Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Pragmatics xxx (2005) xxx–xxx
journal of PRAGMATICS
www.elsevier.com/locate/pragma

Attribution of discourse goals for using concrete- and abstract-tenor metaphors and similes with or without discourse context

Richard Jackson Harris [a,*], Brian M. Friel [b], Nolan Rett Mickelson [a]

[a] *Department of Psychology, Kansas State University, Bluemont Hall, 1100 Mid-Campus Drive, Manhattan, KS 66506-5302, USA*
[b] *Delaware State University, USA*

Received 15 November 2003; received in revised form 17 April 2005; accepted 15 June 2005

Abstract

Although there has been considerable research on the comprehension of figurative language, investigation of its production has been much less common. A particularly neglected aspect has been the study of the attributions by the language receiver of the language producer's intentions in using various types of figurative language. Experiment 1 presented young-adult participants with 16 sentences containing metaphors (*The submarine was a whale*) and similes (*The submarine was like a whale*). They were asked to check as many of twelve discourse goals, reasons why they thought the author might have chosen that figure of speech, as they thought it to be relevant to that particular sentence. Materials were presented (1) orally, in written form, or both, (2) with or without a prior meaningful discourse context, and (3) with a concrete or abstract topic. The most frequently chosen goal was to Compare Similarities, the only goal on which similes (more often chosen) and metaphors differed. The pattern of nine discourse goals differed between the Abstract and Concrete sentences. The factor of discourse context affected the discourse goal attributions on Be Humorous and Compare Similarities, while the factor of modality affected only Add Emphasis. Experiments 2 and 3 examined stylistic preference and perceived aptness of the comparisons expressed in the sentences used in Experiment 1 as possible explanations for the results obtained. Findings were interpreted in terms of

* Corresponding author. Tel.: +1 785 532 0610; fax: +1 785 532 5401.
E-mail address: rjharris@ksu.edu (R.J. Harris).

0378-2166/$ – see front matter
doi:10.1016/j.pragma.2005.06.010

different discourse goals tapping into either the situation level model of representation or only the propositional textbase level.

Keywords: Figurative language; Comprehension versus production of; Metaphors; Similes; Discourse goals; Attribution of; Discourse context; Concrete versus abstract; Situation model of representation

Although there has been considerable research on the comprehension of figurative language (see Cacciari and Glucksberg, 1994; Gibbs, 1994, and Katz, 1998, for reviews), it has mostly centered on a few figures of speech, particularly metaphors and idioms. There are, however, many other types of figurative language, which have been less studied (e.g., metonymy, simile, hyperbole, oxymoron, and irony). The research focus of the psycholinguistics literature on figurative language has been on comprehension, e.g., whether metaphors are comprehended after an initial rejection of a literal interpretation (Glucksberg and Keysar, 1990; Ortony, 1979; Searle, 1979). Greatly neglected has been the production of figurative language, particularly the question of why a writer or speaker would choose one form of figurative language over another or over a literal variant.

One important aspect of the pragmatics of language production is its intentionality. Gibbs (1999) argued that intentionality is a large part of what drives the production of language, and attributions about those intentions are critical in comprehension. For example, a speaker or writer chooses certain words when holding intentions to be polite (''Could you reach the salt?''), sarcastic (''You're certainly cheery today.''), or disapproving (''I can't believe she would come here dressed like that!''). In each case, the hearer's correct attribution of the speaker's intentions is crucial for successful communication to occur. This discourse goals approach to figurative language argues that, rather than primarily assessing literal truth, the listener tries to assess why the speaker used the words he or she did, that is, to recover the discourse goals (Kreuz, 2000).

In terms of figurative language specifically, a speaker or writer intends to be non-literal, or in the case of metaphor, to *metaphorize* (MacCormac, 1985). The communication fails if the hearer or reader does not correctly infer that speech act of intending to be non-literal. However, there are many reasons that someone might choose to speak non-literally in a particular situation. Such reasons are called discourse goals.

The present study used the task of identifying discourse goals to examine the language receiver's attributions of the language producer's reasons for using particular figurative language. It is difficult to assess speakers' goals in choosing certain words in an on-line processing situation, and, even if possible, this would be highly obtrusive and turn the task into an intensely unnatural one. We chose rather to focus more on the comprehender's assessment of the speaker's discourse goals, given that scholars such as Gibbs, MacCormac, and Kreuz argue that such attributions are an integral part of the comprehension of figurative language.

One of the few empirical studies examining the comprehender's attributions about a speaker or writer's discourse goals in using figurative language was Roberts and Kreuz (1994). They instructed each participant about one of eight types of figurative language: metaphor, simile, hyperbole, idiom, irony, indirect request, understatement, and rhetorical

question. They then asked participants to generate three additional examples of that figure and list as many reasons as they could for someone to use that figure of speech. These reasons were content analyzed by two judges (interrater agreement = 75%) into a taxonomy of 19 discourse goals. Each of the eight figures of speech had its own characteristic discourse goal profile. For example, hyperbole was very often seen as used to "emphasize" and "clarify," while indirect requests were seen to "guide another's actions," "be polite," and "protect the self."

The present studies take a closer look at the discourse goals that motivated the use of two specific figures of speech: metaphors and similes. Although metaphors and similes were rated as the two most similar by participants in Roberts and Kreuz' study (1994), there were some important differences in how they were used. While both were perceived as being used to "clarify" (82% vs. 94%) "provoke thought" (35% vs. 39%), and "compare similarities" (35% vs. 33%), metaphors were perceived to be used to "add interest" far more often than were similes (71% vs. 22%), while similes were more often used to "be humorous" (33% vs. 0%). Metaphors were also seen as used more often than similes to "be conventional" (24% vs. 6%).

Alone among the different figures of speech, metaphors and similes are structurally and lexically identical except for the presence of the explicit comparison marker **like** in the simile (Her date was like an octopus, Her eyes were like diamonds); it is otherwise identical to the metaphor (e.g., Her date was an octopus, Her eyes were diamonds). Metaphors and similes also appear to be very similar in meaning. Given their high degree of syntactic and semantic similarity, the rather large differences in the discourse goal findings of Roberts and Kreuz (1994) are somewhat striking and bear further inquiry. The participants in the latter's study responded to a description of a single type of figure of speech and to their own generated examples, in what was essentially a metalinguistic task. In contrast, the present study asked participants to respond to particular instances of *both* metaphors and similes. As such, this task is closer to (and somehow part of) natural language comprehension. In that sense, it is less metalinguistic and more like natural language processing than is Roberts and Kreuz' task.

Many, if not most, models of metaphor used in psychology, linguistics, and philosophy consider both metaphors and smiles as statements of comparison of the *tenor* (or *topic*), what the figure of speech is talking about, and the *vehicle*, what the figure is talking about the tenor in terms of. What the tenor and vehicle have in common is the *ground*, which must be computed to successfully comprehend the metaphor. Many classic models in philosophy (and later in psychology) posited some sort of substitution or comparison model (e.g., Black, 1962; Richards, 1936). Some of these specific models posit the transfer of a semantic feature from the vehicle to the tenor or an imbalance in the salience of properties of the tenor and vehicle (e.g., Ortony, 1979).

Although not all psychological models of figurative language processing make comparative predictions about metaphors and similes, most that do (e.g., Miller, 1979) consider similes a more basic, direct, and explicit comparison than metaphors, because the comparison in a simile is unambiguously signaled by the presence of the word **like**. Metaphors are thus elliptical similes, and they must be recoded back to similes in order to be comprehended. This sort of model is almost totally syntactically based with little reference to discourse context.

Not all models view metaphor as an abbreviated simile or implicit comparison, however. MacCormac (1985) argued that metaphors are more likely to be chosen to evoke consideration of both similarities and differences between the tenor and the vehicle, while similes are used primarily to evoke similarities. Ricoeur (1977) claimed that novel metaphors evoke surprise with their unexpected juxtaposition of the tenor and the vehicle, thus leading to heightened perceptual involvement and deeper processing, which results in more complex ground relations in a metaphor than in its corresponding simile. Glucksberg's (1991, 1998, 2001, 2003; Glucksberg and Keysar, 1990) attributive categorization (class-inclusion) model posits the metaphor as the more direct, psychologically simpler statement, which directly asserts a class inclusion. For example, saying his job is a jail asserts that the job belongs to a category of unpleasant entities denoted by jail. However, a simile actually requires additional processing, because the comparison is qualified by **like** and thus only indirectly asserts a class-inclusion statement. Metaphors need not be implicitly transformed into similes to be understood, but rather are comprehended directly as class-inclusion statements. Support for the attributive categorization model has come from studies using a variety of tasks, including reaction time (Glucksberg et al., 1997; Johnson, 1996), paraphrasing (Glucksberg et al., 1997; McGlone, 1996), and cued recall (Harris and Mosier, 1999; Harris et al., 1999; McGlone, 1996). Some have suggested limits on the class inclusion process; for example, Gentner and Bowdle (2001) argue that only conventional metaphors are understood by means of class inclusion, while novel metaphors are understood by a comparison process.

One important attribute of metaphors and similes is their degree of concreteness. Both the tenor and the vehicle can vary from highly abstract to highly concrete. Using a sentence-completion task where participants had to fill in either "is" or "is like" between a given topic and vehicle, Gibb and Wales (1990a) showed that metaphors were preferred for abstract vehicles, while similes were preferred for concrete vehicles. Concreteness of the tenor mattered far less in the preference for a metaphor or simile. Other research has shown that a greater degree of concreteness is associated with comprehensibility in metaphors (Zelman et al., 1985) and that concrete tenors tend to evoke more intense sensory images (Gibb and Wales, 1990b).

Most models of metaphor and simile comprehension seem to deal specifically with understanding the relationship between the tenor and vehicle, and do not explicitly address the role of inferring discourse goals of the speaker or writer. As such, these models do not clearly motivate predictions for what attributions listeners and readers make about the reasons a speaker or writer might use a metaphor versus a simile. Therefore, the present study's investigation of attributed discourse goals (why a metaphor or simile was used) must be characterized as exploratory.

1. The present research

The present research presented participants with a list of 12 discourse goals and asked them to identify those they thought to be relevant to the production of a specific sentence. As such, it extended the work of Roberts and Kreuz (1994) in four ways. First, a more ecologically valid task was used, assessing the comprehender's attribution of discourse

goals for *particular* sentences, rather than merely asking participants to list what purposes those figures of speech would be used for. This process of attributing the discourse goals is a major part of the task of figurative language comprehension. Secondly, the modality of the sentences was systematically varied, in order to compare attributions in both written and spoken figurative language use. It is possible that some goals may appear more salient in the spoken than in the written modality. Only spoken language has the cues of prosody, which often are very important when establishing a speaker's intention. For example, a particular tone of voice may signal sarcasm, or a rising intonation may turn a syntactic statement into a pragmatic question. The written sentence is devoid of such cues; thus, observed differences in judgments about oral versus written sentences may be due to prosodic factors.

Third, the presence or absence of a discourse context was compared. Context is known to affect the meaning constructed from figurative language, but its effect on the inferred intentionality of figurative language is not known. We would expect sentences in meaningful discourse context to be more likely to draw on the situation model level of representation. Situation models involve the retrieval of prior knowledge as part of the comprehension process. On the other hand, sentences in lists would necessarily be more limited to the text-based level of representation, which involves parsing the string of words and assigning meaning (Kintsch, 1998). We predict that the discourse goals of Add Emphasis, Add Interest, and Clarify will be selected more for similes and metaphors embedded in a discourse context than for those presented in isolation. The rationale for this prediction is that the similes and metaphors at the end of the paragraphs (i.e., in context) can serve to add emphasis or interest to, as well as clarify, the preceding discourse. In the case of figures of speech presented in isolation, there is no other discourse to add emphasis to, add interest to, or clarify.

Finally, sentences with either concrete or abstract tenors were compared, in order to see if discourse goals were differentially attributed as a function of concreteness, a reasonable question given that several other aspects of figurative language have been related to concreteness. It seems plausible that people might infer different discourse goals motivating the use of such concrete vehicles for describing either abstract or concrete tenors, e.g., *a canary is (like) a prisoner* versus *abusive control is (like) a cracking whip*. Since Gibb and Wales (1990a) found no difference in preference for concrete versus abstract tenors, no a priori preference should be assumed to be a factor in the present study. Also, abstract vehicle sentences are somewhat odd, especially in isolation, and for this reason they were not used.

As to specific goals, no directional predictions other than those mentioned above were made; the four major research questions addressed were:

(1) What different discourse goals do people attribute to an interlocutor (in speech or writing) for choosing a metaphor over a simile?
(2) What effect does the particular modality of production (speech or writing) have on attributions about an interlocutor's goals in using a particular metaphor or simile?
(3) How does the presence or absence of a discourse context affect the attribution of discourse goals?
(4) How are concrete- versus abstract-tenor sentences perceived differently in terms of discourse goals?

2. Experiment 1

2.1. Method

2.1.1. Participants

The participants were 242 native English-speaking students sampled from a large Midwestern U.S. university with open admissions. They received course credit for participation and were tested in small groups.

2.1.2. Materials

The experiment used the materials from Harris and Mosier (1999), composed of two lists of 16 sentences each. Half of the sentences in each list were metaphors and half were similes, with the two differing only in the addition of the word *like* in the simile sentences. The two lists of sentences were identical except that any sentence, which contained a simile in one list (*A small mind is like a locked room*) contained a metaphor in the other list (*A small mind is a locked room*). All sentences were in the past tense and had concrete vehicles/predicates. Half of them contained concrete tenors/subjects, e.g., *The desert was (like) an oven,* while half contained abstract tenors/subjects, e.g., *The small mind was (like) a locked room.*

2.1.3. Design

The experiment was a $3 \times 2 \times 2 \times 2$ mixed design. One independent variable was Figure type (metaphor or simile) and was within-subjects. Complete counterbalancing ensured that each list contained half metaphors and half similes, and any given sentence appeared in its metaphor form in one list and its simile form in the other. The second within-subjects variable was Concreteness. Half of the 16 sentences each participant encountered (4 metaphors, 4 similes) contained Concrete topics and half Abstract ones.

The other two variables were between-subjects. The sentences were presented either in a list format (List) or embedded at the end of a brief story (Context). The stories, including the target sentence, varied from 32 to 61 words long (see sample materials in Table 1). Finally, for one-third of the participants in each context condition, the sentences were presented orally (Oral), for one-third in written form (Written), and for one-third in both oral and written form simultaneously (Both).

2.1.4. Procedure

Volunteers assigned themselves randomly to groups by signing up for specific small-group sessions of 3–15 people. Upon arrival, the participants in the Oral condition were told:

> ‘‘When people speak or write, they choose words carefully, and for particular reasons. This is a task assessing your perception of the reasons a person would say something in a particular way. You will listen to a list of sentences and, for each sentence, will be asked to mark all possible reasons for the speaker choosing these words. For example, if the sentence stated, ‘‘Brains are like complex computers,’’ you might think they chose these words to be humorous, to emphasize a point, to provoke thought about a subject, and/or to compare similarities. In another example,

Table 1
Sample stimulus materials: Experiment 1

Sentence list concrete condition:
The playful monkeys were (like) clowns.
The hungry mosquitoes were (like) vampires.

Sentence list abstract condition:
The makeshift lie was (like) a returning boomerang.
The harsh criticism was (like) a stinging bullet.

Context concrete condition:
Marine biologist Anne Marie looked out of her submersible as she glided through the coastal waters looking for traces of damage to the coral reef. Suddenly a huge dark object breezed by out the starboard side. First thinking it was some huge creature, Anne Marie then realized it was one of those submarines from the nearby naval base. The submarine was (like) a whale.

Context abstract condition:
Dylan often tended to fly off the handle at others without much provocation, and people found him very quick to anger. One time he started yelling at his friend Sam over a minor disagreement and before long they were slugging it out by punching each other. His short temper was (like) a smoking volcano.

Note: Metaphor and simile sentences were identical except for the presence of "like" in the simile condition. Every participants heard/read either the metaphor or the simile member of each pair but never both. Equal number of participants heard/read any given sentence in each version.

> "Love is a flower," you might think that writer chose these words to be eloquent, to show positive emotion, or to provoke thought about a subject. Each sentence's use can be placed into as few or as many categories as you feel appropriate. There are no right or wrong answers."

Minor changes in wording, as appropriate, occurred in the instructions for the Written and Both conditions.

The answer sheets consisted of the numbers 1–16 and a 16 × 12 grid of small squares. Across the top were listed 12 of Roberts and Kreuz' (1994) 19 discourse goals (the other seven of their goals were never or very seldom selected for metaphors or similes in their study and thus were not included in the present study). In the Written and Both, but not the Oral condition, the sentences were written next to each number 1–16 by the grid. Participants placed an × in however many of the 12 squares for a given sentence they believed to represent relevant goals for the speaker or writer using that sentence.

In the Oral and Both conditions, the experimenter stopped the tape recorder after each sentence and waited until all participants had responded. Then he moved on to the next sentence. In the Written condition, each participant worked at his or her own speed. In the Context condition, instructions made clear that participants were to select the discourse goals of the final (critical) sentence only.

2.1.5. Results and discussion

Results were scored for the presence each discourse goal marked, and are presented in Table 2. For each discourse goal, the proportions of participants checking it were analyzed in two 3 × 2 × 2 × 2 analyses of variance, with between-subjects factors of Modality

Table 2
Experiment 1: percent selecting discourse goals by condition

Discourse goal	Figure		Context		Modality			Tenor (Conc)	Tenor (Abst)	RK94[a]	
	Met	Sim	No	Yes	Oral	Written	Both			Met	Sim
Compare similarities	57*	64*	65*	56*	60	57	64	69*	52*	35	33
Add emphasis	45	44	41	48	39*	43*	51*	41*	48*	24	11
Show negative emotion	34	34	33	36	34	30	39	14*	55*	17	31
Provoke thought	32	32	32	32	33	29	35	22*	42*	35	39
Add interest	26	26	26	27	24	27	28	30*	23*	71	22
Get attention	25	26	26	25	22	23	31	24	27	12	11
Clarify	20	19	17	23	21	17	21	20	19	82	94
Be humorous	16	15	19*	11*.	15	17	14	25*	6*	0	33
Be conventional	16	14	14	16	17	13	16	17*	13*	24	6
Be eloquent	14	12	13	13	15	13	11	11	15	35	22
Show positive emotion	7	8	6	9	7	7	9	13*	2*	6	6
Contrast differences	6	6	6	7	7	7	5	5*	7*	6	0

Note: Met: metaphor; Sim: simile; Conc: sentences with concrete topics; and Abst: sentences with abstract tenors.
[a] Percent of time goal selected by participants in Roberts and Kreuz (1994) (=RK).
* Mean differ for F^1 and F^2 at $p < .05$.

(Oral, Written, or Both) and Context (Context or List), and the within-subjects factors of Figure (Metaphor or Simile) and Concreteness (Concrete or Abstract tenor). Separate analyses were performed for participants (F^1) and items (F^2). Only results significant at $p < .01$ in both participant and item analyses are discussed.

Overall, of the 12 discourse goals, only Compare Similarities, the most frequently chosen goal overall, differed for metaphors and similes (see Table 2), $F^1(1,235) = 30.78$, $\eta^2 = .01$, and $F^2(1,42) = 16.98$, $\eta^2 = .05$. This goal was more often chosen for similes (64%) than for metaphors (57%), probably because the word ''like'' in the simile signals a highly explicit comparison, as against a metaphor. On no other goal did metaphors and similes differ.

A main effect for the Context variable was obtained for two discourse goals. The sentences in Lists were more often than sentences in context rated as intended to Be Humorous, $F^1(1,235) = 22.32$, $\eta^2 = .03$, and $F^2(1,42) = 14.08$, $\eta^2 = .04$, and to Compare Similarities, $F^1(1,235) = 7.35$, $\eta^2 = .02$, and $F^2(1,42) = 27.73$, $\eta^2 = .07$. The sentence by itself was more likely to stand out as humorous or comparing similarities because that was the only material present.

For the Modality variable, two effects were found. The Add Emphasis goal was selected more often in the condition where the sentence was both heard and read at the same time than it was in the Oral or Written alone conditions, $F^1(2,235) = 5.15$, $\eta^2 = .03$ and $F^2(2,42) = 10.38$, $\eta^2 = .09$. Closer examination showed that this Modality difference was confined to the Sentence List condition, and that the interaction was significant, $F^1(2,235) = 6.42$, $\eta^2 = .03$ and $F^2(2,42) = 11.06$, $\eta^2 = .09$ (see Fig. 1). It may be the case that the meaningful discourse context was rich enough so that the extra modality did not ''Add Emphasis'' as much as it did in the List condition. Perhaps the presence of two modalities in the Both condition offers more opportunity to attract attention, thus suggesting more possible ways to ''Add Emphasis.''

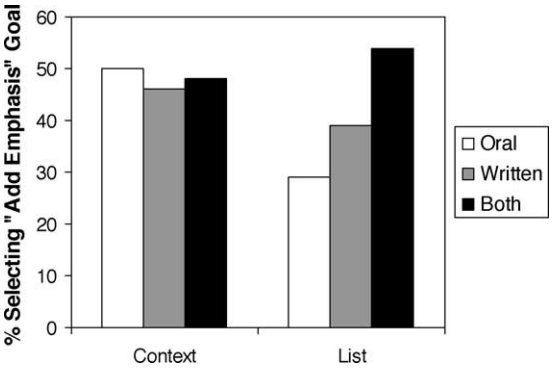

Fig. 1. Percent selecting the discourse goal add emphasis as a function of modality and discourse context.

For the Concreteness variable, several significant main effects occurred. Concrete-tenor sentences were more often than Abstract-tenor sentences seen as being used to Be Conventional, $F^1(1,235) = 15.25$, $\eta^2 = .01$, $F^2(1,42) = 18.79$, $\eta^2 = .06$; Be Humorous, $F^1(1,235) = 276.45$, $\eta^2 = .19$, $F^2(1,42) = 54.51$, $\eta^2 = .26$; Compare Similarities, $F^1(1,235) = 131.38$, $\eta^2 = .06$, $F^2(1,42) = 88.26$, $\eta^2 = .27$; Add Interest, $F^1(1,235) = 18.94$, $\eta^2 = .01$, $F^2(1,42) = 21.97$, $\eta^2 = .08$; and Show Positive Emotion, $F^1(1,235) = 114.97$, $\eta^2 = .12$, $F^2(1,42) = 59.91$, $\eta^2 = .20$. On the other hand, Abstract-tenor sentences were more often used to Contrast Differences, $F^1(1,235) = 10.08$, $\eta^2 = .01$, $F^2(1,42) = 10.97$, $\eta^2 = .05$; Add Emphasis, $F^1(1,235) = 30.20$, $\eta^2 = .01$, $F^2(1,42) = 13.98$, $\eta^2 = .04$; Provoke Thought, $F^1(1,235) = 160.78$, $\eta^2 = .12$, $F^2(1,42) = 68.41$, $\eta^2 = .29$; and Show Negative Emotion, $F^1(1,235) = 717.69$, $\eta^2 = .39$, $F^2(1,42) = 300.87$, $\eta^2 = .50$. There was also one significant interaction of Concreteness with Context (see Fig. 2). In terms of the frequency of the attribution of the discourse goal Compare Similarities, Concreteness made a larger difference for items in discourse context than for those in lists, $F^1(1,235) = 13.43$, $\eta^2 = .01$, $F^2(1,42) = 9.00$, $\eta^2 = .03$. Specifically, Concrete-tenor sentences were selected as having the

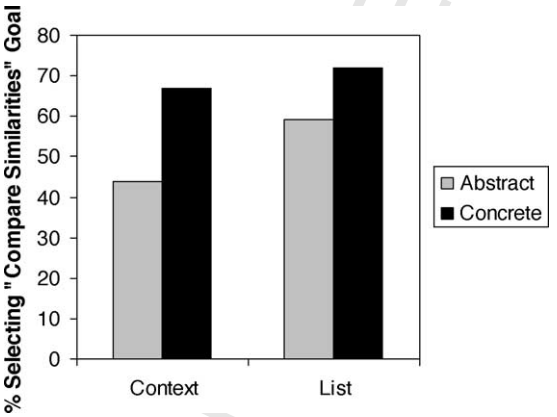

Fig. 2. Percent selecting the discourse goal compare similarities as a function of tenor concreteness and discourse context.

discourse goal of Comparing Similarities with comparable frequency in both Context conditions. However, attribution of Abstract-tenor sentences as reflecting the goal of Comparing Similarities occurred more frequently when items were in lists than when they were presented in discourse context. The Abstract sentences in the List condition were the least able to draw on the situation model representation; thus, the participant was left with only the text-based representation, in the overt form of a comparison, as a possible basis for attributing discourse goals.

The findings from the present studies were not consistent with the predominant discourse goals found for metaphors and similes by Roberts and Kreuz (1994); their data from these two speech figures appear in the last two columns in Table 2. In this connection, it is worth noticing that although both their studies and ours produced percentages of participants reporting each goal, the procedures for arriving at these seemingly comparable figures were very different. Using their metalinguistic task and a between-subjects design, Roberts and Kreuz taught participants about *one* particular figure of speech, after which each participant generated three examples of that figure and listed the reasons why someone might use that figure of speech. These responses were then content-analyzed into the discourse goal categories appearing in Table 2. The present study had participants infer discourse goals of specific sentences.

Differences between Roberts and Kreuz' and the present study particularly stand out for three discourse goals. (1) Although the Compare Similarities goal was by far the most frequently selected goal for both metaphors and similes in the present study, it was identified by only about a third of Roberts and Kreuz' participants. (2) On the other hand, the latter's most frequently chosen goal for both figures of speech was Clarify, a relatively minor goal in the present studies. (3) The discourse goal to show the largest differences between metaphors and similes in the Roberts and Kreuz study was Add Interest, identified for metaphors by 71% of the participants, as against by only 22% for similes. In the present study, both metaphors and similes produced about this same lower level. Clearly, the method of studying discourse goals is hugely important in determining the results, and further research is necessary, preferably using a variety of methodologies.

3. Experiment 2: preference ratings

One potential issue in any study of production of figurative language is the users' stylistic preference for one form over another. It is possible that metaphors or similes may be stylistically preferred, at least in certain conditions. In fact, a fill-in-the-blank task by Gibb and Wales (1990a) showed a preference for metaphors when the vehicles were abstract, but for similes when the vehicles were concrete (as all were in the present Experiment 1). To further examine this issue, Experiment 2 was conducted, using a different type of task in order to assess any a priori stylistic preference for metaphors over similes.

3.1. Method

Experiment 2 asked 40 new participants from the same participant pool as used in the first experiment to compare two alternative forms of the same sentence and say which one

they preferred. Participants were given a sheet with two columns of sentences, the same 16 sentence-pairs used in Experiment 1. A given position in each column was occupied by either the metaphor or simile version of a given sentence, with the alternative version in the other column. The placement of a metaphor or simile in each column was determined at random with the constraint that a given column had eight metaphors and eight similes. There were two variations of the answer sheet, with the simile and metaphor sentences reversed from the first form to the second, in order to control for any position effects; 20 participants completed each form.

Participants were told:

> ''This experiment is studying people's preferences in the choice of wording in sentences. Read each pair of statements below and circle the one you think best expresses the comparison in question, for example, ''Priests are shepherds'' and ''Priests are like shepherds.'' There are no right or wrong answers; we are simply interested in which wording you think expresses the comparison better.''

3.2. Results and discussion

Results were scored for the number of choices of metaphors and similes as the preferred form. There were no differences between the two forms of the answer sheet, so that will not be considered further. Overall, similes were preferred over metaphors (see Table 3), with mean preferences of 72 and 28%, respectively. Our results replicate Gibb and Wales' (1990a) and are similar to the 61% versus 39% found by Chiappe and Kennedy (1999) for ratings of a different set of similar sentences on the same task. The range for mean simile preference across individual sentences varied from 45 to 95%, while for metaphors it varied from 5 to 55%. All but one of the 16 sentences were preferred by more participants in the simile form than in the metaphor form. This strong preference for similes may have been due in part to the lack of any abstract vehicles in the present stimulus set. (It was the abstract-vehicle sentences that were most likely to be preferred in the metaphor form by Gibb and Wales' (1990a) participants.)

Because the three independent variables of figure, modality, and context were completely crossed orthogonally in Experiment 1, and because a main effect of Figure appeared for only one discourse goal (Compare Similarities), this preference bias for similes is unlikely to account for many of the obtained results. It could, however, have contributed to the difference obtained for the Compare Similarities discourse goal for similes and metaphors. However, if the stylistic preference bias greatly affected the discourse goal ratings overall, one would expect evidence of that on more

Table 3
Mean preference (Experiment 2) and aptness (Experiment 3) ratings for metaphors and similes

	Percent preferring	Mean aptness[a]
Metaphors	28	4.88
Similes	72	5.08

[a] 1: very inappropriate comparison between tenor and vehicle and 7: very appropriate comparison between tenor and vehicle.

goals than Compare Similarities; however, metaphors and similes did not differ on any other goals.

Another explanation is suggested by MacCormac (1985), who argued that metaphors direct attention to both the similarities and differences between the tenor and the vehicle, while similes focus primarily on the similarities. This obtained preference bias for similes would be consistent with this difference. Since none of the discourse goals tested included drawing attention to both similarities and differences, no direct test of MacCormac's prediction was possible. However, the fact that the discourse goal Contrast Differences was chosen only 6% of the time for both metaphors and similes does not appear to be consistent with MacCormac's hypothesis. In any event, there is little chance that this preference for similes could have affected the results from the concreteness or context variables, since those were completely crossed with figurative type and there were almost no interactions. Thus it seems unlikely that stylistic preference could account for the large majority of results obtained in Experiment 1.

4. Experiment 3: aptness ratings

Experiment 3 was another rating study of the 16-sentence set used in Experiment 1. This time, the dimension rated was the appropriateness or aptness of the comparison stated by the metaphor or simile. Aptness has been defined as "the extent to which the statement captures important features of the topic" (Chiappe et al., 2003:97) or "how well ... the [figure] expresses its nonliteral meaning" (Blasko and Connine, 1993:297), considering the appropriateness of the fit between the topic and the vehicle. Chiappe and Kennedy (1999) argued that tenors and vehicles vary in the degree to which they are entities appropriate to be compared by any figure of speech, and that this dimension could be a factor in the choice of a metaphor or simile to express an idea. Following their procedure, Experiment 3 gathered ratings for this perceived aptness of the comparison.

4.1. Method

Another 45 participants from the previous participant pool rated a list of 16 sentences. The instructions, adapted as closely as possible from Chiappe and Kennedy (1999), were as follows:

> "This task is an assessment of the extent to which a comparison manages to capture salient properties of the topic. For example, "a train is like a worm" may not be very appropriate, because the comparison between trains and worms does not seem to capture many of the salient features of trains, such as their strength and power. Similarly "tigers are teddy bears" may be inappropriate, because the comparison between tigers and teddy bears fails to capture the more salient properties of tigers, such as their striped coats and ferocity. However, a statement such as "oil is like liquid gold" may seem very appropriate, because the comparison does seem to capture many of the salient features of oil, such as its value and rarity. Similarly, the statement "Minnesota in January is an icebox" may be an appropriate comparison,

because it captures some of the most salient features of Minnesota in January, such as it being very cold. You will see a list of sentences and are asked to rate each one as to its appropriateness by circling a number between 1 and 7. There are no right or wrong answers; we are merely interested in your honest opinions.''

There were two alternative lists (Lists A and B) for this task, each containing eight metaphors and eight similes; the lists were counterbalanced such that any given sentence in metaphor form on one list was in the simile form on the other. Participants received either one list or the other (List A: $n = 22$, List B: $n = 23$), so that each participant rated only one of the two forms of each sentence-pair. Beside each sentence was a 7-point scale with the end anchors ''very inappropriate'' (1) and ''very appropriate'' (7). Participants circled one of the 7 numbers between the two end anchors.

4.2. Results and discussion

Mean aptness ratings were calculated for each sentence on each list. There were no systematic differences across the two lists, so that factor was not considered further. Overall, the mean aptness ratings for metaphor ($M = 4.88$) and simile ($M = 5.08$) sentences did not significantly differ, both being well on the ''appropriate'' side of the scale (see Table 3). For 10 of the 16 sentence-pairs, the simile was seen as more apt, while the metaphor form was preferred for the other 6. Generally speaking, the two versions of any given sentence did not differ much. If one was relatively low in aptness, the other usually was too, and similarly for those sentence pairs rated as high in aptness ($r = +.65$). Thus, the higher attribution of the discourse goal Compare Similarities in Experiment 1 to Similes than to Metaphors cannot be attributed to similes being a generally more apt means of comparison.

The mean aptness ratings for the metaphor and simile version of each sentence were combined to form a mean aptness rating for that sentence pair. Pearson correlation coefficients were calculated comparing this mean aptness rating for each of the 16 sentence pairs with its preference ratings for metaphors and similes, obtained from Experiment 2. This produced a very weak correlation ($r = .11$ for the metaphors; $r = -.11$ for the similes), with these r's not significantly different from zero at an α-level of .05. Thus, the aptness of the tenor-vehicle comparisons did not predict their preference of simile or metaphor form. This failed to replicate the result of Chiappe and Kennedy (1999), who found that more apt comparisons were more preferred as metaphors ($r = .75$). The relatively small number of items (16) may have contributed to this lack of significance.

Thus, while Experiments 2 and 3 have told us that there is an a priori stylistic preference for similes over metaphors, the two do not differ in aptness, and aptness and figure-type preference are not related.

5. General discussion

The present research has extended the findings of Roberts and Kreuz (1994), using a very different paradigm. Our procedure of attributing an author's discourse goals for

individual sentences is a workable methodology that has produced interpretable results. In Experiment 1, Compare Similarities was chosen far more often than any other goal. There also was considerable stability across the different modalities, with the difference in Add Emphasis attributable to the presence of prosodic cues in the Oral condition. All of this points to such judgments being fairly stable and depending largely, though probably not totally, on the nature of the sentence itself.

Two of the goals (Be Humorous and Compare Similarities) did depend in part on the presence of a meaningful context. These differences are interpretable in terms of Kintsch's (1998) levels of representation in memory. Using Kintsch's framework, sentences in lists were more often seen to be used to Be Humorous or Compare Similarities, both of which goals are relatively more micro-level, being able to be realized entirely from the text-based level of representation of an individual sentence without recourse to a situation model. Further research using these discourse goals, as well as studies of attributions about intentionality more generally, should carefully consider the implications that some goal attributions tap into the situation model level of representation, while others can be adequately assessed entirely from the text-base level.

Whether the topic of the sentence was Concrete or Abstract affected the attribution of several discourse goals, but this factor almost never interacted with the Context, Modality, or Figure of Speech variables. Metaphors and Similes in Concrete-tenor sentences were more likely to be seen as having been selected to Be Conventional, Be Humorous, Compare Similarities, Add Interest, and Show Positive Emotion. There was also an interaction with Context for the Compare Similarities goal. With both components being concrete, these sentences may have required less effort to compute the ground, thus leaving more cognitive resources available for appreciating Humor, Interest, Convention, and Positive Emotion. Also, the similarity comparison may have been more salient because of the closer positioning of tenor and vehicle. On the other hand, Metaphors and Similes in Abstract-tenor sentences were more likely to be seen as having been selected to Contrast Differences, Add Emphasis, Provoke Thought, and Show Negative Emotion. These differences may reflect the more disparate nature of the tenors and vehicles in the Abstract-tenor condition. Thus, they may have been more likely to be perceived as having been used to Contrast Differences, Add Emphasis, and Provoke Thought, because in these cases more mental activity is required to compute the ground connecting the tenor and the vehicle. Show Negative Emotion might also indirectly reflect the greater difference of the tenor and vehicle. These findings suggest a greater difficulty in processing the Abstract as against the Concrete-tenor sentences, and are consistent with previous research showing Abstract sentences to be more difficult to remember (Harris and Mosier, 1999; Harris et al., 1999). A useful extension of this work would be to test abstract-vehicle sentences, which Gibb and Wales (1990a) found to be the one type preferred in the metaphor form, although such sentences might seem very odd in the sentence-list condition.

The only discourse goal to show a difference between metaphors and similes was the most commonly chosen goal overall, Compare Similarities, a goal more often attributed to similes. Given that similes and metaphors differ primarily in the fact that similes are more explicit comparisons and the fact that there was an a priori preference for similes, this finding is not surprising. What is more striking is that no differences were found in Experiment 1 for any other discourse goal.

The findings regarding Modality showed a difference only in the Add Emphasis goal. This makes sense, given that Emphasis is very often indicated by the prosodic cue of stress, although in this case it was only increased in the Both condition in the sentences in lists. Perhaps it is in the isolated sentences that adding emphasis seems to be more salient, especially when the participant is both reading and hearing the sentence. The fact that Modality differences were obtained at all is noteworthy in that prosodic cues were probably less salient and less important in this study than they would be in ordinary language use, since the speech was tape-recorded for standardization purposes.

Another interesting finding is the almost total lack of significant interactions. Figure, Modality, Context, and especially Concreteness sometimes mattered, but seldom did so in any interactive fashion. This suggests that future research could more confidently investigate these factors in non-factorial designs.

Further research on various aspects of the production of figurative language is desirable. The same methodology could be extended to other figures of speech (Kreuz and Roberts, 1993). There are possible variations on the present procedure. For example, participants could assign 100 "points" of discourse goal value across the various goals, in order to assess differential weightings of different goals. For example, both Compare Similarities and Provoke Thought might be checked, but the former might receive 70% and the latter 30%. The present design had no way to assess differential weightings of the various discourse goals selected. It also might be worth including a discourse goal of "Comparing Similarities and Highlighting Differences" in order to test all aspects of MacCormac's (1985) prediction.

Research using the present paradigm, as well as that of Roberts and Kreuz (1994), could also be useful in developing further details of models of intentionality along the lines of Gibbs (1999). The choice of type of figurative language may well occur, at least in part, at a level even more abstract than the situation model and text-based levels of representation. Kintsch's (1998) most macro level of representation, the pragmatic level, might include the intention to be figurative. The next more micro level (the situation model) would include somewhat more specific discourse goals such as Be Humorous or Get Attention; these would involve the situation model level. Finally, the most micro level goals, such as Compare Similarities or Clarify, would occur at the propositional text-base level. Depending on which level of representation is involved, factors such as Modality and Concreteness have different effects.

Uncited references

Gibbs and Gerrig (1989), Graesser et al. (1997) and Hess et al. (1995).

Acknowledgements

Thanks are expressed to Tuan Tran, Kristen Geri, and Bruce Braden for helpful comments on previous drafts of the manuscript. Portions of these data were presented at meetings of the Midwestern Psychology Association, Chicago, May 2001, and the American Psychological Society, Toronto, June 2001.

References

Black, Max, 1962. Models and Metaphors. Cornell University Press, Ithaca, NY.
Blasko, Dawn G., Connine, Cynthia M., 1993. Effects of familiarity and aptness on metaphor processing. Journal of Experimental Psychology: Learning, Memory, and Cognition 19, 295–308.
Cacciari, Cristina, Glucksberg, Sam, 1994. Understanding figurative language. In: Gernsbacher, M.A. (Ed.), Handbook of Psycholinguistics. Academic Press, San Diego, pp. 447–478.
Chiappe, Daniel L., Kennedy, John M., 1999. Aptness predicts preference for metaphors or similes, as well as recall bias. Psychonomic Bulletin and Review 6, 668–676.
Chiappe, Daniel L., Kennedy, John M., Smykowski, T., 2003. Reversibility, aptness, and the conventionality of metaphors and similes. Metaphor and Symbol 18, 85–105.
Gentner, Dedre, Bowdle, B.F., 2001. Convention, form, and figurative language processing. Metaphor and Symbol 16, 223–247.
Gibb, Heather, Wales, Roger, 1990a. Metaphor or simile: psychological determinants of the differential use of each sentence form. Metaphor and Symbolic Activity 5, 199–213.
Gibb, Heather, Wales, Roger, 1990b. Psychological responses to the concreteness characteristics of literal and metaphoric expressions. Language and Cognitive Processes 5, 29–52.
Gibbs Jr., Raymond W., 1994. The Poetics of Mind. Cambridge University Press, Cambridge.
Gibbs Jr., Raymond W., 1999. Intentions in the Experience of Meaning. Cambridge University Press, Cambridge.
Gibbs Jr., Raymond W., Gerrig, Richard J., 1989. How context makes metaphor comprehension seem ''special''. Metaphor and Symbolic Activity 4, 154–158.
Glucksberg, Sam, 1991. Beyond literal meanings: the psychology of allusions. Psychological Science 3, 146–152.
Glucksberg, Sam, 1998. Understanding metaphors. Current Directions in Psychological Science 7, 39–43.
Glucksberg, Sam, 2001. Understanding Figurative Language: From Metaphors to Idioms. Oxford University Press, New York.
Glucksberg, Sam, 2003. The psycholinguistics of metaphor. Trends in Cognitive Science 7, 92–96.
Glucksberg, Sam, Keysar, Boaz, 1990. Understanding metaphoric comparisons: beyond similarity. Psychological Review 97, 3–18.
Glucksberg, Sam, McGlone, Mathew S., Manfredi, Deanna, 1997. Property attribution in metaphor comprehension. Journal of Memory and Language 36, 50–67.
Graesser, Arthur C., Millis, Keith K., Zwaan, Rolf A., 1997. Discourse comprehension. Annual Review of Psychology 48, 163–189.
Harris, Richard J., Mosier, Noah J., 1999. Memory for metaphors and similes in discourse. Discourse Processes 28 (3), 257–270.
Harris, Richard J., Tebbe, Michael R., Leka, Gary E., Garcia, R. Coral, Erramouspe, Raquel, 1999. Monolingual and bilingual memory for English and Spanish metaphors and similes. Metaphor and Symbol 14, 1–16.
Hess, D.J., Foss, Donald J., Carroll, P., 1995. Effects of global and local context on lexical processing during language processing. Journal of Experimental Psychology: General 124, 62–82.
Johnson, Andrew T., 1996. Comprehension of metaphors and similes: a reaction time study. Metaphor and Symbolic Activity 11, 145–159.
Katz, Albert (Ed.), 1998. Figurative Language and Thought. Oxford, New York.
Kintsch, Walter, 1998. Comprehension: A Paradigm for Cognition. Cambridge University Press, Cambridge, UK.
Kreuz, Roger J., 2000. The production and processing of verbal irony. Metaphor and Symbol 15, 99–107.
Kreuz, Roger J., Roberts, Richard M., 1993. The empirical study of figurative language in literature. Poetics 22, 151–169.
MacCormac, Earl R., 1985. A Cognitive Theory of Metaphor. MIT Press, Cambridge MA.
McGlone, Matthew S., 1996. Conceptual metaphors and figurative language interpretation: food for thought? Journal of Memory and Language 35, 544–565.
Miller, George A., 1979. Images and models, similes and metaphors. In: Ortony, A. (Ed.), Metaphor and Thought. Cambridge University Press, Cambridge, UK, pp. 202–250.
Ortony, Andrew, 1979. Beyond literal similarity. Psychological Review 86, 161–180.
Richards, I.A., 1936. The Philosophy of Rhetoric. Oxford University Press, London.

Ricoeur, Paul, 1977. The Rule of Metaphor: Multidisciplinary Studies of the Creation of Meaning in Language. Routledge and Kegan Paul, London.

Roberts, Richard M., Kreuz, Roger J., 1994. Why do people use figurative language? Psychological Science 5, 159–163.

Searle, John, 1979. Expression and Meaning. Cambridge University Press, Cambridge.

Zelman, Juraj, Sípos, Ivan, Horecky, Ján, 1985. Dimensions of metaphors or metaphors of dimensions? Studia Psychologica 27 (1), 75–84.

Richard Jackson Harris is professor of Psychology at Kansas State University (USA), where he teaches courses in Psycholinguistics, Cognitive Psychology, and Psychology of Mass Communication. His research interests include figurative language, bilingualism, and autobiographical memory for media experiences. He received his PhD in psycholinguistics from the University of Illinois at Urbana-Champaign. He has taught as a Visiting Fulbright Professor at the Federal University of Minas Gerais in Belo Horizonte, Brazil, and the Catholic University of Uruguay in Montevideo, Uruguay, and been a Visiting Scholar at Massey University in Palmerston North, New Zealand.

Brian Michael Friel is assistant professor of Psychology at Delaware State University (USA). He received his MS from the University of Oklahoma in 2000 and his PhD from Kansas State University in 2004. His research interests are in bilingualism, lexical access in memory, and the role of syntactic and semantic context in word perception.

Nolan Rett Mickelson is a 2001 graduate of Kansas State University (USA) with a major in Psychology.