved, on form (simile or be-form), and contextual factors?

Now, there is a highly salient contrast between existing comparison and categorization accounts: categorization theories have assumed much more contribution-disparity than comparison theories have. The determination of a partial structured analogy between T and S in, for instance, CoM, is usually assumed to be entirely symmetrical. The same mappings (in reversed form) would arise if those concepts were interchanged. This symmetry arises because the concepts are treated in an equal way. The symmetry is the crux of, for instance, Wolff & Gentner (2011) experiments. But, by contrast, in existing categorization theories the determination of S* uses S and T very differently. The disparity is most evident when suitable superordinate categories are already connected in the understander's mind to S (term or concept). Plausibly, the process checks whether one of the available categories fits the target side and is contextually appropriate. We see a clear contribution disparity: S supplies categories, T merely supplies a filter. Topic dimensions arising from T and context (Glucksberg 2001) could provide a quicker check on the suitability of a superordinate category, or could actually drive the choice of one. But still we have a major disparity: the topic dimensions do not themselves offer a category for consideration. If T and S were interchanged, a very different superordinate would typically be found. This may explain what happens when metaphors are reversed: see, e.g., Campbell & Katz (2006), Chiappe, Kennedy & Smykowski (2003) and Glucksberg (2001). A symmetrical-comparison theory of metaphor needs to appeal to other aspects of understanding, such as candidate inferences, to explain reversal effects (see especially Wolff & Gentner 2011).

When already-available categories S* fail to be appropriate, then different, possibly less disparate possibilities arise. Glucksberg (2001, 2008) implies that a process of comparison may be needed (see also Kennedy & Chiappe 1999). But categorization theory need not hastily resort to comparison (a point that even categorization theorists have not exploited very much). For instance, it may be possible to take a common generalization of two or more existing superordinates (if any), by abstracting away from their non-shared details. This is a highly contribution-disparate process.

16

Despite all the above, the contrast between comparison and categorization theories as regards disparity is not a necessary one. There is nothing about categorization in itself that dictates that T and S should be treated disparately in finding S*, even if comparison is not used. (The practice of using a star label such as S* in categorization accounts, including relevance-theory ones, alluding notationally only to the source, is itself misleading. It would always have been theoretically more neutral to use a label such as [T/S]*.) Equally, there is no reason why an analogy-finding process cannot treat T and S very differently (as Chiappe & Kennedy 2001 also imply). For example, partial-analogy finding might have more tolerance for leaving out information on one side than on the other.

Also, what specific analogy is found in the time available may depend on whether the process is guided by looking first at target features to see if they correspond to source side ones, or vice versa, or pursuing some other procedure. If there is a strong time limitation, and more than one (partial) analogy is possible, then source/target disparity in the way the process is guided could lead to different analogies being found depending on which way round the metaphor is, even if there were no dependence given unlimited time. Indeed, the "greedy," pragmatic version of SME described by Forbus & Oblinger (1990) does refrain from finding all possible analogies, to improve efficiency, and is guided towards pragmatically fruitful ones by a target-side query, thus building disparity in.

But disparity in comparison does not depend on a sequential element in processing: the connectionist ACME system (Holyoak & Thagard 1989) for comparison, which has been applied to metaphor, allows disparate pragmatic guidance through enhanced activation of any collection of source or target elements that are deemed pragmatically important (Holyoak, Novick & Melz 1994). The enhancement is via extra activation of all source/target element correspondences that involve such an element.

So, both comparison and categorization can exhibit contribution-disparity in mediator-finding, and both can avoid it. Thus it is disparity or the lack of it, not the comparison/categorization distinction, that generates certain experimental results, most obviously reversal results. For example, Wolff & Gentner (2011) find that early stages of metaphor understanding are insensitive to reversal (hence, no important contribution-disparity) whereas later stages are sensitive to it. The early insensitivity is consistent with their symmetric comparison process—but also with categorization lacking contribution disparity in S*-determination.

As for form differences:- If, say, a disparate process of comparison is used for be-form metaphor but a non-disparate version for simile, then experimental results concerning be-form will be more affected by reversal than will results concerning simile form—as indeed found by Chiappe *et al.* (2003). But some experimental work, e.g. Gokcesu (2009), failed to find a simile/be-form differential effect on acceptability of reversals, and Campbell & Katz (2006) showed that context may strongly influence reversal effects. The central point for us is not what reversal effects actually hold, but merely the point that results that have been claimed to reveal a C/C distinction can instead result simply from differences in the extent of contribution disparity achievable either by comparison or categorization.

Plausibly, the more conventional a particular source item is, or that a target/source pairing is, the more the understanding process will show contribution disparity, because the more it can simply access one of the entrenched senses. The more novel the source or pairing is, the more the target *might* play a role more like that of the source. This is consistent with Gokcesu's (2009) finding that conventional target/source pairings are more affected by reversal than novel ones. However, novelty is not a reliable predictor of significant novel investigation, whether by comparison or categorization. Even some metaphor or simile involving a completely novel source can sometimes be simply understood by appeal to existing associated properties, categories, etc. Thus, to give an extreme example to make the point, someone might say "My TV is [like] a heap of rotting marmoset tails" [source-term novelty confirmed by web search in January 2012]. Nevertheless, the undesirability inferred from rotting marmoset tails without any, or only minimal, guidance from the target could be carried over immediately to the TV. This does not preclude other properties being ascribed through more complex, target-reliant processing, but in context it could well be that the undesirability is all that is relevant. Thus some novel cases may involve great contribution disparity.

The disparity dimension as defined above is about how the T and S *terms* affect understanding. But we should, at least theoretically, decompose this effect into two stages. The first is the transition from T term to T concept, and, similarly, S term to S concept (if used: understanding might proceed directly from S term to an S* category when available). The second stage is the way the T *concept* and S *concept* (if used) affect understanding. This decomposition inspires two important caveats.

First, even if the comparison or categorization process itself exhibits no disparity, there could be disparity between the T and S term-to-concept transitions. For example, a more detailed conceptual structure might

tend to be created for T than for S. This is a salient possibility to consider. T is typically the more important item to glean new information about, and also much detailed information about T may already be available, whereas only schematic information about S might be brought to bear. Thus, reversal of the sentence could change the detailed conceptual structures that the comparison or categorization process works upon: structures S′ and T′ rather than the original T and S structures respectively. Reversal could thus still cause a difference in mediator found. This effect is potentially excluded by Wolff & Gentner's (2011) finding of reversal-insensitive early processing. But if the effect does remain a worry, more attention to the term-to-concept transitions will be needed in the Debate. The Debate usually (and usually tacitly) assumes that terms lead to the same (literal) concepts whichever side of a simile or be-form metaphor they are on. It could also mean that our Contribution Disparity dimension should be replaced by at least two: one about the T and S concepts, and one about the term-to-concept transitions.

The second caveat is about the relationship between contribution-disparity and notions of asymmetry. If a mental process such as mediator-finding has absolutely no contribution-disparity then it is automatically symmetric in the sense that interchanging the target and source concepts will lead to the same result. Hence, if there is result-wise asymmetry there must be contribution disparity. However, if there is disparity then result-wise asymmetry will only be a strong tendency, not a definite outcome, and contrapositively if there is result-wise symmetry then there will only tend to be a lack of disparity. For instance, disparity in how the two sides guide analogy finding may not actually affect the result if enough time is available for all useful analogies to be found.

Finally, we can briefly consider disparity outside of mediator-finding. Existing comparison and categorization accounts differ less here, both tending to considerable disparity. In comparison, we have, notably, the projection of candidate inferences. For categorization, the situation is vaguer but authors generally refer to the T concept being augmented by inheritance of features of S*.

5.2 Dimension 2: Target/Source Mediacy-Preservation

This dimension is relevant only to simile and to cases of be-form metaphor which are not held to be understood *merely* through direct access of some conventional metaphorical meaning from the S term, neglecting

the literal S concept. The type of be-form metaphor that is mainly of interest is where the literal S concept is activated and compared to T or used to help find an S*. But the dimension does also have relevance to conventional metaphor where the literal S concept, and not *only* the directly-accessed metaphorical meaning, is activated (as suggested by Gibbs 2011 in his embodiment-based view of metaphor).

In either simile or be-form metaphor whose understanding uses the literal S concept, some relationship is established between the literal S concept and the T concept. In analogy-based comparison such as in CoM, the mediator is a system of correspondences between aspects of S and T, so is itself a relationship between those concepts. In categorization theory, the relationship is the fact of co-categorization of S and T within the mediator, S*. Because of these close connections of the S/T relationship to the mediator, we will call the relationship the S/T *mediacy*. So, in the comparison case the mediacy is just the mediator, whereas in categorization it is a little more.

The Mediacy-Preservation dimension is intuitively about the following issue. Does the main information that an understander gleans from a 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 T/S mediacy itself? To put it another way, is the mediacy *merely* found as a stepping-stone to information about T in its own right, or is it *preserved* as part of the meaning?

We can put the issue a little more usefully as follows, avoiding difficult linguistic/philosophical issues about what "meaning" is. When understanders have understood a simile or be-form metaphor, they presumably have derived some new information asserted about T by the utterance and/or have cognitively highlighted some known aspects of T. Let us call the new information or highlighting the "comment" on T derived via the mediator. For some short or long period this T-comment will be active in the understander's mind.[1] The question is whether the understanding process also makes the mediacy itself keep a comparably high level of activity during that period, or instead allows the mediacy activation to drop away once the comment on T is determined.

Plausibly, simile usually has a stronger mediacy-preservation tendency than (be-form) metaphor has, for a

---

[1]Here we appeal to the normal widespread assumption in psychology that meanings of sentences, features or concepts of which meanings are built, etc. are in some sense more "activated" in an understander's mind than they would normally be. We also assume highlighting is a matter of greater "activation."

given T and S. Consider what the simile and corresponding metaphor are *apparently saying*. This apparent-saying is the core of Grammatical Concordance (see section2). The simile is apparently saying that T and S, taken literally, are alike, presumably in some contextually relevant sense to be uncovered by the understander. The corresponding metaphor is apparently saying that T is to be categorized under S, taken literally. But only in simile does reality match appearance: the point is to convey that T and S are indeed alike in certain respects, if only perhaps in a very deep, abstruse sense. But the point of the metaphor is, of course, not to convey that T is within the literal S category. At most, it conveys that T is within a different category, S*, if we follow a categorization account. Thus, in the case of simile, the understander fleshes out a stated likeness in some particular way, *accepting* the apparent likeness as something that is actually being asserted, whereas with be-form metaphor the understander must *discard or bypass* what it apparently asserts and replace it with something else: a different categorization, or a comparison, for instance.

This suggests that in simile the mediacy itself tends to be a central part of the meaning of (or main information gained from) the utterance, whereas in be-form metaphor it tends merely to be a stepping-stone to the T-comment. This is an intuitive point about meaning rather than cognitive activation, but the more that any information is taken by the understander to be part of meaning the more, presumably, its mental activation will be preserved.

As for experimental evidence, Krennmayr (2011) studied the extent to which subjects include (what I regards as) source/target mediators in their understanding, and found that use of simile form encourages such inclusion. The findings below about "double predication" are also consistent with this claim.

However, metaphor understanding may also sometimes preserve the mediacy in the meaning. Stern (2010) argues this in discussing "knowledge by metaphorical character." Barnden (2010) shows how such preservation (called there "survival of the source/target linkage") needs to hold in certain types of example.

A further comment is needed about simile asserting a real likeness, not just appearing to do so. Many authors within the Debate and outside, including Glucksberg (2001) and Tirrell (1991), distinguish between literal and figurative comparison statements, only the latter being similes. The distinction may only be fuzzy if tenable at all, (cf. Carston & Wearing 2011, O'Donoghue 2009), but in any case it seems to be based on the types of similarity seen between the two sides, not on whether there is indeed similarity (likeness). Now,

many authors (including O'Donoghue 2009, Ritchie 2006 and Tirrell 1991) say furthermore that in simile, as opposed to literal comparison, allegedly shared features are often not truly shared, e.g. that for "Juliet is like the sun" the feature of warmth might be cited, whereas of course different sorts of warmth (one itself metaphorical) are involved. This might make it seem that actually there is no genuine similarity in such simile. However, even if it were correct that there is not, the Debate at least does cast simile as conveying genuine similarity. For example, the comparison process such as that in CoM is never portrayed by either its proponents or its detractors as not *really* achieving comparison or determining similarities (the shared structures found, the common abstractions derived, etc.).

But anyway in cases such as as Juliet/sun, if we dig deep we *can* find a genuine, if possibly subjective and psychological, sharing of properties. For instance, the sharing could be that Juliet causes (in the speaker at least) a strong feeling of emotional pleasure that is the same as that caused by physical warmth of the sun. Even in a yet more highly figurative simile such as "Love is like a banana," some actual property sharing is conveyed, though it may be very abstract and difficult to formulate. In this example, the likeness could be that both are good for you over a period of time that is significant with respect to the time-scale in its subdomain of human life (the subdomain of important relationships in life in the love case, and the subdomain of eating in the banana case), or that you have to apply considerable effort (e.g. in understanding the main features of the loving situation, and in peeling the banana) before you can get the benefits.

We now turn to novel T/S pairings *versus* more familiar ones, irrespective of whether within simile or within be-form metaphor. The more familiar the pairing, the less reason for the understander to suppose that the utterance is pointing out a relationship of T to the literal S concept rather than just commenting on T; hence the less appropriate it is to preserve the mediacy along with the T-comment. For instance, if the understander is already accustomed to organizations of various types being cast as being, or being like, dictatorships, then the fact that some particular organization of those types, or even an organization of a new but related type, is cast as (like) a dictatorship is in itself relatively uninteresting. On the other hand, the more novel the pairing, the more natural it is for the understander to suppose that the utterance is not only commenting on T but also pointing out an unfamiliar linkage between T and the literal S. For instance, if an utterance casts someone's garden as a dictatorship (e.g., because it is dominated by one large tree) the understander is liable to find the relationship to literal dictatorships interesting. The argument applies equally to similes and be-form

metaphors, but it is possible that in simile there is a higher default level of preservation to begin with.

The comparison/categorization distinction appears tangential to the mediacy-preservation dimension. There does not appear to be anything about comparison or categorization in general that would dictate an extent to which they preserve or fail to preserve mediacy. Yet the extent of mediacy-preservation that exists in a particular case of understanding could have important relationships with experimental effects studied in the Debate.

For example, Bowdle & Gentner (2005) and Gokcesu (2009) report experiments looking at frequency of *double predications.* This is where an experimental subject couches an interpretation in terms of a shared predication about both target and source, rather than just about the target. An example is provided by "a child is a snowflake." The interpretation "Both are unique" is a double predication whereas "a child is unique" is a *single predication.* People's interpretations of novel similes and metaphors were found to involve more double predication than their interpretations of (relatively) conventional statements. Those authors propose that double predication is a symptom of comparison, single predication a symptom of categorization. But an alternative explanation is just that single predication is a symptom of treating the source/target mediacy as a stepping-stone and double predication is a symptom of including it in the meaning.

Krennmayr's (2011) findings suggest that novelty, as well as use of simile, encourages inclusion of the mediator in subjects' mental representations of meaning. In fact, she found that there was no (two-way) interaction with use of simile, so that novelty may affect mediacy-preservation largely independently of simile use. Nevertheless, this does not prevent novelty affecting form preference, and other matters, via mediacy-preservation. To the extent that greater novelty of a pairing encourages greater mediacy-preservation, and to the extent that simile form is more aligned with mediacy preservation than be-form is, we would expect relatively novel pairings to lead to simile form being preferred, or leading to faster processing, etc. This is consistent, for instance, with the following findings: that novel sources cause slower processing, and also more difficulty for children, in be-form than in simile form (Bowdle and Gentner 2005); that subjects prefer simile form more strongly for novel than for conventional figuratives (Gentner and Bowdle 2001); that novel be-form metaphors make less sense to people than the corresponding similes do (Gokcesu 2009); and that conventionality has a significant role in influencing form preference towards be-form (Utsumi 2007).

Now, a key sub-Debate has been whether experimental effects that have been attributed to the dimension of novelty (of sources or source/target pairings) *versus* conventionality (of sources) or familiarity (of pairings) should instead be attributed to other dimensions such as aptness or amount of similarity. Contributions to this sub-Debate have included Chiappe, Kennedy & Smykowski (2003), Gentner & Bowdle (2008), Glucksberg (2008), Glucksberg & Haught (2006a,b), Gokcesu (2009), Jones & Estes (2006), Pierce & Chiappe (2009), Thibodeau & Durgin (2011), Utsumi (2007) and Xu (2010). Therefore, it is difficult to say definitively at present how mediacy preservation's loose connection to novelty fits with experimental results overall.

However, there is one type of argument in the sub-Debate that is worth addressing briefly here, while meriting a longer discussion. It is about utterances such as "My lawyer is [like] a well-paid shark" where the source term contains a modifier that is appropriate for the target but not the literal source, and makes the source term as a whole novel even if the noun has conventional metaphorical uses. The findings of Glucksberg & Haught (2006a,b) indicate that such statements are preferred in be-form, rated more apt in that form, and quicker to process in that form: in short, the be-form statements are not so felicitously paraphrased in simile form. Glucksberg and Haught (*ibid.*) claim that the results damage other authors' claims that novelty encourages preference for simile form. The paraphrase failure is claimed to arise because, taken literally, "a well-paid shark" does not make sense.

However, the argument is flawed. Certainly, it is natural to take the be-form metaphor as if it were paraphrased as "My lawyer is a shark and is well-paid" (see O'Donoghue 2009), giving an easy interpretation to the extent that the first, metaphorical, clause is easily interpreted. And a similar paraphrase for the simile case as "My lawyer is like a shark and is well-paid" does indeed seem less plausible (in fact O'Donoghue *ibid.* claims it is not allowed at all—but that seems too strong). Indeed, if it were as readily available as with be-form metaphor then presumably the above experimental results would not have arisen. So, the understander is led to make sense of "a well-paid shark" as a source term. Now, Glucksberg & Haught (*ibid.*) and O'Donoghue (*ibid.*) assume that the only interpretation route for this term is to take it literally. But, of course, it can itself be taken as an embedded adjective-noun metaphor, as follows.

The adjective "well-paid," precisely in virtue of suggesting a person, by itself encourages an interpretation of "a well-paid shark" as if it had been "a person who is a well-paid shark," which is akin to the be-form metaphor "Some person is a well-paid shark." So, let's say "a well-paid shark" provides a metaphorical

meaning such as "a person who is aggressive and well-paid." It does not matter to our discussion what the mechanism used for this step is. It could be categorization, for instance using an existing shark* category. But now the understander is faced with seeing how the lawyer could be *like* a well-paid aggressive person.[2] This is linguistically a non-felicitous comparison, because a lawyer could well just literally *be* a well-paid aggressive person. Thus, in the simile case the understander not only has to try to undertake (a) a comparison (between lawyer and well-paid aggressive person) as well as (b) the metaphor-processing step within the simile source term, but moreover this comparison is infelicitous. In contrast, the be-form metaphor "My lwayer is a well-paid shark" only requires a metaphor step (lawyer as shark) and an untroublesome attribution afterwards of "well-paid." That combination is much like merely step (b) in the simile case. This contrast can easily explain the experimental results. They are nothing directly to do with the *novelty* of the source term, or with what mechanism is used for any of the metaphor steps in question, but rather just with the addition of an extra layer of processing complexity. This layer arises through embedding a metaphor within the source term of a like-form sentence.

Note that, in the be-form case too, the vehicle term could alternatively be interpreted metaphorically to mean a well-paid aggressive person just as in the simile case, leaving the lawyer just to be categorized straightforwardly and literally as a well-paid aggressive person. On this analysis the "is" in the sentence provides no further figurative effect. This comprehension route is again cheaper and more felicitous than that needed for the simile case.

Although phrasal metaphor such as adjective/noun metaphor is not much discussed in the Debate, Utsumi & Sakamoto (2007) is an exception. Goatly (1997) discusses adjective/noun metaphor at some length in a linguistic vein.

---

[2]The phrases "a well-paid shark" and "a person who is aggressive and well-paid" are assumed here to be read predicatively (attributively) rather than referentially, in common with usual assumptions in the Debate and elsewhere about how source terms in similes and be-form metaphors are generally read. However, Glucksberg & Haught (*ibid.*) appeal crucially to a referential reading of the source term in another example, "[Some company] will be like the next Enron." Going into this matter would require further lengthy discussion.

5.3 Dimension 3: Mediator-Carefulness

The dimension discussed here is the extent to which the understanding process is careful in determining an appropriate mediator, though not necessarily with the understander being conscious of being careful. The degree of carefulness is the degree to which the process is thorough in exploring possible mediators (existing or new) in an attempt to make the best sense possible of the utterance in context. What best sense amounts to is a highly variable and context-sensitive matter, and might for instance be cashed out in terms of an optimal level of relevance as proposed in Relevance Theory (whose application to figurative language is discussed in Sperber & Wilson 2008 and Carston & Wearing and 2011).

If an appropriate mediator is already available, then merely accessing it in some straightforward way and then adopting it without looking further would show a low level of carefulness. Thus, in particular, just directly accessing a single entrenched metaphorical meaning shows little carefulness. Accessing known target/source mappings of the sort proposed in conceptual metaphor theory (Lakoff & Johnson 2003) only illustrates carefulness to the extent that it may take exploration to see which conceptual metaphor is appropriate and how. More generally, making a considered choice between several available mediators constitutes an appreciable extent of carefulness. Attempting to work out a new mediator even though one is already available shows a higher level of care, and considering several possible new mediators shows yet greater carefulness. Another aspect of care is effort put into evaluating how appropriate the mediator is (depending on the needs of context and on the understander's purposes): e.g., evaluating the degree of similarity achieved, or establishing which candidate inferences actually apply to the target.

Ultimately it may become necessary to distinguish, more than we will do here, between different ways of being careful, e.g., what spaces of possibilities are relatively thoroughly explored, or what sorts of checking of the mediator are done. They do not necessarily lead to the same effects and could even lead to some conflicting ones. Similarly, they could arise from different aspects of metaphors and similes. Some may empirically turn out to be more important than others.[3]

Carelessness, in tolerating less appropriate mediators, may lead to less understanding in one sense, but on the other hand allows licence to be more open-ended in what features, feature-correspondences, etc. are

---

[3]The account here of carefulness already restricts to just one aspect of care as considered in Barnden (forthcoming).

used in the mediator. Carelessness could thus lead to more creative mediators in one sense. On the other hand, carefulness could help to uncover non-obvious possibilities, such as deep analogies that do not involve much superficial similarity, supporting a different dimension of creativity.

The dimension has an interesting relationship to the previous one, Mediacy-Preservation. In principle, a carelessly-determined mediator could be strongly preserved and a carefully-determined one could be ignored as soon as the T-comment is found. However, mediacy-preservation plausibly has considerable positive correlation with mediator-carefulness. In particular, if the mediacy is preserved for the particular reason that it is included in meaning as opposed to being just a stepping-stone to the T-comment, then the understander can be expected to apply more care. Thus, to the extent that be-form metaphor involves lower mediacy preservation than simile does, as suggested in the previous subsection, we conjecture that be-form metaphor will tend to lead to lower mediator-carefulness than simile will. For example, the understanding process will be more easily satisfied with determining and stopping with any mediator that provides a T-comment relevant enough for the discourse purposes at hand, and is not constrained by concerns about the quality of the mediator itself. Thus, as a broad generalization, less carefulness will be applied for be-form metaphor than for simile.

This careless tendency in be-form metaphor does not contradict intuitions about such metaphor being more emphatic or shocking than simile (at least in novel cases) because of more directly juxtaposing source and target. Emphasis might place more pressure on the understander to come up with *some* understanding, but it does not of itself imply a high quality-level for the mediator or for understanding. While one effect of emphasis might be to make the understander more interested and therefore apply more care, an opposite effect could even be to encourage quick, ill-considered understanding. But relative carelessness is merely an overall tendency of metaphor: a sufficiently novel metaphor could well be a case where there is high carefulness because of causing increased interest.

Metaphor's tendency claimed here towards lower carefulness may also seem at first not to sit well with comments in the previous subsection about apparent-saying. Surely, if metaphor provides less guidance than simile as to what sort of mediator is involved, one needs to explore more possibilities in order to work it out, and would that not imply *more* care? However, precisely in being relatively careless, the understander is at liberty to limit the need to explore multiple possibilities, and instead to pursue whatever line of processing

27

looks initially most promising and to stop when an adequate T-comment is found. Higher availability of possibilities does not force more investigation of possibilities.

A further, related, consideration is that it may be more difficult to detect that a given be-form utterance is metaphorical than it is to detect a simile. Thus more care is needed in some sense. However, this is a different dimension of care from the one we are considering, which is about care applied either on the implicit or explicit assumption that, or after it has been determined that, or while considering whether, the utterance is metaphorical. Clearly, the detection issue is important in its own right and could affect experiments, but it is not much addressed in the Debate and is not covered in this article.

Mediator-carefulness differences could lie behind experimental results (Glucksberg 2008) that in simile the understander has more of a tendency than in metaphor to go back to low-level properties of the literal source term rather than (just) using available superordinate categories/properties associated with that term. More generally, a simile can lead to different interpretations from its corresponding metaphor, even for a conventional T/S pairing (Glucksberg & Haught 2006a). These findings would follow from the understander tending to be more careful about the mediator in simile, as argued above, thereby being predisposed to seek a new mediator, considering S and T afresh.

Again, if simile generally leads to higher mediator-carefulness, we have an explanation of Gregory & Mergler's (1990) result (cited by Bowdle & Gentner 2005) that similes are more likely than be-form metaphors to highlight non-obvious similarities, because the process is more disposed towards thorough investigation of possibilities. Similarly, we have an explanation of Roncero, Kennedy & Smyth (2006)'s result that simile form tends to be employed when the meaning is unusual or difficult (even when the metaphor form is conventional), and be-form metaphor to be employed when the meaning is usual or easily comprehended. A difficult or unusual meaning takes more care in order to be discovered. It also could explain Jones & Estes's (2006) finding that novel similes are comprehended more slowly than novel metaphors, to the extent that more care generally implies slower processing.

O'Donoghue's (2009) claims are very much in line with our consideration of carefulness. She claims that simile often gives distinctive effects as compared to metaphor because simile invites "more measured reflection" of the points of comparison between the two concepts, and metaphor "involves less direct online

contemplation of the precise terms which have prompted the juxtaposition of the two concepts..." However, our discussion above more strongly brings out that both metaphor and simile are capable of different levels of carefulness (depending on context and the particular source and target terms). Also, we explicitly analyze the connection between simile form and carefulness as being indirect, via mediator-preservation in meaning.

Finally: Mediator Carefulness cuts across the comparison/categorization distinction. Either type of process can be arbitrarily careless or careful, and the level of care can dynamically change according to circumstances.

5.4   Mediator Carefulness and Interpretative Diversity

Utsumi (2007; see also 2011) studied the *interpretive diversity* (ID) of target/source pairings, such as the pairing *anger/sea* (in their Japanese versions). ID measures the "semantic richness of the figurative interpretation of" a pairing. Utsumi found that a pairing's ID was correlated positively with preference for be-form metaphor over simile for that pairing, and with the relative comprehensibility of be-form. High-ID pairings were equally comprehensible in be-form and simile form, whereas lower ID ones were more easily comprehended in simile form. Indeed, he found that ID was a more important factor in such be-form/simile differences than aptness, similarity and conventionality. Given the prominence of those factors in the Debate, Utsumi's study is potentially of great significance. We will seek to illuminate his results using the carefulness dimension.

The ID of a target/source pairing depends on "both the number of features involved in the interpretation and the uniformity of [the] salience distribution of those features." Intuitively, according to various statements in Utsumi (2007, 2011), a be-form metaphor or simile using the pairing mentally activates a set of features that are shared between source and target, with different levels of mental salience. ID does not use any absolute values of salience, but instead assumes that each feature has a relative salience value between 0 and 1, with all the values for a given pairing adding to 1. By then combining the relative salience values for the pairing in an entropy-like formula, we get an ID value.[4] The higher the number of features associated with a pairing the higher the ID; and the more uniform the relative salience of those features the higher the ID.

---

[4]The ID of a pairing is the sum, over the features mentioned by subjects, of $S_f log_2 \frac{1}{S_f}$ where S_f is the relative salience of feature f. NB: Utsumi just uses the term "salience," but "relative salience" is clearer.

By way of example, when there is just one feature, ID takes on its minimum value of 0. When there are n features all of equal salience, ID takes its maximum possible value (for that n) of $log_2 n$. For a given n, say 10, ID will be low for a pairing with a very non-uniform salience distribution, such as when two have relative salience of 0.4 and the remaining eight having the same low relative salience of 0.025. It must be remembered here that we are dealing with *relative* salience values for a given pairing, and these cannot be compared between pairings. It is only the different level of uniformity that matters.

We now proceed to comment on how Utsumi (2007, 2011) seeks to explain the connection between ID and his comprehensibility and form preference results. We will argue that considerations of mediator carefulness provide a better connection. Afterwards, we will also comment briefly on how ID is experimentally measured.

First, we must caution that Utsumi tends to use the terms "meaning" and "feature" interchangeably, and often seems to take an individual meaning for a metaphor/simile to consist of one feature, though at other times talks of an interpretation involving several features. In effect, he conflates the case of multiple features within a single interpretation and the case of a person's alternative interpretations having different features. This conflation is also evidenced by the way ID is experimentally measured, as explained below. The conflation does not explicitly affect Utsumi's own reasoning, and we can understand his claims by assuming that a pairing activates, with degrees depending on their different saliences, all the different features the person associates with a pairing, irrespective of whether those features are within the same or different interpretations. (This view is consistent with his model of understanding analyzed in section 6 below.)

Utsumi (2007, 2011) addresses the connection between ID and his comprehensibility and form preference results by arguing that high ID allows understanders to perform categorization whereas low ID requires them to use comparison, and then appealing to Grammatical Concordance to align categorization with be-form and comparison with simile form as usual. However, the argument is only brief and partial, for instance in not properly addressing salience uniformity. He seems merely to assume, without argument, that the features involved in categories are relatively uniform in salience, and that low uniformity does not allow adequate categories to be accessed. Also, he seems tacitly to assume that comparison is not made more difficult by low uniformity.

But we can argue alternatively as follows. First, we argue that mediator-carefulness tends to oppose ID: i.e., lower/higher carefulness encourages higher/lower ID, respectively. If the level of care taken is relatively high, there will arguably be a higher quality bar on features activated, making fewer features seem relevant than would otherwise be the case. At the same time, the features will tend less to be just the more salient ones (the ones that would be chosen with little care). Thus, there will be a higher spread of absolute salience levels than with a more careless process. This then leads to the *relative* salience levels also being more spread out, i.e. less uniform. (This argument assumes that the more careful process does keep at least one high-salience feature.) Hence, greater carefulness tends to lower both the number of features and their relative-salience uniformity, thereby lowering ID.

Secondly, recall from section 5.2 our conjecture that simile tends to lead to higher carefulness than be-form metaphor does. Putting this together with the argument just above, we have that simile tends towards lower ID than be-form metaphor does. This potentially explains Utsumi's findings about low ID being associated with advantages for simile, without needing to consider categorization versus comparison or make assumptions about how those processes relate to numbers and saliences of features.

We now comment briefly on Utsumi's experimental estimation of ID values. The method reflects the conflation mentioned above of multiplicity of features within an interpretation and multiplicity between interpretations, where indeed now the different interpretations can be by different people. For a given target/source pairing, the relative salience of a feature is simply the popularity of that feature with the experimental participants when they give interpretations of metaphors or similes involving that pairing, compared to other features for that pairing: it is the number of mentions of the feature by the participants divided by the total number of feature-mentions by the participants.[5] Now, the participants were asked to produce an interpretation, listing at least three features, so we can grant that normally each participant does produce one coherent interpretation. But, of course, different participants often produce different interpretations, with different sets of features.

Nevertheless, the use to which Utsumi puts ID and his general comments about ID (e.g., that it measures "semantic richness" of interpretation) suggests that he assumes that ID measures the potential for a typical

---

[5]The notions of shared feature and salience used here are those used in Experiment 2 in Utsumi (2007) and mentioned again in Utsumi (2011). Somewhat different notions are mentioned for Experiment 1, with salience there bringing in participants' own subjective salience ratings.

individual understander to produce a diversity of features for a given metaphor or simile. Utsumi's way of measuring ID is reasonable if we can assume that different participants have in their minds roughly similar interpretations of similes and be-form metaphors such as those in his experiments, but that in the experiments the participants only draw upon one or a small subset of these. It also helps to consider that individual features they mention could potentially be interpretations by themselves.

6 Discussion of Utsumi's Semantic-Space Model

An interesting application of various themes of this article can be made in a specific way in the case of Utsumi's (2011) model. He provides precise, simple, computer-simulated algorithms *Compa* and *Categ* intended to be construed as performing comparison and categorization respectively, within a unified cognitive framework. Utsumi (2011) relates the model to interpretive diversity (see section 5.4) but we will not be commenting on this aspect.

Utsumi's model is based on the Predication Model of Kintsch (2000, 2001). This uses a semantic space of high-dimensional numerical vectors, where the particular vectors that are initially included correspond to words (lexical word-forms). Intuitively, each vector encodes something about the meanings of a word, although the vector merges together the different senses of the word. The vectors are derived by computation over word occurrences in a language corpus. (The details of this are unimportant here.) Individual vectors in the space, whether corresponding to words or not, can be viewed as representing properties, categories or concepts, depending one one's theoretical predilections. Also, *sets of* vectors can be viewed as defining categories (cf. Utsumi 2011).

Both Kintsch and Utsumi apply the approach to understanding be-form metaphors linking two words. One method for doing this in Utsumi (2011) is the *Categ* algorithm, borrowed from Kintsch (2001). In broad strokes this involves computing a new vector—the *metaphor vector*—from the target word T's vector and the source word S's vector. The metaphor vector is considered to encapsulate the meaning of the metaphor. The computation proceeds via intermediate word vectors that are related to both the T vector and S vector. Kintsch (2008) suggests, and Utsumi (2011) more strongly claims, that the algorithm can viewed as a realization of the categorization approach. We can view the intermediate word vectors as representing a

superordinate category S* covering S and T, and the metaphor vector as the result of putting T in S* (although there is an obstacle to this view that will become clear below). Thus, in the case of *A business is a dictatorship* the *business* and *dictatorship* vectors are used to calculate some intermediate vectors. These can be viewed as forming a representation of a superordinate *dictatorship** category that is appropriate for application to businesses. The intermediates together with the *business* and *dictatorship* vectors themselves are then used to compute the metaphor vector.

The Utsumi model also includes an algorithm *Compa* that he views as performing comparison. Again, the algorithm computes a metaphor vector from the T and S vectors, via some intermediate vectors. However, the computation of the intermediates is markedly different from the method in *Categ*, and the way that the final metaphor vector is computed from the intermediates is also significantly different.

The categorization and comparison algorithms can be described as follows, with Fig. 1 and Fig. 2 respectively providing schematic illustrations. The "m-neighborhood" of a vector V for some number m is the vector-set consisting of the m vectors closest to V. "Closeness" or similarity of two vectors is measured using the cosine of their mutual angle. It is important to keep in mind that all the vectors in the space at the start of the processes are *word* vectors: i.e., each vector represents a specific word. However, the metaphor vectors created by the algorithms would only be word vectors through (huge) coincidence. In the following descriptions and elsewhere we will often use T and S to mean the T vector and S vector for brevity. The steps are numbered differently from Utsumi (2011).

Categ (with numerical parameters m and k, where $k < m$):

(1a) Find the the m-neighborhood of S.

(1b) Within that neighborhood find the k vectors closest to T.

(2) Compute the metaphor vector as the centroid (i.e., average) of the k vectors from Step 1b together with *both* S and T.

Compa (with numerical parameter k):

(1) Find the [at-least-]k word-vectors that are *closest ranking-wise to T and S taken jointly*. That is, find the smallest number m such that T's m-neighborhood and S's m-neighborhood have k vectors in common. These k vectors are "common neighbors" of S and T.[6]

(2) Compute the metaphor vector as the centroid of the k vectors from Step 1 together with T (*but not* S).

((FIGURES 1 and 2 ABOUT HERE))

In each case we can take the mediator to be the k word-vectors selected by Step 1 as related to both the T vector and the S vector. In *Compa* this is viewed as a set of common neighbors, in *Categ* as defining the superordinate category S*. In each algorithm, Step 2 is the application of the mediator to generate the metaphorical meaning.

We will now proceed to discuss the algorithms in the light of the three processing dimensions presented above and our skepticism about the categorization/comparison distinction. We first consider the issue of source/target contribution disparity. In *Compa*, the mediator-finding (i.e. Step 1) is completely contribution-equal and symmetric result-wise: the same result would arise with T and S interchanged. But the application of the mediator, i.e., Step 2, is markedly disparate in not involving S. This disparity is needed to account for metaphor-reversal phenomena. *Categ* is the opposite to *Compa* in both respects. It is clearly disparate in mediator-finding (Step 1), while the mediator-application Step 2 is *not* disparate, since the averaging involves both S and T.

Categ is thus an exception to the generalization in section 5.1 that in a categorization approach the use of the mediator is generally contribution-disparate. This is ironic, as Utsumi's (2007, 2011) own informal comments on the role of categories in metaphor understanding give no role to S once S* has been found, in line with other authors' writings on the categorization account. It is unclear why *Categ's* Step 2 is

---

[6]Utsumi does not explain what happens when *more* than k vectors are shared by the two m-neighborhoods when they are first large enough to share at least k vectors. The problem arises because increasing m by 1 can introduce two new, different vectors into the intersection, one in the new neighborhood for T and one in that for S, where the new vector for T is already in the *old* neighborhood for S, and vice versa. For the purposes of our discussion, we continue to adopt the simplifying fiction that this situation never arises. But it makes no difference to our observations if really more than k common neighbors are found instead of exactly k.

different from *Compa's* Step 2, which more naturally captures the idea of applying the found mediator to T. Another view of *Categ* would be to regard the whole algorithm as just computing S*, with Step 1 being an intermediate stage. Our discussion below could be suitably modified. However, Utsumi (2011) does say with regard to the example "A rumor is a virus" that the vectors from Step 1 represent the category of contagious things, i.e. the superordinate category according to a categorization approach, and, following Kintsch, regards the vector resulting from Step 2 as the result of *applying* the predicate (here, being-contagious) to the argument (the rumor).

More importantly, we question the comparison/categorization distinction claimed as existing between the algorithms. Someone who had been told only that *Categ* was a categorization algorithm could be forgiven for thinking that *Compa* was simply another, rather similar, categorization algorithm. It differs mainly in giving equal influence to T and S in building the superordinate category, rather than proceeding in the disparate manner of *Categ's* Step 1. The intuitive motivation for *Compa* being thought of as a *comparison* algorithm may lie in Utsumi's informal characterization of its Step 1 as being a matter of finding k "common neighbors" of T and S. By contrast, he merely refers to the vectors found by *Categ's* Step 1b as those "neighbors" of S that have highest similarity [i.e., are closest] to T—he refrains from saying that the vectors are neighbors of T. But this terminology is highly misleading, as follows.

The first observation is that the metaphorical notion of "neighbor," though mathematically standard, is misleading. As *Compa's* Step 1 illustrates, a neighbor of a vector V can be arbitrarily far from V, because there is no limit to the size of the m in that step. *Any* vector in the space could potentially qualify as a neighbor of V. This is the background for a second, more important, observation, namely that, against initial appearances, there is little basis for saying that the vectors found by *Compa's* Step 1 are any more common than those found by *Categ's* Step 1b. In both cases, they are vectors that are related merely by certain very loose criteria of closeness to T and S: and not even by actual closeness according to the vector similarity measure—but only in terms of *ranking* of actual closeness. The word "common" is presumably used so as to resonate with the notion that a comparison process finds common properties. But there is no in-built criterion in the semantic-space model for saying that one vector has another vector as a property, any more than there is for a vector to be a neighbor of another. The "common neighbors" cannot be classified as such on any basis to do with being particularly close to S and T, whether in terms of of actual closeness

or of closeness rank, because of our first observation about there being no limit on m. Nor are the alleged common neighbors guaranteed to be reasonably uniform in actual closeness to S and T: the vectors that Step 1 produces can differ to an unlimited extent, in principle, in actual closeness to S and T. Just because they are in an m-neighborhood of T and an m-neighborhood of S, for the same m, has no bearing in general on what the actual distances are and how they compare to each other. If word vectors are packed more much tightly round S than around T, for example, S's m-neighborhood will be nestled much more tightly round S than T's is around T, so the intersection of the neighborhoods will be much tighter around S than around T. See Fig. 3.

((FIGURE 3 ABOUT HERE))

Instead of saying that one of the algorithms does comparison and the other categorization, a more fruitful and justifiable way of comparing and contrasting them is as follows.

1. The mediator-finding parts of *Compa* and *Categ* (Step 1 in each case) are on a rich spectrum of possibilities as regards how disparately the T and S vectors can be treated. Indefinitely many other possibilities exist for selecting k vectors given T and S, including methods that combine aspects of *Compa's* and *Categ's* methods. In addition, there are methods that use actual closeness instead of or as well as closeness rank.

   One qualitative difference between the way *Categ's* Step 1 uses S and T is that T is only used after an initial neighborhood of S has been found. The calculation of that initial neighborhood is in no way guided by T. But one could have such guidance while still keeping a stronger influence from S.

2. The mediator-application parts of *Compa* and *Categ*, namely their Step 2s, are similarly on a rich spectrum of possibilities, except this time *Compa* is the more contribution-disparate. One simple continuum is produced by having numerical weights on the contributions of T and S.

3. *Compa* is more careful in finding the mediator, because it attempts to put more constraint on the distance from T than *Categ's* Step 1 does. *Compa's* carefulness in this regard is increased by its incremental stepping up through possible values of m in Step 1.

4. As regards mediacy preservation, both algorithms include the mediacy strongly in the generated meaning by using the k mediator vectors in the averaging. The relative degree of usage goes up with increasing k. That that usage can be a matter of degree, not a black and white issue, is itself usefully demonstrated.

We should note, however, that none of our observations diminish Utsumi's overall achievement in capturing aspects of human behavior on metaphor in his model.

7 Conclusion and Final Remarks

We conclude that the comparison/categorization battle, at least as currently framed, is misdirected. There is an alternative to the battle—a more pressing, more immediately relevant, and better arranged one: a battle to be waged within a theoretical landscape defined by various mental-processing dimensions to which the comparison/categorization distinction is only weakly relevant. Hence, those notions do not capture the truly important distinctions.

The proposed dimensions do not constitute understanding processes, but merely characteristics of how processes can broadly proceed. Theories of understanding may naturally wish to propose particular processes. There is therefore no claim that theories should not be framed in terms of comparison and/or categorization, but rather that, if they are, they need to be more specific about the detailed nature of comparison and categorization proposed, and more aware of different variants of comparison and categorization. Furthermore, the theories could fruitfully attend explicitly to where they lie on our dimensions. The comparison/categorization fight, if there is one to be had, is really between particular theories involving them, not the general notions, and a battle within one camp may be more important than a war between them.

We have not shown that a C/C borderline is devoid of value, since we have not shown that all experiments deployed in the Debate can be explained by means of the dimensions. Rather, we have provided some suggestive evidence that the experimental findings can be so explained, and have motivated a search for further or alternative dimensions if the three proposed ones (Target/Source Contribution Disparity, Target/Source Mediacy Preservation, and Target/Source Mediator Carefulness) turn out not to be adequate in themselves.

Indeed we saw that Contribution Disparity and Mediator Carefulness are multi-faceted and may ultimately need to be split into more dimensions. Something beyond the scope of this article but also important for future research is the connection of carefulness to deliberate use of metaphor as identified by Steen (2008) and currently being actively debated (Gibbs 2011 and responses in same journal issue). Deliberateness could be seen as involving a type of carefulness on the speaker side, but could lead to extra hearer-side carefulness or a different type of hearer-side carefulness.

Finally, we put forward certain specific, conjectured tendencies concerning metaphor or simile in regard to the dimensions. The most salient ones are as follows:

- (From section 5.1): The more conventional/novel a particular source item is, the more/less (respectively) contribution disparity will arise in the understanding process. But even a simile or be-form metaphor involving a completely novel source can sometimes be very simply understood by a highly contribution-disparate process appealing to existing source properties, categories, etc. that apply directly to the target.

- (From section 5.2): Simile tends to lead to greater preservation of the target/source relationship (the mediacy) in the metaphorical meaning and in cognitive activation than be-form metaphor does.

- (From section 5.2): Higher familiarity/novelty of a T/S pairing encourages lower/higher mediacy-preservation respectively.

- (From section 5.3): Metaphor tends to encourage lower mediator-carefulness than simile does.

These are all fruitful areas for future research.

Acknowledgments

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*, 311–338. Cambridge, U.K.: Cambridge University Press.

Barnden, J.A. (2010). Metaphor and metonymy: Making their connections more slippery. *Cognitive Linguistics, 21*(1), pp.1–34.

Barnden, J.A. (forthcoming). Metaphor and simile: Categorizing and comparing categorization and comparison. Chapter based on invited talk at *Conference on Metaphor and Communication*, Cagliari, Italy, May 2011.

Barnden, J.A., Glasbey, S.R., Lee, M.G. & Wallington, A.M. (2004). Varieties and directions of interdomain influence in metaphor. *Metaphor and Symbol, 19*(1), pp.1–30.

Barnden, J.A., Helmreich, S., Iverson, E. & Stein, G.C. (1996). Artificial intelligence and metaphors of mind: within-vehicle reasoning and its benefits. *Metaphor and Symbolic Activity, 11*(2), pp.101–123.

Bowdle, B.F. & Gentner, D. (2005). The career of metaphor. *Psychological Review, 112*(1), pp.193–216.

Campbell, J.D. & Katz, A.N. (2006). On reversing the topics and vehicles of metaphor. *Metaphor and Symbol, 21*(1), pp.1–22.

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*(2): pp.283–312.

Chiappe, D.L. & Kennedy, J.M. (2001). Literal bases for metaphor and simile. *Metaphor and Symbol, 16*(3&4), pp.249–276.

Chiappe, D., Kennedy, J.M. & Smykowski, T. (2003).Reversibility, aptness, and the conventionality of metaphors and similes. *Metaphor and Symbol, 18*(2), pp.85–105.

Davidson, D. (1978). What metaphors mean. *Critical Enquiry, 5*, pp.31–47.

Falkenhainer, B., Forbus, K.D. & Gentner, D. (1989). The Structure-Mapping Engine: algorithm and examples. *Artificial Intelligence, 41*(1), 1–63.

Fogelin, R.J. (1994). Metaphors, similes and similarity. In J. Hintikka (Ed.), *Aspects of Metaphor,* pp.23–39. Norwell, MA: Kluwer Academic Publishers.

Fogelin, R.J. (2011). *Figuratively speaking.* Revised ed. New York, Oxford: Oxford University Press.

Forbus, K.D., Ferguson, R.W. & Gentner, D. (1994). Incremental structure-mapping. In *Procs. Sixteenth Annual Conference of the Cognitive Science Society*, pp.313–318. Hillsdale, N.J.: Lawrence Erlbaum.

Forbus, K.D. & Oblinger, D. (1990). Making SME greedy and pragmatic. In *Procs. Twelfth Annual Conference of the Cognitive Science Society* (pp.61–68).

Gentner, D. (1983). Structure-mapping: a theoretical framework for analogy. *Cognitive Science, 7* (2), 95–119.

Gentner, D. & Bowdle, B.F. (2001). Convention, form and figurative language processing. *Metaphor and Symbol, 16*(3&4), pp.223–247.

Gentner, D. & Bowdle, B. (2008). Metaphor as structure-mapping. In R.W. Gibbs, Jr. (Ed.), *The Cambridge Handbook of Metaphor and Thought*, pp.109–128. Cambridge, U.K.: Cambridge University Press.

Gibb, H., & Wales, R. (1990). Metaphor or simile: Psychological determinants of the differential use of each sentence form. *Metaphor and Symbolic Activity, 5*, 199–213.

Gibbs, R.W., Jr. (2011). Are 'deliberate' metaphors really deliberate? A question of human consciousness and action. *Metaphor and the Social World, 1*(1), pp.26–52.

Glucksberg, S. (2001). *Understanding figurative language: From metaphors to idioms.* New York: Oxford University Press.

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, U.K.: Cambridge University Press.

Glucksberg, S. (2011). Understanding metaphors: The paradox of unlike things compared. In K. Ahmad (Ed.), *Affective Computing and Sentiment Analysis: Emotion, Metaphor and Terminology,* pp.1–12.

Glucksberg, S. & Haught, C. (2006a). On the relation between metaphor and simile: When comparison fails. *Mind and Language, 21*(3), pp.360–378.

Glucksberg, S. & Haught, C. (2006b). Can Florida become like the Next Florida? When metaphoric comparisons fail. *Psychological Science, 17*(11), pp.935–938.

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

Goatly, A. (1997). *The language of metaphors.* London and New York: Routledge.

Gokcesu, B.S. (2009). Comparison, categorization, and metaphor comprehension. In *Procs. Conference of the Cognitive Science Society,* pp.567–572.

Gregory, M.E. & Mergler, N.L. (1990). Metaphor comprehension: In search of literal truth, possible truth, and metaphoricity. *Metaphor and Symbolic Activity, 5*, pp.151–173.

Hampton, J.A. (2007). Typicality, graded membership, and vagueness. *Cognitive Science, 31*, pp.355–383.

Holyoak, K.J. & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13*(3), pp.295–355.

Holyoak, K.J., Novick, L.R. & Melz, E.R. (1994). Component processes in analogical transfer: mapping, pattern completion, and adaptation. In K.J. Holyoak & J.A. Barnden (Eds), *Advances in Connectionist and Neural Computation Theory, Vol. 2,* pp.113–180. Norwood, N.J.: Ablex Publishing Corp.

Jones, L.L. & Estes, Z. (2006). Roosters, robins, and alarm clocks: Aptness and conventionality in metaphor comprehension. *J. Memory and Language, 55*, pp.18–32.

Kennedy, J.M. & Chiappe, D.L. (1999). What makes a metaphor stronger than a simile? *Metaphor and Symbol, 14*(1), pp.63–69.

Kintsch, W. (2000). Metaphor comprehension: A computational theory. *Psychonomic Bulletin and Review, 7*, pp.257–266.

Kintsch, W. (2001). Predication. *Cognitive Science, 25*, pp.173–202.

Kintsch, W. (2008). How the mind computes the meaning of metaphor: A simulation based on LSA. In R.W. Gibbs, Jr. (Ed.), *The Cambridge Handbook of Metaphor and Thought*, pp.129–142. Cambridge, U.K.: Cambridge University Press.

Kintsch, W. & Mangalath, P. (2011). The construction of meaning. *Topics in Cognitive Science, 3*, pp.346–370.

Kolodner, J.L. (1993). *Case-based reasoning.* San Mateo, CA: Morgan Kaufmann.

Krennmayr, T. (2011). *Metaphor in newspapers.* Utrecht: LOT – Netherlands Graduate School of Linguistics.

Lakoff, G. & Johnson, M. (2003). *Metaphors we live by.* 2nd Ed. Chicago: University of Chicago Press.

Lee, M.G. & Barnden, J.A. (2001). Reasoning about mixed metaphors with an implemented AI system.

Metaphor and Symbol, 16 (1&2), pp.29–42.

Levinson, S.C. (1983). *Pragmatics.* Cambridge, U.K.: Cambridge University Press.

Miller, G.A. (1979). Images and models, similes and metaphors. In A. Ortony (Ed.), *Metaphor and Thought*, pp.357–400. Cambridge, U.K.: Cambridge University Press.

Nowottny, W. (1965). *The language poets use.* University of London: The Athlone Press.

O'Donoghue, J. (2009). Is a metaphor (like) a simile? Differences in meaning, effects and processing. *UCL Working Papers in Linguistics* 21, pp.125–149.

Ortony, A. (1979). The role of similarity in similes and metaphors. In A. Ortony (Ed.), *Metaphor and Thought,* pp.186–201. Cambridge, U.K.: 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), pp.1–19.

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), pp.377–388.

Roncero, C., Kennedy, J.M. and Smyth, R. (2006), Similes on the Internet have explanations. *Psychonomic Bulletin & Review, 13*(1), pp.74–77.

Rosch, E. (1975). Cognitive representations of semantic categories. *J. Experimental Psychology: General, 104*, pp.192–232.

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, U.K.: Cambridge University Press.

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

Stern, J. (2000). *Metaphor in context.* Cambridge, MA and London, UK: Bradford Books, MIT Press.

Thibodeau, P.H. & Durgin, F.H. (2011). Metaphor aptness and conventionality: A processing fluency account. *Metaphor and Symbol, 26*(3), pp.206–226.

Thomas, M.S.C. & Mareschal, D. (2001). Metaphor as categorization: A connectionist implementation. *Metaphor and Symbol, 16*(1&2), pp.5–27.

Tirrell, L. (1991). Reductive and nonreductive simile theories of metaphor. *J. Philosophy, 88*(7), pp.337–358.

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

Utsumi, A. (2011). Computational exploration of metaphor comprehension processes using a semantic space model. *Cognitive Science, 35*, pp.251–296.

Utsumi, A. & Sakamoto, M. (2007). Computational evidence for two-stage categorization as a process of adjective metaphor comprehension. In *Procs. Second European Cognitive Science Conference* (EuroCogSci2007), pp.77–82.

Verbeemen, T., Vanpaemel, W., Pattyn, S., Storms, G. & Verguts, T. (2007). Beyond exemplars and prototypes as memory representations of natural concepts: A clustering approach. *J. Memory and Language, 56*, pp.537–554.

Vervaeke, J. & Kennedy, J.M. (1996) Metaphors in language and thought: Falsification and multiple meanings. *Metaphor and Symbolic Activity, 11*(4), pp.273–284.

Voorspoels, W., Storms, G. & Vanpaemel, W. (2011). Contrast in natural language concepts: An exemplar-based approach. In L. Carlson, C. Hoelscher & T.F. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 531–536). Austin, TX: Cognitive Science Society.

Voorspoels, W., Vanpaemel, W. & Storms, G. (2008). Exemplars and prototypes in natural language concepts: A typicality-based evaluation. *Psychonomic Bulletin and Review, 15*(3). pp.630-637.

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

Xu, X. (2010). The relationship between the content and form of metaphorical statements. *J. Psycholinguist Research, 39*, pp.165–178.
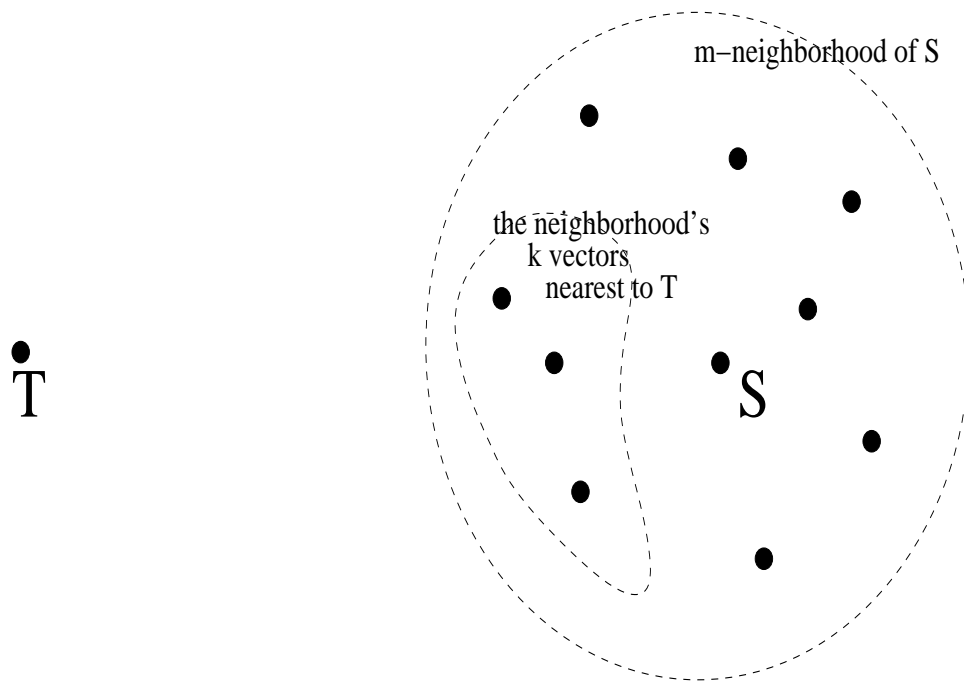
Figure 1: Impressionistic view of *Categ* algorithm's effect. Each black dot illustrates the end point of a word vector. (The vectors, normalized to have the same spatial length, can be thought of as leading from the centre of a high-dimensional sphere to its surface. The figure then shows a region of the surface.) The T and S vectors are indicated. Closeness of dots in the figure symbolizes similarity of the vectors. The m-neighborhood of S (dashed ellipse) contains the m vectors closest to the S vector. The vectors in the smaller dashed region are the k vectors in S's m-neighborhood that are closest to the T vector.
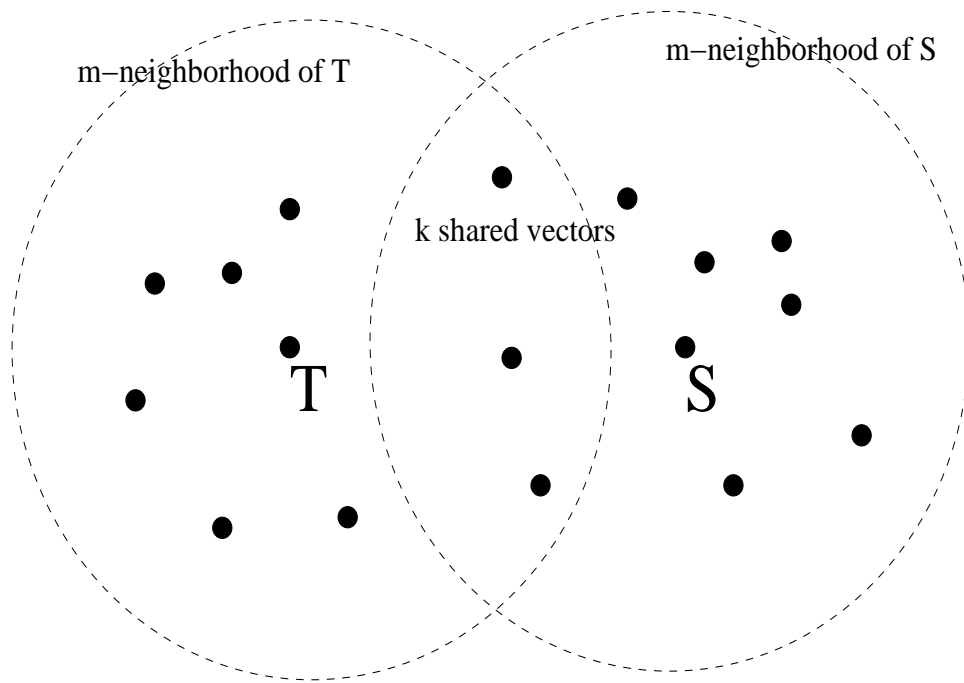
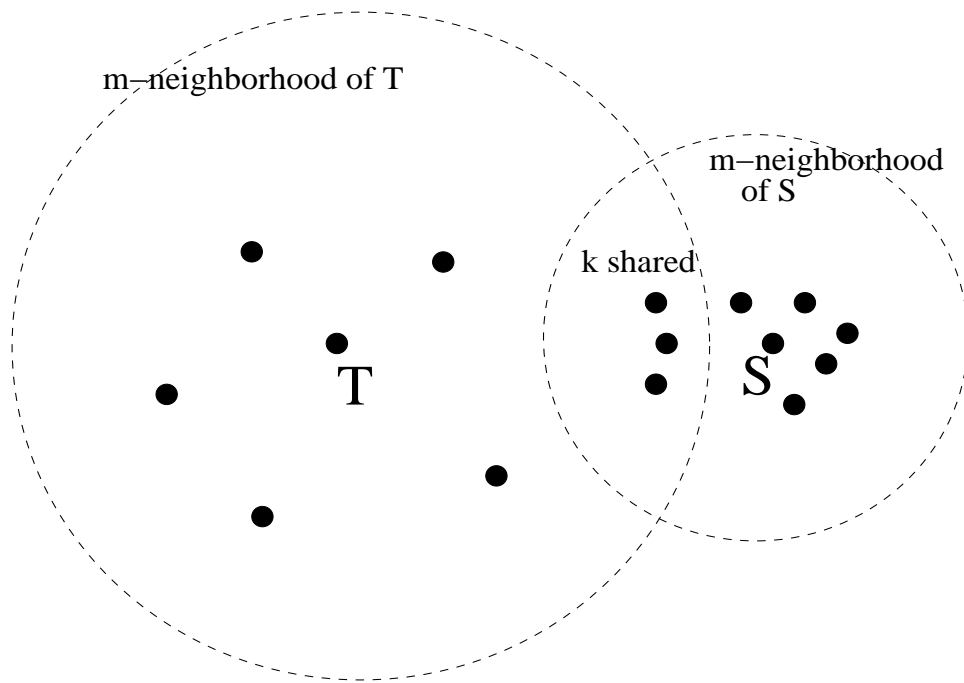Figure 2: Impressionistic view of *Compa* algorithm's effect. Diagram conventions are as in Fig. 1.

Figure 3: Variant of Fig. 2, suggesting that the m vectors closest to S may be much closer in actual distance to S than the m vectors closest to T are to T. The opposite is similarly possible, of course.