

# Effects of Discourse Goals on the Process of Metaphor Production

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

Kota Nakamura (kota@utm.inf.uec.ac.jp)

Maki Sakamoto (sakamoto@inf.uec.ac.jp)

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

## Abstract

Only a few attempts have so far been made at exploring the process of metaphor production, although a large number of studies have addressed metaphor comprehension. Therefore, in this paper, we address the problem of how people generate metaphors or identify an apt vehicle for a given topic of metaphors. Specifically, we examine how the process and product of metaphor production differ between two discourse goals of metaphor, namely an explanatory purpose (e.g., to clarify) and a literary purpose (e.g., to aesthetically pleasing). Experiment 1 analysed the metaphors (or vehicles) generated in the metaphor production task, and demonstrated that people identified more prototypical exemplars of the property attributed to the topic as a vehicle for explanatory metaphors than for literary metaphors. In addition, it was found that explanatory metaphors were more apt and conventional, and had high topic-vehicle similarity than literary metaphors, while literary metaphors were more familiar and imageable than explanatory metaphors. Experiment 2 used a priming paradigm to assess the online availability of prototypical and less prototypical members of the topic property during metaphor production. The result was that both prototypical and less prototypical members were activated in producing literary metaphors, while neither members were activated in the production of explanatory metaphors. These findings indicate that the process of metaphor production is affected by discourse goals of metaphor, and suggest that only prototypical members of the category are rapidly searched for a vehicle during the production of explanatory metaphors, while both prototypical and less prototypical members are searched to generate literary metaphors.

**Keywords:** Metaphor production; Discourse goal; Priming

## Introduction

Metaphor has been a main topic of research in cognitive science, because metaphorical expressions are frequently observed in our everyday use of language. Hence, a large number of studies have been made on how people comprehend metaphors (e.g., Bowdle & Gentner, 2005; Gibbs, 2008; Glucksberg, 2001; Utsumi, 2011). In contrast, only a few studies have addressed the process of metaphor production (for a notable exception, see, for example, Chiappe & Chiappe, 2007; Katz, 1989). This paucity of research on metaphor production is especially problematic, given that by its very nature a metaphor is an intentional, linguistic device employed to convey ideas that might be otherwise inexpressible. To ameliorate this situation, in this paper, we empirically explore the process of metaphor production.

Previous studies on metaphor production analysed the products of metaphor production (i.e., metaphorical expressions produced) in terms of the qualitative dimensions of metaphors and/or individual differences. Concerning the qualitative dimensions of metaphor products, Fainsilber and Ortony (1987) demonstrated that descriptions of emotional states

contained more metaphorical language than did descriptions of actions. Katz (1989) examined the properties of metaphor vehicles by asking participants to choose, from a set of alternatives, a vehicle that completes a given sentence frame (e.g., Chemistry is the \_\_\_\_\_ of science) as comprehensible and apt metaphors. The result was that participants were likely to choose the vehicles that were moderately distant from the topic and referred to concrete domains. Concerning individual differences, it was demonstrated that the quantity and quality of metaphor products were affected by individual differences such as writing experience (experienced or novice) (Williams-Whitney, Mio, & Whitney, 1992), reasoning and imagery ability (Katz, 1989), gender (male or female) (Hussey & Katz, 2006), and working memory capacity (Chiappe & Chiappe, 2007). For example, Chiappe and Chiappe (2007) demonstrated that people with high working memory capacity produced more apt metaphors than did low capacity individuals.

Although shedding light on the specific aspects of metaphor production, these studies did not address one important aspect of metaphor production, namely *discourse goals of metaphor*. Because metaphors (and other figurative language) are intentionally used to accomplish certain communication goals (Roberts & Kreuz, 1994), it is obviously crucial to explore the effects of discourse goals on the process of metaphor production. Discourse goals that are accomplished by the use of metaphorical expressions can be classified broadly into two classes: explanatory and literary purposes (Steen, 1994; Utsumi, 2005). These two goals are quite different and sometimes incompatible with each other; explanatory metaphors are used to clarify certain properties of the topic, while literary metaphors are used to evoke an aesthetically pleasing feeling by enriching the meanings conveyed by the metaphors. It naturally follows that discourse goals are likely to affect the process of generating metaphors, or more specifically choosing the vehicles of metaphors. This paper aims at examining how the process and product of metaphor production differ between these two discourse goals.

In metaphor production, people often have in mind a topic that they want to express and some properties that they intend to attribute to the topic. They must identify an appropriate or apt vehicle to convey the intended meaning (i.e., the information that the topic has the property). Hence, metaphor production essentially involves the process of searching for or retrieving an apt vehicle. According to the categorization (or attributive category) theory of metaphor (Glucksberg, 2001; Glucksberg & Keysar, 1990), apt vehicles must not only have the intended property but also be a prototypical exemplar of

that property. For example, consider the topic “rumor” and the property “quickly spreads from person to person.” People have to identify an appropriate vehicle to express by metaphor that a rumor quickly spreads from person to person. This property is true of a number of things such as “virus,” “swine flu,” and “louse,” but a virus is a prototypical exemplar of the category “things that quickly spread from person to person.” Hence, the metaphor “The rumor is a virus” seems to be more apt than “The rumor is a swine flu” and “The rumor is a louse.”

The research question to be answered here is how people identify or select a vehicle that accomplishes their communication purposes. This question can be rephrased as how the process of searching a set of things (i.e., search space) for a vehicle differs according to whether explanatory metaphors or literary metaphors are intended. To tackle this problem, we consider the observed differences between explanatory and literary metaphors. Some studies have found that semantic aptness (Steen, 1994; Utsumi, 2005), clarity (Gentner, 1982), and interpretive richness (Gentner, 1982; Utsumi, 2005) are distinctive properties for distinguishing between explanatory and literary metaphors. Explanatory metaphors are more apt (or appropriate) and clearer than literary metaphors, while literary metaphors are interpretively richer than explanatory metaphors. These findings suggest that semantically apt and clear vehicles may be preferably searched for in producing explanatory metaphors, while less apt vehicles that enrich the metaphorical meaning may be searched for in producing literary metaphors. According to the attributive category theory, vehicles of apt metaphors are prototypical members of the category characterized by the intended property of the topic (Glucksberg & Keysar, 1990). On the other hand, for a metaphor to be semantically rich and involve a number of metaphorical interpretations, its vehicle must be less prototypical because highly prototypical vehicles evoke only the intended property.

From these discussions, we can derive the following hypothesis about the process of metaphor production:

**Hypothesis about process:** *Only prototypical members of the category are searched for a vehicle during the production of explanatory metaphors, while less prototypical members are also considered to generate literary metaphors.*

This hypothesis presupposes that people first search prototypical members of the category before searching less prototypical members, regardless of discourse goals (e.g., Giora, 2003; Rosch & Mervis, 1975). Hence, this hypothesis implies that when producing explanatory metaphors, they do not have to search less prototypical members, because prototypical members are sufficient for apt vehicles. However, when producing literary metaphors, people have to search less prototypical members after prototypical members are activated.

The hypothesis about process implies another hypothesis about the metaphors (or vehicles) produced.

**Hypothesis about products:** *More prototypical and apt vehicles are chosen for explanatory metaphors than for literary metaphors.*

Considering that prototypicality and other related properties such as conventionality and familiarity can be classified into a more general notion of salience (Giora, 2003), it is also predicted that conventional and familiar vehicles are preferably chosen for explanatory metaphors. In addition, metaphor aptness refers to the extent to which the vehicle’s metaphoric category captures an important feature of the topic, and thus it reflects the similarity between the vehicle and topic of a metaphor (e.g., Chiappe & Kennedy, 1999). Hence, we also predict that vehicles chosen for explanatory metaphors are more similar to the topic than those chosen for literary metaphors.

In this paper, we test these hypotheses through two experiments, namely, an offline generation experiment and an online priming experiment. In the metaphor generation experiment (i.e., Experiment 1), we examined the validity of the hypothesis about products by analysing the vehicles generated for explanatory or literary metaphors. We compared the vehicles for explanatory metaphors and literary metaphors in terms of several factors: vehicle prototypicality, metaphor conventionality, metaphor aptness, topic-vehicle similarity, vehicle familiarity, vehicle word frequency, and vehicle imageability. In the online priming experiment (i.e., Experiment 2), we used a priming paradigm to test the hypothesis about the internal process of metaphor production, particularly to examine what words (or concepts) are activated during identifying an appropriate vehicle. In the priming experiment, participants were presented with a topic (i.e., an incomplete metaphorical sentence) and its property as a prime and asked to determine a vehicle appropriate for a given discourse goal. Afterwards, a target word was presented and participants were asked to make a lexical decision about it. The target conditions were a word that was highly prototypical of the category made up of the property, a word that was less prototypical of the category, and a control word unrelated to the property.

It must be noted here that the previous studies mentioned earlier analysed only the products of metaphor production using an offline paper-and-pencil experiment paradigm; they did not directly examine the online processes of metaphor production. To the best of our knowledge, this is the first study to apply an online priming paradigm to the study of metaphor production.

## Experiment 1

In Experiment 1, we tested the hypothesis about the product by collecting metaphor vehicles in a metaphor generation task and the ratings of their properties in a vehicle rating task.

### Method

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

**Materials** Twenty pairs of a topic and a property to be attributed to the topic were used for the experiment. Topic words were selected from Japanese abstract nouns comprising two kanji characters, and their properties were expressed in a Japanese short phrase referring to a salient feature of the

topic. In order to eliminate an undesirable effect of the abstractness of the topic words on the process of categorization (Glucksberg & Keysar, 1990), we equalized the degree of abstractness of the topic words by choosing them from abstract categories of words at almost the same depth (i.e., depth of 7 or 8) in the hierarchical structure of the Japanese thesaurus. For example, the topic word “plan” (“*keikaku*” in Japanese) was paired with the property “does not always go as scheduled” (“*yotei doori-ni ikanai*” in Japanese).

**Procedure** This experiment comprised two tasks, namely a metaphor generation task and a vehicle rating task.

The metaphor generation task was conducted by 20 participants. Each participant was assigned all the 20 topic-property pairs, one half of which were used to generate explanatory metaphors and the other half of which were used to generate literary metaphors. Topic-vehicle pairs were counterbalanced across conditions so that each pair appeared 10 times in both conditions. The presentation order of the pairs in each group was randomized for each participant. Participants, who were run individually, were seated in front of a computer screen. They were first given an overall instruction of the experiment and presented with four practice trials (two for explanatory metaphors and two for literary metaphors) followed by two groups of 10 experimental trials. In the explanatory metaphor condition, participants were instructed to generate a vehicle that clearly explains the given property of the topic, while in the literary metaphor condition they were instructed to generate a metaphor aesthetically pleasing enough to use in literary works. On each trial, they were presented with a sentence frame including a topic (e.g., “A plan is (like) a \_\_\_\_\_”) and a property (“does not always go as scheduled”) in the center of the screen, and asked to generate an apt vehicle that completes the sentence at their own pace. When participants came up with a suitable vehicle, they indicated it by pressing the appropriate key on the keyboard. Afterwards, they typed the vehicle as quickly as possible. Reaction times were measured from the onset of the topic-property pair until the appropriate key was pressed (metaphor production time), and from the key press until the input of the vehicle was completed (vehicle typing time).

In the vehicle rating task, another 20 participants were presented with all the 20 topic-property pairs and their vehicles generated (i.e., metaphors and their property attributed to the topic) in the metaphor generation task. They were asked to rate each metaphor or vehicle on the following four 7-point scales: vehicle prototypicality (7 = *prototypical*, 1 = *not at all prototypical*), metaphor conventionality (7 = *conventional*, 1 = *novel*), metaphor aptness (7 = *apt*, 1 = *not at all apt*), and topic-vehicle similarity (7 = *similar*, 1 = *dissimilar*).

## Results and Discussion

First, we analysed the metaphor production time for explanatory and literary metaphors in the metaphor generation task. We eliminated from the analysis extreme outliers (i.e., production times shorter than 2s or longer than 40s, and production times of the trial whose vehicle typing times were longer than 30s), and averaged the remaining production times. As

Table 1: Means ( $M$ ) and standard deviations ( $SD$ ) of seven factors of explanatory and literary metaphors in Experiment 1

Factor	Explanatory		Literary	
	$M$	$SD$	$M$	$SD$
Vehicle prototypicality***	5.14	0.64	4.70	0.43
Metaphor conventionality***	4.21	0.90	3.31	0.60
Metaphor aptness*	5.16	0.67	4.86	0.42
Topic-vehicle similarity***	3.98	0.43	3.60	0.59
Vehicle familiarity*	5.90	0.33	6.09	0.28
Vehicle word frequency	3.14	0.52	3.17	0.54
Vehicle imageability*	5.05	0.33	5.32	0.44

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

a result, literary metaphors ( $M=77.39s$ ,  $SD=50.18s$ ) took longer to generate than explanatory metaphors ( $M=55.82s$ ,  $SD=29.74s$ ). This difference was marginally significant in the participant analysis,  $F_p(1, 19) = 4.18$ ,  $p = .055$ , although not significant in the item analysis,  $F_i(1, 19) = 2.13$ ,  $p > .1$ . (Note that, in all the ANOVAs reported in this paper, the data were analyzed by participants  $F_p$  and by items  $F_i$ .) This result is consistent with the hypothesis about the production process, in that people must search less prototypical members of the category after searching prototypical members, and consequently require more time to generate literary metaphors than explanatory metaphors. Note, however, that this result may not be reliable enough to warrant the hypothesis, because we had no efficient methods for confirming that participants really spent all the time producing metaphors. (Hence, Experiment 2 was conducted to collect reliable data on the internal process of metaphor production.)

We then compared the explanatory and literary metaphors generated in terms of four factors (i.e., vehicle prototypicality, metaphor conventionality, metaphor aptness, and topic-vehicle similarity). Table 1 shows the mean rating values of these four factors for explanatory and literary metaphors. Explanatory metaphors were rated as significantly higher on all the four factors than literary metaphors,  $F_p(1, 19) = 44.24$ ,  $p < .001$ ,  $F_i(1, 19) = 9.44$ ,  $p < .01$  for vehicle prototypicality;  $F_p(1, 19) = 1256.76$ ,  $p < .001$ ,  $F_i(1, 19) = 16.76$ ,  $p < .001$  for metaphor conventionality;  $F_p(1, 19) = 4.76$ ,  $p < .051$ ,  $F_i(1, 19) = 3.70$ ,  $p = 0.70$  for metaphor aptness; and  $F_p(1, 19) = 23.53$ ,  $p < .001$ ,  $F_i(1, 19) = 8.49$ ,  $p < .01$  for topic-vehicle similarity. These results are entirely consistent with the hypothesis about products, confirming that more prototypical and apt vehicles are chosen for explanatory metaphors than for literary metaphors.

Furthermore, we exploratorily analysed three additional factors of vehicle words — i.e., vehicle word familiarity, vehicle word frequency, and vehicle word imageability — that may differ between explanatory and literary metaphors. These values were derived from the database of Japanese lexical properties “Nihongo No Goi Tokusei.” In this database, the familiarity and imageability of a word were given as the mean rating scores on a 7-point scale ranging from 1 to 7, and the frequency of a word was given as the number of times the word occurred in a newspaper corpus. In the analysis, these

scores were used for vehicle familiarity and imageability, and the common logarithm of the frequency score was used for vehicle word frequency. The last three rows of Table 1 show the mean values of these three factors across vehicle words. Two out of three factors were significantly different; the vehicles of literary metaphors were more familiar and imageable than those of explanatory metaphors,  $F_i(1, 19) = 5.25, p < .05$  for vehicle familiarity; and  $F_i(1, 19) = 5.52, p < .05$  for vehicle imageability. The finding on vehicle imageability is intuitively acceptable, because literary works often evoke mental imagery, although Katz, Paivio, and Marschark (1985) found an opposite result that the vehicles of poetic metaphors were less imageable. On the other hand, the finding of vehicle familiarity is seemingly surprising, but can be explained by the assumption that the use of familiar words as a vehicle may enrich the interpretation of metaphors. Note that this finding is consistent with the finding of Katz et al. (1985) that poetic metaphors were rated as being more familiar.

## Experiment 2

In Experiment 2, we tested the prediction about the metaphor production process using a priming paradigm, in which an incomplete metaphorical sentence was presented first with the property to be attributed to the topic, and the task was to make a lexical decision about a target word presented after the metaphorical sentence. The target conditions were a word highly prototypical of the category characterized by the property (HPT), a word less prototypical of the category (LPT), and a control target (CNT) that is unrelated to the metaphor.

Faster lexical decisions in comparison with the CNT indicate online activation. If only the prototypical members of the category are searched for a metaphor vehicle, the HPT would be faster to make a lexical decision than the CNT, but the LPT would not be faster. If less prototypical members are also searched, both the LPT and HPT would be faster than the CNT. Hence, if our hypothesis about the process is right, only the HPT would be facilitated when explanatory metaphors are intended, while both the HPT and LPT would be facilitated when literary metaphors are generated.

## Method

**Participants** Fifty-four undergraduate and graduate students participated as volunteers. All participants were native speakers of Japanese.

**Materials** Thirty Japanese pairs of a topic and a property attributed to the topic (including all the 20 pairs used in Experiment 1) were used as primes. For each prime pair, three target words (i.e., HPT, LPT, and CNT) comprising two kanji characters were prepared. The HPTs and LPTs were selected on the basis of ratings collected from independent groups of participants in a norming study. The CNTs were selected randomly from a dictionary so that they were unrelated to both the topic and the property, and their word frequency was approximately equal to that of the HPTs and LPTs.<sup>1</sup> For exam-

<sup>1</sup>The mean logarithms of word frequency score were 3.41 for the HPT, 3.62 for the LPT, and 3.60 for the CNT. They were not significantly different,  $F(2, 58) = 0.492$ .

ple, the pair of the topic “all-night activity” (“*tetsuya*”) and its property “unhealthy” (“*karada ni aku eikyo wo ataeru*”) was combined with the HPT “drug” (“*mayaku*”), the LPT “edacity” (“*oogui*”), and the CNT “speech” (“*enzetsu*”). In addition, another 20 topic-property pairs were prepared and used as filler sentences for nonword targets.

**Norming study** For each of the 30 topic-property pairs, six words (comprising two kanji characters) were prepared that referred to an object or a concept with that property but different degree of prototypicality. Ten undergraduate students rated these words on the 7-point scale of prototypicality ranging from 1 (*not at all prototypical*) to 7 (*prototypical*). For each topic-property pair, the word with the highest rating was selected as an HPT, and the word rated as the lowest of the words with the prototypicality degree of 4 (i.e., the midpoint) or higher was selected as an LPT. The prototypicality rating of the HPTs ( $M = 5.91, SD = 0.51$ ) was significantly higher than that of the LPTs ( $M = 4.37, SD = 0.35$ ),  $F(1, 29) = 211.23, p < .001$ , confirming that LPTs and HPTs were appropriately selected.

**Procedure** A within-participants design was used with each participant processing all the 50 topic-property pairs (i.e., 30 pairs with word targets and 20 pairs with nonword targets) under all conditions. The 50 pairs were equally divided into two groups, each of which comprised 15 pairs with word targets and 10 pairs with nonword targets. One group was used to generate explanatory metaphors and another group was used to generate literary metaphors. Topic-vehicle pairs with word targets were counterbalanced across all conditions (i.e., two conditions of discourse goal and three target conditions) so that each pair appeared an equal number of times in all conditions. The presentation order of two groups of discourse goals (i.e., whether explanatory or literary metaphors were generated first) and the order of the pairs in each group were randomized for each participant.

Participants, who were run individually, were seated in front of a computer screen. They were first given an overall instruction of the experiment and then presented with four practice trials (two for explanatory metaphors and two for literary metaphors) followed by two groups of 25 experimental trials. On each trial, they were presented with a sentence frame with a topic in the subject position (“A plan is (like) a \_\_\_\_\_”) and a property (“does not always go as scheduled”) as a prime in the center of the screen for 7000 ms and asked to consider an appropriate vehicle word to fill in the blank. A target word (HPT, LPT, CNT, or nonword) was then presented 500ms after the offset of the topic-property pair. Participants were asked to decide whether the target word was a word or a nonword as quickly as possible; they indicated decision by pressing the appropriate key on the keyboard. Finally, they typed the vehicle that occurred to them before the lexical decision. Reaction times were measured from the onset of the target word until the appropriate key was pressed.

## Results and Discussion

Only reaction times of correct decision were used in the analysis. In addition, reaction times greater than 10,000ms were

Table 2: Means ( $M$ ) and standard deviations ( $SD$ ) of correct lexical decision times in milliseconds for Experiment 2

Metaphor condition	HPT (High Prototypicality)			LPT (Low Prototypicality)			CNT (Control)	
	$M$	$SD$	$DIF$	$M$	$SD$	$DIF$	$M$	$SD$
Explanatory	1450	550	51	1474	584	27	1501	727
Literary	1431	475	155	1303	420	283	1586	743

Note.  $DIF$  = difference from control target.

eliminated from the analysis.

Table 2 shows mean lexical decision times and standard deviations for the correct “yes” responses. The time difference ( $DIF$ ) from the CNT indicates the extent of a priming effect. In the explanatory metaphor condition, the HPT produced a moderate priming effect (51ms faster than the CNT), but the LPT showed only a small priming effect (27ms faster). On the other hand, in the literary metaphor condition, both targets showed a much larger priming effect. In particular, the priming effect of the LPT (283ms) was larger than that of the HPT (155ms), indicating that less prototypical exemplars of the category were activated during the production of literary metaphors. These results are entirely consistent with our hypothesis that only prototypical members of the category are searched for a vehicle during the production of explanatory metaphors, while less prototypical members are also considered to generate literary metaphors.

To confirm these differences statistically, we conducted a two-way ANOVA of Target (HPT, LPT, or CNT)  $\times$  Discourse goal (explanatory or literary) on lexical decision times. These two factors were within participants and within items. First of all, the main effect of Discourse goal was not significant in either analyses, but the main effect of Target was significant in the participant analysis,  $F_p(2, 106) = 4.54, p < .05$ . Post-hoc pairwise comparisons showed that the priming effect of the LPT was significant ( $p < .05$ ) and the priming effect of the HPT was marginally significant ( $p = .051$ ). This result indicates that discourse goals did not affect the overall processing time of lexical decision and the priming effect was observed in this experiment. It confirms that this priming experiment was successful and the priming methodology can be applied effectively to the study of metaphor production.

The most important result was that the interaction between two factors was significant in the participant analysis,  $F_p(2, 106) = 3.40, p < .05$ , although not significant in the item analysis. The nature of this interaction was that, the simple main effect of Target was significant in the literary metaphor condition,  $F_p(1, 159) = 4.57, p < .05$ , but not significant in the explanatory metaphor condition. Post-hoc pairwise comparisons ( $p < .05$ ) revealed that both the LPT ( $M = 1303$ ms) and HPT ( $M = 1431$ ms) were significantly faster than the CNT ( $M = 1586$ ms). Again, these results are consistent with the hypothesis of this paper, although the absence of a significant priming effect of the HPT in the explanatory metaphor condition was not consistent with the hypothesis.

## General Discussion

The two experiments reported in this paper provided empirical evidence in favor of our view that discourse goals of metaphor affect the process of metaphor production. Specifically, we focused on two types of metaphors — i.e., explanatory metaphors and literary metaphors — that accomplish different discourse goals, and demonstrated that the production of literary metaphors required activation of both prototypical and less prototypical members of the category characterized by the topic property, while the production of explanatory metaphors did not. This processing difference leads to the finding on the products that explanatory metaphors generated in the metaphor production experiment were more prototypical than literary metaphors. In addition, it was found that explanatory metaphors were more conventional and apt, and had higher topic-vehicle similarity, while the generated vehicles for literary metaphors were more familiar and imageable.

The finding that less prototypical members of the category were activated during the processing of literary metaphors and as a result less prototypical vehicles are selected for literary metaphors is especially interesting, because it indicates that some metaphors cannot be explained by the attributive category theory of metaphor (Glucksberg, 2001; Glucksberg & Keysar, 1990). We have demonstrated elsewhere that the comprehension of predicative metaphors (i.e., figurative expressions that involve the metaphorical use of a verb or an adjective) cannot be explained by the attributive category theory of metaphor, and proposed an indirect categorization theory of metaphor (Utsumi & Sakamoto, 2007, 2011). The finding of this study suggests a possibility that some nominal metaphors (in particular, literary metaphors) may be processed by an indirect categorization or other mechanisms.

On the other hand, the finding that prototypical members were not activated during the production of explanatory metaphors is not consistent with the hypothesis. One possible explanation of this result would be that the process of identifying a vehicle for explanatory metaphors was so rapid that prototypical members were no longer activated 7500 ms after the onset of the prime. To test this possibility, we must repeat a priming experiment by varying the stimulus-onset asynchrony (SOA) of prime and target, which is left for future research.

The discourse goals of metaphorical or figurative expressions are affected by their grammatical form, i.e., whether figurative comparisons are expressed in metaphor form “An X is a Y” or in simile form “An X is like a Y.” Glucksberg and Keysar (1990) argued that metaphors are more forceful

than similes and alert an addressee that a specific set of properties is intended. Roberts and Kreuz (1994) demonstrated that metaphors and similes differ in that metaphors are used to add interest, but similes are used to deemphasize. If the discourse goals really affect the process of metaphor production as demonstrated in this paper, these pragmatic differences between metaphor and simile imply that the grammatical form has an important influence on the process of metaphor production. Furthermore, a number of studies have revealed that several properties addressed in Experiment 1 (e.g., aptness, conventionality, topic-vehicle similarity) determine the preferred form of figurative comparison (e.g., Bowdle & Gentner, 2005; Chiappe & Kennedy, 1999; Jones & Estes, 2006). Hence, the finding of Experiment 1 may serve to explain the relation between the grammatical form and the production process.

Computational modeling is also an efficient methodology for the study of metaphor production. Recently, computational studies based on a semantic space model such as latent semantic analysis (LSA) have addressed metaphors and shed new light on the process of metaphor comprehension (e.g., Kintsch, 2000; Utsumi, 2011; Utsumi & Sakamoto, 2007). In particular, the predication algorithm in a semantic space model proposed by Kintsch (2000) can embody the attributive category theory, and has been shown to achieve good performance of simulating the process of metaphor comprehension. It follows that the study of metaphor production can also benefit from the same computational modeling framework. For example, the predication algorithm computes the meanings of metaphors by combining neighbors of the vehicle (i.e., word vectors similar to the vehicle vector) with the topic and the vehicle. The process of metaphor production can also be modeled in the same way; an appropriate vehicle can be selected from the neighbors of the vector for the metaphorical meaning expressed by a given topic-vehicle pair, and a search space can be manipulated by varying the number of and quality of neighbor vectors. Indeed, Chiappe and Chiappe (2007) attempted to explain their findings on metaphor production using the predication algorithm.

These research topics for advancing the study of metaphor production are worth pursuing in further research.

### Acknowledgments

This study was supported by a Grant-in-Aid for Scientific Research B (No.23300098) from the Japan Society for the Promotion of Science.

### References

- Bowdle, B., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, 112(1), 193–216.
- Chiappe, D., & Chiappe, P. (2007). The role of working memory in metaphor production and comprehension. *Journal of Memory and Language*, 56, 172–188.
- Chiappe, D., & Kennedy, J. (1999). Aptness predicts preference for metaphors or similes, as well as recall bias. *Psychonomic Bulletin & Review*, 6, 668–676.
- Fainsilber, L., & Ortony, A. (1987). Metaphorical uses of language in the expression of emotions. *Metaphor and Symbolic Activity*, 2(4), 239–250.
- Gentner, D. (1982). Are scientific analogies metaphors? In D. Miall (Ed.), *Metaphor: Problems and perspectives* (pp. 106–132). Sussex, England: The Harvester Press.
- Gibbs, R. W. (Ed.). (2008). *The Cambridge handbook of metaphor and thought*. Cambridge University Press.
- Giora, R. (2003). *On our mind: Saliency, context, and figurative language*. Oxford University Press.
- Glucksberg, S. (2001). *Understanding figurative language: From metaphors to idioms*. Oxford University Press.
- Glucksberg, S., & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity. *Psychological Review*, 97, 3–18.
- Hussey, K. A., & Katz, A. N. (2006). Metaphor production in online conversation: Gender and friendship status. *Discourse Processes*, 42(1), 75–98.
- Jones, L., & Estes, Z. (2006). Roosters, robins, and alarm clocks: Aptness and conventionality in metaphor comprehension. *Journal of Memory and Language*, 55, 18–32.
- Katz, A. (1989). On choosing the vehicles of metaphors: Referential concreteness, semantic distances, and individual differences. *Journal of Memory and Language*, 28, 486–499.
- Katz, A., Paivio, A., & Marschark, M. (1985). Poetic comparisons: Psychological dimensions of metaphoric processing. *Journal of Psycholinguistic Research*, 14(4), 365–383.
- Kintsch, W. (2000). Metaphor comprehension: A computational theory. *Psychonomic Bulletin & Review*, 7(2), 257–266.
- Roberts, R., & Kreuz, R. (1994). Why do people use figurative language? *Psychological Science*, 5(3), 159–163.
- Rosch, E., & Mervis, C. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573–605.
- Steen, G. (1994). *Understanding metaphor in literature: An empirical approach*. Longman Publishing Group.
- Utsumi, A. (2005). The role of feature emergence in metaphor appreciation. *Metaphor and Symbol*, 20(3), 151–172.
- Utsumi, A. (2011). Computational exploration of metaphor comprehension processes using a semantic space model. *Cognitive Science*, 35(2), 251–296.
- Utsumi, A., & Sakamoto, M. (2007). Computational evidence for two-stage categorization as a process of adjective metaphor comprehension. In *Proceedings of the Second European Cognitive Science Conference (EuroCogSci2007)* (pp. 77–82).
- Utsumi, A., & Sakamoto, M. (2011). Indirect categorization as a process of predicative metaphor comprehension. *Metaphor and Symbol*, 26(4), 299–313.
- Williams-Whitney, D., Mio, J., & Whitney, P. (1992). Metaphor production in creative writing. *Journal of Psycholinguistic Research*, 21(6), 497–509.