UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Interpretive Diversity as a Source of Metaphor-Simile Distinction

Permalink
https://escholarship.org/uc/item/1n5110jm

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 27(27)

ISSN
1069-7977

Authors
Lacroix, Joyca P.W.
Murre, Jaap M.J.
Postma, Eric O.

Publication Date
2005

Peer reviewed

Interpretive Diversity as a Source of Metaphor-Simile Distinction

Akira Utsumi (utsumi@se.uec.ac.jp)

Yuu Kuwabara (kuwabara@utm.se.uec.ac.jp)

Department of Systems Engineering, The University of Electro-Communications
1-5-1 Chofugaoka, Chofushi, Tokyo 182-8585, Japan

Abstract

In this paper, we argue that the metaphor form of a comparison (i.e., a topic-vehicle pair) is preferred over, and more comprehensible than, the simile form when the interpretive diversity for that comparison is high. Interpretive diversity refers to the richness of the figurative meaning of a comparison; it is high to the extent that more features or properties are related to the figurative meaning and their values of salience are more uniformly distributed. We tested this claim through three experiments. Experiment 1 examined preference for the metaphor form and found that, as the interpretive diversity of the features shared by the topic and the vehicle increased, the metaphor preference increased as well. Furthermore, interpretive diversity was found to have a greater effect on metaphor preference than topic-vehicle similarity, which Chiappe and Kennedy (2001) claimed was a source of metaphor preference. Experiment 2 assessed familiarity of topic-vehicle pairs and demonstrated that interpretive diversity was more important in determining the metaphor preference of less familiar pairs than that of familiar pairs. Experiment 3 addressed comprehension of metaphors and similes and found that metaphors were comprehended more easily than similes when the interpretive diversity of the figurative meaning is high. These findings provide empirical evidence in favor of our argument and suggest that metaphor-simile distinction is better explained by interpretive diversity.

Introduction

When we describe a concept (i.e., topic) figuratively in terms of a different concept (i.e., vehicle), we can express the comparison in two ways: in metaphor form (*X is Y*) and in simile form (*X is like Y*). Since metaphor and simile express almost the same figurative meaning, it is often said that a metaphor is an indirect paraphrase of a simile. If so, however, why do we use both expressions without bringing them together?

Recently some experimental studies have tackled the problem of how metaphor and simile differ. One series of experimental studies (Chiappe and Kennedy 1999, 2001; Chiappe, Kennedy, and Chiappe 2003) has been made on what properties affect people's preference for the metaphor form. These studies demonstrated that metaphor preference was explained by similarity between a topic and a vehicle, or aptness of the comparison. The metaphor form was preferred over the simile form when the topic and the vehicle of the comparison were highly similar or when the comparison was highly apt, while the simile form was preferred when the topic-vehicle similarity or the aptness of comparison was low. Another series of experimental studies (Bowdle and Gentner 1999; Gentner and Bowdle 2001; Zharikov and Gentner 2002) has addressed the conventionality of the figurative meaning of the

vehicle (i.e., the base term) as a source for metaphor-simile distinction. These studies demonstrated that the simile form was preferred when the vehicle of the comparison was novel, but the metaphor form became preferred when that vehicle was conventionalized by repeated figurative use.

In this paper, we propose *interpretive diversity* of a comparison as a key property which explains the metaphor-simile distinction. By the term 'interpretive diversity', we mean the richness of the figurative meaning of a comparison. The richness of the figurative meaning depends on two factors, that is, the number of features (or predicates) involved in the meaning and the salience distribution of those features (Utsumi 2005). In general, the richness of a comparison is high to the extent that more features are related to the figurative meaning and that the values of salience of those features are more uniformly distributed. For example, the comparison "*Deserts are (like) ovens*" may possibly convey one highly salient meaning of "They are burning hot" with a relatively much less salient meanings such as "They are dry" or "Their temperature is greatly changed". Such interpretation can be seen as less rich or less diverse. On the other hand, the comparison "*History is (like) footprints*" can be considered as having a highly rich, diverse interpretation because many equally salient meanings such as "It remains behind", "It is a thing of the past" and "It is a living proof" are contained in the figurative interpretation.

Interpretive diversity can be assessed by using an idea of Shannon's entropy $H(X)$ defined by the following equation.

$$H(X) = -\sum_{x \in X} p(x) \log_2 p(x) \qquad (1)$$

A higher value of entropy means that a single random variable X has more states and their probability $p(x)$ is more uniformly distributed. The entropy is maximal, $H(X) = \log_2 N$ with N the number of states in X, when all states have the same probability $1/N$, and the entropy is minimal, $H(X) = 0$, when one state has probability 1 and all others have probability 0. Therefore, if we regard the salience distribution of features for the figurative interpretation of a comparison as a probability distribution $p(x)$, the entropy becomes a reasonable measure of richness of that interpretation.

We then argue that, when the interpretive diversity for a comparison is high, i.e., when the figurative meaning is highly diverse, the metaphor form of that comparison is preferred over, and more comprehensible than, the simile form. According to our argument, for example, "*History is footprints*" may be preferred over and comprehended more easily than "*History is like footprints*" because the comparison

is interpretively diverse, while in the case of the comparison "*deserts – ovens*" the simile form "*Deserts are like ovens*" may be preferred and more comprehensible.

In the rest of this paper, we report three experiments conducted to test our claim. In the first two experiments (Experiments 1 and 2), we addressed the preference rating for the metaphor form, i.e., to what degree people judged the metaphor form to be a better way of expressing a given comparison. In Experiment 1, we examined how the metaphor preference could be explained by the interpretive diversity of the features shared by the topic and the vehicle, and compared it with people's similarity rating, which Chiappe and Kennedy (2001) claimed was an important source of metaphor preference. In Experiment 2, we focused on how the familiarity (or conventionality) of a comparison was related to the metaphor preference of that comparison. In the last experiment (Experiment 3), we examined how comprehensibility of a comparison changed between when it was presented in the metaphor form and when it was presented in the simile form by asking participants to judge ease of comprehension after they interpreted the metaphor or the simile. We then tested whether difference in comprehensibility between both forms was explained by the interpretive diversity of the figurative interpretation generated by participants.

Experiment 1

Method

Participants   Ninety undergraduate and graduate students participated as volunteers. All participants were native speakers of Japanese.

Materials   Thirty pairs of Japanese topic and vehicle words were used for the experiment. These pairs were derived from a pilot study.

Pilot study   For a pilot study, we used 40 Japanese topic-vehicle pairs. They were composed of 22 pairs chosen as appropriate for Japanese figurative comparisons from among a list of 30 English pairs used in Chiappe and Kennedy's (1999) experiments, and 18 pairs chosen from a list of expressions frequently used for Japanese metaphors or similes.

Forty undergraduate and graduate students participated in the pilot study. None of them participated in the main study. They were assigned to all the 40 topic-vehicle pairs and asked to rate each pair with respect to similarity on a 5-point scale (1 = *not at all similar*, 5 = *extremely similar*). Based on the ratings, 30 pairs were chosen for the main study such that they differ in similarity as gradually as possible.

Procedure   The main study of Experiment 1 consisted of three tasks: preference rating task, shared feature listing task and salience rating task. Participants were assigned to 30 topic-vehicle pairs and carried out only one of the three tasks. The presentation order of 30 pairs was randomly determined for each participant.

In the preference rating task, 30 participants were asked to read each pair and to rate it on a 5-point scale of metaphor preference (1 = *simile is preferable*, 5 = *metaphor is preferable*). In the shared-feature listing task, 30 participants were asked to read each pair and to list features shared by the topic and the vehicle. For the features listed for each pair, closely

Table 1: Regression analysis of interpretive diversity and similarity for predicting the metaphor preference (n=30).

Variable	B	SE B	β
Interpretive diversity	0.76	0.26	0.46**
Topic-vehicle similarity	0.29	0.13	0.35*

Note. $R^2 = .40$, $F(2, 27) = 8.98$, $p = .001$.
*$p < .05$. **$p < .01$.

related words and phrases were accepted as the same feature. After that, any features listed by only one participant were dropped. As a result, 4 to 13 features per pair (a total of 233 features) were selected as shared. These features were used for the salience rating task. In the salience rating task, 30 participants were assigned to all the 30 pairs with the shared features listed in the feature listing task. They were asked to read each pair and its shared features, and to judge how salient these features were when they were seen as shared by the topic and the vehicle using a 7-point scale (0 = *not at all salient*, 6 = *extremely salient*).

Results and Discussion

For each pair, the similarity rating in the pilot study and the metaphor preference rating in the main study were averaged across participants. The mean similarity rating across 30 pairs was 2.96 (SD = 0.83), ranging from 1.45 (*freedom – the dark*) to 4.25 (*hope – light*). The mean preference rating was 2.82 (SD = 0.69), ranging from 1.57 (*tree – umbrella*) to 4.20 (*history – footprints*).

Interpretive diversity of shared features for each pair was calculated by Equation 1 as follows. First, the salience rating for each feature x was averaged across participants. Then $p(x)$ was assessed in such a way that the mean salience rating of the feature x was divided by the sum of all the mean salience ratings for the topic-vehicle pair. For example, for the comparison "*deserts – ovens*" four shared features (*burning, hot, dry, great change of temperature*) were listed and their mean salience ratings were 4.88, 4.77, 3.42 and 2.65. Hence the value of interpretive diversity for this comparison was calculated as $-(4.88/15.72)\log(4.88/15.72) - (4.77/15.72)\log(4.77/15.72) - (3.42/15.72)\log(3.42/15.72) - (2.65/15.72)\log(2.65/15.72) = 1.95$. The mean interpretive diversity across 30 pairs was 2.87 (SD = 0.42), ranging from 1.95 (*deserts – ovens*) to 3.68 (*life – journey*).

As predicted, the interpretive diversity of shared features was positively correlated with the metaphor preference rating, $r = .53$ ($p < .005$). The similarity rating was also positively correlated with the preference rating, $r = .45$ ($p < .05$). To examine which of the two variables better explains the preference rating, regression analysis was conducted with preference rating as the dependent variable. Table 1 shows the result of the regression analysis. Both interpretive diversity and similarity accounted for a significant portion of the variance in metaphor preference, but the standardized regression coefficient was larger for interpretive diversity than for similarity. Furthermore, the commonality analysis revealed that 50% of the explained variance (i.e., 20% of the total variance)

in metaphor preference was associated uniquely with interpretive diversity, while 28% of the explained variance (i.e., 11% of the total variance) was associated uniquely with similarity. These results suggest that interpretive diversity is a more important factor, but similarity does not lose its role as a supplementary source of the metaphor preference.

Chiappe and Kennedy (2001) interpreted their finding of the positive correlation between topic-vehicle similarity and metaphor preference as indicating that "the classification [i.e., metaphor] form is used when the topic and the vehicle share many properties, whereas the similarity [i.e., simile] form is used when the topic and the vehicle share fewer common properties" (p.250). However, our finding indicates that the correlation between similarity and metaphor preference may not reflect the effect of the number of shared features. Indeed, the correlation between similarity rating and the number of shared features was not significant, $r = .22$. Rather, our finding shows that interpretive diversity embodies such effect and at the same time that although playing a secondary role, similarity has a unique impact on metaphor preference.

Experiment 2

Method

Participants Twenty undergraduate and graduate students participated as volunteers. All participants were native speakers of Japanese.

Materials The 30 topic-vehicle pairs used in Experiment 1 were used for this experiment. Each participant was assigned to all the 30 pairs. The order of the material was randomly determined for each participant.

Procedure Participants were asked to read each topic-vehicle pair and to rate how familiar they were with the expression comparing the topic word of the pair with the vehicle word on a 7-point scale (0 = *not at all familiar*, 6 = *extremely familiar*).

Results and Discussion

The familiarity rating of each pair was averaged across participants. The mean familiarity rating across 30 pairs was 2.93 (SD = 1.27), ranging from 0.53 (*cigarettes – time bombs*) to 4.89 (*life – journey*).

The correlation between familiarity and metaphor preference was positive and significant, $r = .47$ ($p < .01$). However, familiarity was also highly correlated with similarity, $r = .82$ ($p < .0001$). Hence, the regression analysis was conducted with familiarity, similarity and interpretive diversity as the independent variable. The result was that the standardized regression coefficient for familiarity was 0.002 and the squared partial correlation for familiarity was 0.000. This result indicates that familiarity did not have a unique contribution to determining metaphor preference.

This finding on the relation between similarity and familiarity replicates the finding of Chiappe and Kennedy's (2001) study. From this finding, they argued that "Familiarity's role is to enhance similarity" (p.263). If their argument is right, interpretive diversity is predicted to have a greater effect on preference when topic-vehicle pairs are unfamiliar than when they are familiar. To test this prediction, we divided the 30 topic-vehicle pairs into two groups — familiar pairs

Table 2: Summary of regression analysis predicting the metaphor preference for familiar and less familiar comparisons.

Variable	B	SE B	β	p
Familiar pairs (n=16)				
Interpretive diversity	0.73	0.36	0.45	0.06
Topic-vehicle similarity	0.51	0.26	0.43	0.07
Less familiar pairs (n=14)				
Interpretive diversity	0.88	0.43	0.53	0.07
Topic-vehicle similarity	0.23	0.36	0.16	0.54

Note. $R^2 = .38$, $F(2, 13) = 3.90$, $p < .05$ for familiar pairs; $R^2 = .40$, $F(2, 11) = 2.11$, $p = .17$ for less familiar pairs.

(i.e., those whose familiarity rating was the midpoint of 3 or higher) and less familiar pairs (i.e., those whose familiarity rating was less than 3) —, and conducted regression analysis separately for the two groups of pairs. Table 2 shows the two regression equations with interpretive diversity and similarity as independent variables.

The result shown in Table 2 is consistent with the prediction. Concerning familiar pairs, both interpretive diversity and similarity accounted for a marginally significant portion of the variance in preference, and more importantly the effect of similarity on preference was greater than when all the pairs were entered into the regression analysis. Indeed, the commonality analysis showed that the percentage of the variance explained uniquely by similarity increased to 19% (i.e., 50% of the variance explained by both factors), as compared to 11% of the variance in preference for all pairs. On the other hand, for less familiar pairs only interpretive diversity accounted for a marginally significant portion of the variance in preference. It was also found that 98% of the explained variance was associated uniquely with interpretive diversity, while similarity uniquely explained only 10% of the explained variance. [1] From these findings it follows that interpretive diversity was more important in determining metaphor preference for less familiar pairs than for familiar pairs. In other words, people may judge preference of novel topic-vehicle pair, in which preexisting similarity is unlikely to be perceived, primarily on the basis of the interpretive diversity of created similarity.

Experiment 3

Method

Participants One hundred and twenty-two undergraduate students participated for a requirement of computer literacy course. All participants were native speakers of Japanese.

---

[1] It must be noted that common variance, i.e., the variance explained by a joint effect of multiple variables, may have a negative value, which means a lack of a joint effect. For example, the proportion of variance for less familiar pairs shared by interpretive diversity and similarity was −0.074.

Materials Forty pairs of Japanese topic words and vehicle words, different from the pairs used in Experiments 1 and 2, were used for this experiment. These pairs were constructed from ten groups each of which consisted of two topic words and two vehicle words. For example, from the two topics, 'anger' and 'sleep', and the two vehicles, 'sea' and 'storm', four topic-vehicle pairs were created: *anger – the sea, anger – storm, sleep – the sea* and *sleep – storm*. Topic and vehicle words were selected from an experimental study on Japanese metaphor and a list of words frequently used for Japanese metaphors. Topic-vehicle pairs were presented in metaphor form (e.g., "*Anger is the sea*" ["*Ikari ha umi da*" in Japanese]) or in simile form (e.g., "*Anger is like the sea*" ["*Ikari ha umi no you da*" in Japanese]).

Procedure This experiment was conducted separately for metaphor comprehension and simile comprehension. The experiment for metaphor comprehension was a part of larger experiment reported elsewhere (Utsumi 2005).

In the experiment for metaphor comprehension, metaphor comprehension task was preceded by the feature listing task for single words. Eighty participants each were assigned to 10 words for the feature listing task for single words and 10 metaphors which did not include any of the assigned words. In other words, for example, participants who saw a word 'sleep' did not see the 'sleep' metaphors, either "*Sleep is the sea*" or "*Sleep is a storm*". Words and metaphors of each group were counterbalanced so that they were assigned to 20 participants. In the feature listing task for single words, participants were asked to list at least three features of a concept expressed by a single word. In the metaphor comprehension task, participants carried out three subtasks. First, they were asked to consider the meaning of a metaphor and to list at least three features of the topic (presented with an underline) that were being described by the vehicle of that metaphor (feature listing task). Second, they were asked to describe their own interpretation of the metaphor freely by sentences (free description task).[2] Finally, they were asked to rate the metaphor with respect to ease of interpretation on a 7-point scale ranging from 1 (*not at all comprehensible*) to 7 (*extremely comprehensible*) (comprehensibility rating task).

In the experiment for simile comprehension, 42 participants each were assigned to 20 similes which shared neither the topic nor the vehicle (e.g., "*Anger is like the sea*" and "*Sleep is like a storm*"). Similes were counterbalanced across participants so that they were assigned to 21 participants. Each participant carried out the same three subtasks as those of the metaphor comprehension task (i.e., feature listing task, free description task and comprehensibility rating task).

Results and Discussion

For each metaphor and simile, the comprehensibility rating was averaged across participants. The mean comprehensibility rating across 40 metaphors was 4.32 (SD = 1.05), ranging from 1.90 (*perfume – ice*) to 6.14 (*child – jewelry*). The mean comprehensibility rating across 40 similes was 4.52 (SD = 1.00), ranging from 2.09 (*perfume – ice*) to 5.90 (*lover – sun*). Difference of comprehensibility for each topic-

---

[2] In this paper, the date obtained in the free description task was not used in the analysis.

Table 3: Correlations of the comprehensibility difference with the interpretive diversity of figurative meaning.

	Difference of comprehensibility	
	All (n=40)	Comprehensible (n=25)
Interpretive diversity		
Metaphor	0.21	0.47*
Simile	0.43**	0.55**
Sum	0.43**	0.61**

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

vehicle pair was then calculated by subtracting the mean comprehensibility rating for the simile form from the mean comprehensibility rating for the metaphor form. Hence the difference of comprehensibility is positive when the metaphor form is more comprehensible than the simile form. The mean difference of comprehensibility was -0.21 (SD = 0.48), ranging from -1.40 (*character – fire*) to 0.63 (*love – journey*). It was significantly different from 0, $t(39) = 2.73$ ($p < .01$), thus suggesting that similes were comprehended more easily than metaphors.

Interpretive diversity of the metaphorical meaning was calculated separately for metaphors and similes, using the features listed in the feature listing task for metaphor or simile comprehension. First, closely related words or phrases in the same list of features were accepted as the same feature if they met one of some criteria, such as one that they belonged to the same deepest category of a Japanese thesaurus and one that they shared the same root form. After this feature combination process, any features mentioned by only one participant were eliminated from the list of features. Finally, for each feature x in the amended list of features, $p(x)$ of Equation 1 was assessed by dividing the number of tokens of that feature (i.e., the number of participants who listed that feature) by the total number of tokens involved in the list. The mean interpretive diversity across 40 pairs was 3.01 (SD = 0.42) when the pairs were presented in the metaphor form, and 3.08 (SD = 0.41) when the pairs were presented in the simile form.

To examine whether difference in comprehensibility between metaphors and similes is explained by interpretive diversity of figurative meaning, we correlated difference of comprehensibility with three values, i.e., metaphor's diversity, simile's diversity and the sum of both diversities. The result of correlation is presented in Table 3. Overall, comprehensibility difference was positively correlated with the interpretive diversity of the figurative meaning; two out of three correlations were significant. The metaphor form was rated as more comprehensible than the simile form when the interpretive diversity was high, whereas the simile form was perceived as more comprehensible when the interpretive diversity was low. Furthermore, to avoid an undesirable effect of semantically irrelevant features which may be included in the list of features generated for incomprehensible or less comprehensible pairs, we limited the correlation analysis to comprehensible pairs whose comprehensibility rating for both forms was the midpoint 4 or higher. The result was
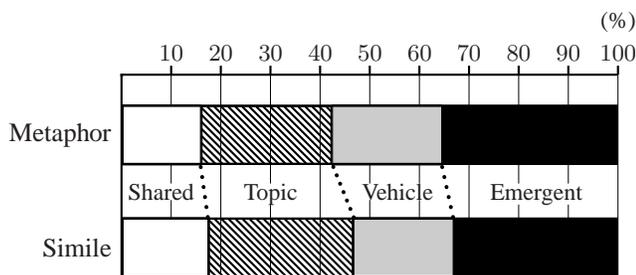
Figure 1: Comparison of the proportions of four feature types in the figurative interpretation between metaphors and similes.

that, as shown in Table 3, higher correlations were observed between comprehensibility difference and interpretive diversity, and all the correlations became significant. These findings clearly show that the metaphor form is more comprehensible than the simile form to the extent that a potential figurative meaning is diverse.

One question that arises here is whether or not the generated figurative meanings differ between metaphors and similes, in other words, whether the grammatical form affects comprehension of figurative statements. The negative result was that there was not a significant difference between metaphor's diversity and simile's diversity, $t(39) < 1$. In addition, the difference of comprehensibility was not correlated with the difference of interpretive diversity between both forms (i.e., metaphor's diversity minus simile's diversity), $r = -.16$.

In contrast, a positive result was also obtained on qualitative difference of figurative meaning; more shared and topic features were generated for interpretation of similes than for interpretation of metaphors, as shown in Figure 1. Shared features were those listed for both the topic and the vehicle in the feature listing task for single words, and more shared features were generated for similes than for metaphors, $t(39) = 2.15$ ($p < .05$). [3] Topic features were those listed only for the topic and more topic features were generated for similes than metaphors, $t(39) = 2.45$ ($p < .05$). Concerning other types of features — vehicle features (i.e., those listed only for the vehicle) and emergent features (i.e., those listed for neither the topic nor the vehicle) —, there was no significant difference in the number of features between similes and metaphors.

These findings seem to suggest that metaphors may be comprehended so that its figurative interpretation diverges to multiple equally salient meanings, while that similes may be comprehended so that its figurative interpretation converges into one or only a few highly salient meanings. The rationale behind this suggestion is that "shared features" here can be assumed to be much more salient than other types of features on the grounds that shared features were listed by many participants although the number of different shared features

was very small (Nueckles and Janetzko 1997; Utsumi 2005). [4] Therefore, the finding that more shared features were generated when topic-vehicle pairs were interpreted as similes than when interpreted as metaphors may mean that simile comprehension is geared to a few highly salient meanings.

General Discussion

The three experiments reported in this paper provided empirical evidence in favor of our argument that interpretive diversity is an important source of metaphor-simile distinction. Moreover, interpretive diversity was found to be more important in determining metaphor preference than Chiappe and Kennedy's (2001) topic-vehicle similarity. At the same time, however, similarity had a unique effect on preference which could not be achieved by interpretive diversity.

Then, what is a source of the unique contribution of similarity? We speculate that the source of such unique effect may involve the total salience of shared features, upon which interpretive diversity does not depend. For example, when two topic-vehicle pairs $X1 - Y1$ and $X2 - Y2$ have the equal number of equally salient shared features, they are equal in interpretive diversity. However, imagine further that the features for $X1 - Y1$ are more salient than the features for $X2 - Y2$. In this case, they are still equal in interpretive diversity, but they differ in the total salience and possibly in similarity; it seems reasonable to suppose that the pair $X1 - Y1$ is more similar than the pair $X2 - Y2$. This assumption may be supported by the result of Experiment 1. The similarity rating was significantly correlated with the total salience of shared features (i.e., the sum of the salience rating for each pair), $r = .43$ ($p < .05$), although it was not correlated with the number of shared features, $r = .22$.

In this paper, we did not address an effect of the conventionality of the vehicle (Bowdle and Gentner 1999; Zharikov and Gentner 2002), but it deserves special mention. The theory underlying these studies is the career of metaphor (Bowdle and Gentner 2005; Gentner et al. 2001) which reconciles two rival theories of metaphor, the structure mapping theory (Gentner 1983) and the attributive category theory (Glucksberg 2001). The career of metaphor hypothesis claims that figurative statements with novel vehicles (i.e., their figurative meanings are not conventionalized) are processed as comparisons, but conventional figuratives (i.e., the figurative meanings of the vehicles are conventional) are processed as categorizations (but also comparisons). Therefore, it predicts that similes are preferred when vehicles are novel, and metaphors become preferred when vehicles are conventionalized.

However, as Chiappe and Kennedy (2001) pointed out, the conventionality of the vehicle alone cannot provide a reasonable explanation for how metaphor preference (and comprehensibility difference) differs among different topic-vehicle pairs. Even though a vehicle has one conventional figurative meaning, other possible figurative meanings of that vehicle are not necessarily conventional. Hence figurative statements

---

[3]Features listed in the feature listing task for single words were preprocessed in the same way as the features listed for metaphors and similes. For each word (a topic or a vehicle), closely related words were combined into one feature, and any features mentioned by only one participant were eliminated from the list of features for that word. After these lists of features were generated, the features in the amended list for metaphors and similes were classified into the four types of features.

[4]It must be noted here that the shared features in Experiment 3 differ from the shared features in Experiment 1. The latter shared features were obtained when a topic and a vehicle were presented together (and thus they might capture a whole figurative meaning), but the former ones were listed both for a topic and for a vehicle when they were presented separately. Hence, the shared features here can be seen as most salient of the shared features in Experiment 1.

with the same vehicle can differ in preference when the topics make different properties relevant to the comparison. In this respect, familiarity of topic-vehicle pairs, not the familiarity of vehicles irrespective of topics, is more appropriate for examining the effect of the conventionality of figurative meanings. And, as Experiment 2 showed, the familiarity of topic-vehicle pairs, too, may not be a primary source of metaphor-simile distinction. Rather, the interesting question for further research is what properties of topic-vehicle pairs have a joint effect with familiarity or conventionality. For example, Nakamoto and Kusumi (2004) focused on this question, and showed that the effect of conventionality of the vehicle on metaphor preference was greater when comparisons were highly apt than when they were less apt.

Nevertheless, we suggest that interpretive diversity may be consistent with the comparison-categorization distinction advocated by the career of metaphor. Novel topic-vehicle pairs are processed as comparison involving structural alignments between the topic and the vehicle, and thus it will succeed even when only a few highly salient features or predicates are common to both concepts, i.e., even when they are interpretively less diverse. On the other hand, conventional or familiar pairs are processed as categorization which requires direct mapping of many salient features involved in the figurative meaning of the vehicle into the topic; categorization statements are inappropriate if only some of the features can be mapped to the topic. Therefore, the pairs for which the metaphor form is preferred must be interpretively diverse. Perhaps the reverse may be possible. Figurative statements with diverse meanings can be comprehended as categorization expressing that the topic is a member of an ad hoc superordinate category exemplified by the vehicle, as argued by Glucksberg (2001), but if they have only a few salient meanings they cannot be comprehended as categorization. Hence, in such cases, they are likely to be comprehended as comparison.

Finally, we should also mention the third existing account of metaphor preference. Aisenman (1999) proposed that metaphor-simile distinction is closely tied to the type of predicates shared by the topic and the vehicle: If shared features are primarily relational then the metaphor form is preferred, while if attributional predicates are shared then the simile form is preferred. This hypothesis receives some support from experimental studies examining how relationality and conventionality affect the metaphor preference (Gokcesu and Bowdle 2003; Zharikov and Gentner 2002). Although how relationality is associated with interpretive diversity must await further research, we offer a possible speculation. Relational predicates may encourage a rich variety of mappings between the topic and the vehicle, and thereby the figurative meaning is likely to be diverse and the metaphor form is preferred. Attributional predicates, on the other hand, are salient on their own, and thus the figurative meaning tends to be less diverse and the simile form is preferred.

To conclude, our findings suggest that interpretive diversity can be a better candidate for the source of metaphor-simile distinction. It would be interesting and vital for further research to examine which of the properties for topic-vehicle pair — interpretive diversity, similarity, aptness, conventionality, relationality — dominates metaphor-simile distinction and how these properties interact with each other.

References

Aisenman, R. (1999). Structure-mapping and the simile-metaphor preference. *Metaphor and Symbol*, *14*(1), 45–51.

Bowdle, B. and Gentner, D. (1999). Metaphor comprehension: From comparison to categorization. In *Proceedings of the 21st Annual Conference of the Cognitive Science Society*, pp. 90–95.

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

Chiappe, D. and Kennedy, J. (1999). Aptness predicts preference for metaphors or similes, as well as recall bias. *Psychonomic Bulletin & Review*, *6*, 668–676.

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

Chiappe, D., Kennedy, J., and Chiappe, P. (2003). Aptness is more important than comprehensibility in preference for metaphors and similes. *Poetics*, *31*(1), 51–68.

Gentner, D. (1983). Structure mapping: A theoretical framework for analogy. *Cognitive Science*, *7*, 155–170.

Gentner, D., Bowdle, B., Wolff, P., and Boronat, C. (2001). Metahor is like analogy. In Gentner, D., Holyoak, K., and Kokinov, B. (Eds.), *Analogical Mind: Perspectives from Cognitive Science*, pp. 199–253. MIT Press.

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

Glucksberg, S. (2001). *Understanding Figurative Language: From Metaphors to Idioms*. Oxford University Press.

Gokcesu, S. and Bowdle, B. (2003). Conventionality, similarity, and the metaphor/simile distinction. In *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*, p. 1350.

Nakamoto, K. and Kusumi, T. (2004). The effect of repeated presentation and aptness of figurative comparisons on preference for metaphor forms. In *Proceedings of the 26th Annual Meeting of the Cognitive Science Society*, p. 1611.

Nueckles, M. and Janetzko, D. (1997). The role of semantic similarity in the comprehension of metaphor. In *Proceedings of the 19th Annual Conference of the Cognitive Science Society*, pp. 578–583.

Utsumi, A. (2005). The role of feature emergence in metaphor appreciation. *Metaphor and Symbol*, *20*(3), in press.

Zharikov, S. and Gentner, D. (2002). Why do metaphors seem deeper than similes?. In *Proceedings of the 24th Annual Conference of the Cognitive Science Society*, pp. 976–981.