

Number of topic-vehicle shared features influences the aptness of metaphors

Ryunosuke Oka^{1*}, Takashi Kusumi².

¹ Information Technology R&D Center, Mitsubishi Electric Corporation, Kamakura, Kanagawa 247-8501, Japan.

² Graduate School of Education, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan.

*Corresponding author at: Mitsubishi Electric Corporation, Kamakura, Kanagawa 247-8501, Japan.

E-mail address: Qualia1006@gmail.com (R. Oka)

Abstract

This study investigated whether the number of topic-vehicle shared features influences readers' evaluation of metaphor aptness. Experiment 1 compared the number of topic-vehicle shared features on metaphor aptness. A two topic-vehicle shared features condition showed a higher aptness rating than a one topic-vehicle shared feature condition. Experiments 2 and 3 compared the number of topic-vehicle shared features and unique features on metaphor aptness. The two topic-vehicle shared features condition showed a higher aptness rating than the one topic-vehicle shared feature and the one vehicle/topic unique feature and two unique features condition. There were differences in aptness between the metaphor and simile in a one topic-vehicle shared feature and a one topic/vehicle unique feature condition. Therefore, topic-vehicle shared features determine the aptness of metaphors, and the effect of topic/vehicle unique feature mixed with topic-vehicle shared features were different between metaphors and similes.

Keywords: metaphor; aptness; shared features

Number of topic-vehicle shared features influences the aptness of metaphors

This study aimed to examine whether the number of topic-vehicle shared features influences readers' evaluation of metaphor aptness. Metaphor (in particular, nominal metaphor) is when two unrelated concepts are juxtaposed, resulting in literally false assertions (Al-Azary & Katz, 2021). For example, "her smile is a flower" is a nominal metaphor. A metaphor comprises a topic and a vehicle. A topic is the concept explained in terms of the vehicle. The vehicle is the concept that explains the property of the topic. In "her smile is a flower", "her smile" is a topic, and "a flower" is a vehicle. Topic-vehicle shared features refer to the common attributes in a topic word (or concept) and a vehicle word (or concept). For example, "gorgeous" and "bright" are the topic-vehicle common features of "smile" and "flower" because these features were commonly listed when we asked participants to list the features of each word. The property attribution model (Glucksberg et al., 1997) is one of the models that treat metaphor aptness (i.e., how well the vehicle concept captures topic-significant features). Based on this model, some previous studies (Glucksberg & Haught, 2006; Haught, 2013; Oka & Kusumi, 2020) revealed that topic-vehicle shared features affect participants' aptness rating on a metaphor. However, it was still unclear whether the number of topic-vehicle shared features and topic/vehicle unique features affected participants' aptness rating on metaphors. For example, "nice" is a "her smile" (topic)-unique feature, and "bloom" is a "flower" (vehicle)-unique feature. This study will reveal what type of feature (i.e., topic-vehicle shared features) affects metaphor aptness and the effect of the number of the topic-vehicle shared/unique features has on metaphor aptness.

Property attribution model and metaphor aptness

When we use metaphors to describe an object or action, we must choose a vehicle with a meaning different from the topic, but that captures the topic's significant properties.

The property attribution model (Glucksberg et al., 1997) explains this idea well. The property attribution model treats metaphor as the sentence whose vehicle asserts an attributive category that provides properties that may be attributed to the topic. The point of the property attribution model in metaphor comprehension is the interaction between the topic and vehicle concept; that is, the vehicle provides the candidate properties attributed to the topic, and the topic constrains the candidate properties provided by the vehicle. For example, in the sentence “My lawyer is a shark,” the vehicle unique property “has a fin” (that is, vehicle unique feature) is suppressed, while “tenacious” is enhanced. Through this process, we perceive “My lawyer is a shark” as “my lawyer is tenacious” (Gernsbacher et al., 2001).

In view of the property attribution model, in understanding metaphors, it is important to understand how well the category is exemplified by the vehicle capturing the salient properties of the topic. This extent is called the aptness of metaphor (Chiappe & Kennedy, 1999; Chiappe, Kennedy, & Chiappe, 2003; Jones & Estes, 2005, 2006). Aptness reflects how salient the vehicle properties are and how well the vehicle semantic dimension can apply to the topic. For example, the vehicle “an alarm clock” has the salient property of “noisy”; thus, this property can easily apply (i.e., is apt) to the topic “a rooster” but is difficult to apply (i.e., less apt) to the topic, “a robin” (Jones & Estes, 2006).

Research comparing the candidate factors that affect metaphor processing showed aptness is a good predictor of processing fluency of metaphors. Chiappe et al. (2003) compared the comprehensibility (i.e., how easy it is to understand a comparison between the topic and vehicle concept) and aptness of how well people correctly memorize a metaphor sentence as a metaphor and a simile sentence as a simile. In a simile, two unrelated concepts are juxtaposed, resulting in laterally false assertion with hedges, which implies comparing two things (e.g., like, as). For example, “her smile is like a flower” is a nominal simile. Results showed that each factor explains the accuracy of performance, but the aptness has a unique

variance. In addition, Jones and Estes (2006) independently manipulated the conventionality (the strength of association between a metaphor vehicle and its figurative meaning) and aptness. They found that only aptness predicts the preference for a metaphor over a simile, the ease of a metaphor comprehension rating, and the extent to which a topic concept is included in the vehicle concept. An especially important point in Jones and Estes (2006) is that the aptness explains the category membership of the vehicle concept. Thus, aptness is a good predictor of metaphor as a categorical assertion sentence, as explained by the property attribution model.

The type and number of topic-attributed features on metaphor aptness

Previous studies of the property attribution model showed that the type of topic-attributed features affects the metaphor aptness. Glucksberg and Haught (2006) and Haught (2013) compared the effect of four types of topic-attributed features on metaphor and simile aptness: the topic unique feature (the feature whose meaning is only applicable to the topic; e.g., “My lawyer is a well-paid shark”), the vehicle unique feature (the feature whose meaning is only applicable to the vehicle; “My lawyer is a razor-toothed shark”), the topic-vehicle shared feature (the feature whose meaning applies to both the topic and vehicle; e.g., “My lawyer is an old shark”), and no feature (e.g., “My lawyer is a shark”). In this study, aptness was determined by how good each metaphor or simile was (Haught, 2013; p. 266). They reported two interesting results. First, aptness ratings on metaphors (metaphor and simile) with the topic-vehicle shared feature condition and the no feature condition showed no significant difference. This result suggested that although the metaphor modified by the topic and vehicle applicable features is novel to the participants, both forms of metaphors might be processed as a categorical sentence. Second, aptness ratings on metaphor and simile showed a significant difference in the topic-applicable adjectival modifier condition. This result suggested that the metaphor form, which treats the sentence as a category assertive sentence,

was more apt than the simile form, which treats the sentence as a feature alignment in the topic-vehicle shared feature condition. These results supported the property attribution model: even the sentence is novel, metaphors and similes are equally good (that is, both preferred categorical processing).

In addition, Oka and Kusumi (2020) showed that the subtype of a topic-vehicle shared feature also affects the metaphor aptness. In their study, aptness was determined as the extent to which the vehicle captured the important feature of the topic, which is the same as the definition in Chiappe et al. (2003) and Jones and Estes (2006). In their study, participants were asked to rate the aptness of a metaphor with three types of topic-attributed features: the distinctive features of the target metaphor (the feature whose meaning applies to both the topic and the vehicle in the target metaphor, and the feature is only applicable to the target metaphor; e.g., “Her smile is gorgeous and blooming”, for the target metaphor “Her smile is a flower”), the distinctive feature of the competitive metaphor (the feature whose meaning is applicable to both the topic and the vehicle in competitive metaphor, and the feature is only applicable to the competitive metaphor; e.g., “Her smile is shining and hot”, for the competitive metaphor “Her smile is a sun”), and the common feature (the feature whose meaning is applicable to both the topic and the vehicle in the target and competitive metaphor; e.g., “Her smile is bright and pleasant”). Both the distinctive features of the target metaphor and the common features are topic-vehicle shared features. In this study, all three conditions have two attributed features. Results showed that the aptness rating on the distinctive features of the target metaphor was higher than that of the common features, which suggested that a distinctive (not shared by another vehicle) feature has more effect on metaphor aptness.

Previous studies, as explained above, showed that the topic-vehicle shared features and its variant affects the metaphor aptness. However, it was still unclear how the number of

the topic-attributed features affects the metaphor and simile aptness. All previous studies explained above fix the number of topic-attributed features: one in Glucksberg and Haught (2006) and Haught (2013), and two in Oka and Kusumi (2020). As the definition of aptness adopted in Jones and Estes (2006) and Oka and Kusumi (2020), the aptness reflects how well the vehicle concept captures the salient properties of the topic. It implies that not only the quality of the feature attributed to the topic (as examined in the previous studies) but also the quantity (i.e., how many features are attributed to the topic) of the feature attributed to the topic affect the metaphor aptness.

In addition, the interaction between the type (metaphor or simile) and the number of features attributed to the topic of metaphor aptness was also unexplored. As Chiappe and Kennedy (2001) explained, because a metaphor is a category assertive sentence, a metaphor vehicle needs to entail all the features attributed to the topic. In contrast, because a simile is a similarity assertive sentence, a simile vehicle needs not to entail all but some features attributed to the topic. Inspired by these ideas, we predicted that whether the topic-attributed feature is a topic-vehicle shared feature or not, the effect of the number of topic-attributed features on aptness is similar if there is at least one topic-vehicle shared feature in a simile. However, we predicted that this pattern would not be observed in a metaphor.

The Present Study

We examined whether the number of shared topic-vehicle shared features affects the aptness of a metaphor. We hypothesized that the number of topic-vehicle shared features affects the metaphor aptness (Hypothesis 1), and there is an interaction between the type of topic attributed feature (whether the topic attributed feature is a vehicle feature or a topic feature) and metaphor type (Hypothesis 2).

We conducted three experiments. Participants in each experiment had to rate the aptness of a metaphor/simile in each topic-attributed condition. In each experiment, we

manipulated the number of topic-vehicle shared features (Experiment 1) and the combination of the topic-vehicle shared features and topic/vehicle unique features (Experiments 2 and 3). Previous studies asked participants to rate the metaphor whose topic or vehicle was modified by the feature (e.g., Glucksberg & Haught, 2006; Haught, 2013; Kusumi, 1985). However, this study separately presented a literal sentence whose topic is modified by the topic-attributed features and a metaphorical sentence. We asked participants to rate how well the vehicle captured the important features of the topic in the literal sense. For example, participants are presented with the sentence “her sarcasm is painful” and asked to rate how “her sarcasm is a dagger” captures the important property of the presented sentence (a more detailed procedure is described in the methods section). This is because we wanted to know how the aptness of a given metaphor changed per the type and the number of topic-attributed features.

As explained above, there were two definitions for metaphor aptness – “how good each metaphor (Glucksberg & Haught, 2013)” and “how well the category is exemplified by the vehicle capturing the salient properties of the topic (Jones & Estes, 2006)”. This study adopts the latter definition of metaphor aptness for two reasons. On the one hand, Jones and Estes’s (2006) definition is better for explaining the vehicle’s role in metaphor comprehension. On the other hand, Glucksberg and Haught’s (2013) definition is confusing because some studies called it “metaphor goodness”. For example, Katz, Pavio, Marschark, and Clark (1988) collected the “metaphor goodness” with the definition “how good or apt metaphors appear (p.194)”. Similarly, Marschark, Katz, and Pavio (1983) collected the metaphor goodness with the definition “metaphors described a situation in an apt and pleasing way (p.21)”.

In Experiment 1, we examined whether the number of topic-vehicle shared features affected the reader’s evaluation of metaphor aptness. In this experiment, we compared one

topic-vehicle shared feature condition and two topic-vehicle shared features conditions. If the number of topic-vehicle shared features affects the metaphor aptness, the aptness rating of metaphor shows higher in a two-features condition than in a one-feature condition.

Experiments 2 and 3 examined whether the type and the number of topic-vehicle shared features affect the reader's evaluation of metaphor and simile aptness. These experiments set three conditions: a two topic-vehicle shared features condition, a one topic-vehicle shared feature and a one topic/vehicle unique features condition, and a two topic/vehicle unique features condition. Experiment 2 explored the effect of the vehicle unique feature, and Experiment 3 explored the topic feature. If the vehicle of a metaphor requires to explain all the features attributed to the topic but not the vehicle of a simile as described by the Chiappe and Kennedy (2001), the aptness rating of one topic-vehicle shared feature and one topic/vehicle unique features condition showed a higher aptness rating in a simile condition than in a metaphor condition. We also investigated whether the number of topic-vehicle shared features has more effect on the aptness of a metaphor/simile than those of the topic/vehicle unique features, as indicated by Glucksberg and Haught (2006) and Haught (2013). We also anticipate that the effect of the topic unique feature attributed to the topic has less impact on metaphor/simile aptness compared with the vehicle unique feature attributed to the topic. In this study, we defined aptness as how well the vehicle captures an important feature of the topic, as in Jones and Estes (2006). Compared to the metaphor goodness definition adopted in Glucksberg and Haught (2006) and Haught (2013), definition adopted in our research pay more attention to the vehicle function, which supposed that the vehicle provides the candidate features attributed to the topic (Glucksberg et al., 1997). This will make participants focus more on the vehicle rather than the topic feature. Because the scope of what topic entails become wider and the vehicle features were evaluated as the topic-vehicle shared features, participants may show higher aptness rating on simile than metaphor

in vehicle-unique features condition.

This study differs from Oka and Kusumi (2021) in terms of aim and method. The current study focuses on the type of topic attributed features (especially topic-vehicle shared feature and topic/vehicle unique feature) and its effect on the aptness of a metaphor and simile. We asked participants to answer the aptness of the presented metaphor and simile in isolation (i.e., participants were presented with only metaphor/simile sentences, without any other options). In contrast, Oka and Kusumi (2021) focus on the type of topic attributed features (especially the number of topic-vehicle shared and topic unique features) and their effect on the preference for metaphorical expressions over literal expressions. They asked participants to choose the options that best paraphrased the given sentence (i.e., participants were presented with metaphor/simile, literal, and other options). Therefore, though the metaphor preference and the aptness of metaphors are positively correlated (e.g., Oka & Kusumi, 2020), our study and Oka and Kusumi (2021) differ significantly.

The contribution of this research is three folds. First, this study examines the number of topic-vehicle shared features on metaphor aptness. Though previous studies did not control the number of topic attributed features, because metaphor aptness includes both the quality of the topic attributed feature and the quantity of the topic attributed features, investigating the number of topic-vehicle shared features on metaphor aptness is also important. This study emphasizes the importance of the quantity of topic attributed features on metaphor aptness. Second, this study examines the features highlighted in metaphor and simile comprehension. Glucksberg and Haught (2006) examined the difference between metaphor and simile on metaphor goodness. In contrast, this study defines aptness as how well the vehicle captures the important property of the topic and topic-attributed features. This expands our understanding of the ways in which metaphor aptness was determined by the number and the type of topic attributed features. Third, as a psychological study of similarity and metaphor,

this study empirically examines the effect of topic-vehicle shared features and unique features on metaphor aptness. As Tversky (1977) and Ortony (1979) pointed out, the number of shared features and unique features is important in the similarity judgment of two concepts. This study examines both features and their combination of metaphor aptness in a controlled setting.

The datasets during and/or analyzed during the current studies are available from https://github.com/okaexp/shared_features_and_metaphor_aptness/tree/master.

Experiment 1

Method

Participants. We recruited 60 participants (39 males and 21 females) between the ages of 23 and 64 years ($M = 40.9$, $SD = 7.7$) using Crowdtasks, a crowdsourcing service in Japan. All the participants were recruited anonymously, and each participant was compensated with 50 yen (that is, almost equal to 50 cents). In this experiment, the sample criteria were defined to include only native speakers of Japanese.

Design. Experiment 1 adopted a 2 (number of topic-vehicle shared features: one, two) $\times$ 2 (lists: lists A, B) experimental design. The number of topic-vehicle shared features served as a within-participants and -items variable; however, the lists served as a between-participants variable.

Stimuli. We used two types of stimuli: metaphors and topic-vehicle shared features. Table 1 presents examples of the stimuli that were used in Experiment 1.

Table 1. Examples of stimuli that were used in Experiment 1. English translations of the original Japanese text are provided.

Metaphor	Number of	Example
----------	-----------	---------

Topic-Vehicle		
Shared Features		
Her sarcasm is a dagger	One	Her sarcasm is painful.
(彼女の皮肉は短刀だ)		(彼女の皮肉は痛い)
	Two	Her sarcasm is painful and harmful.
		(彼女の皮肉は痛く、人を傷つける)
The pearl is a drop of water	One	This pearl is beautiful.
(この真珠は水滴だ)		(この真珠は美しい)
	Two	This pearl is beautiful and round.
		(この真珠は美しく、丸い)

1

2 To collect metaphor stimuli, we identified 36 Japanese metaphors that have been

3 anthologized by past researchers (Nakamoto & Kusumi, 2004). These metaphors have been

4 used in past studies on metaphor (or simile) comprehension among Japanese participants

5 (Kusumi, 1985; Nakamoto, 2003).

6 We collected the topic-vehicle shared features in two steps. First, we conducted three

7 pilot questionnaires: a listing metaphor interpretations questionnaire, a listing topic features

8 questionnaire, and a listing vehicle features questionnaire. In the listing metaphor

9 interpretations questionnaire, participants ($N = 51$) were shown metaphors (e.g., “Her smile is

10 a flower”) and asked to answer at least two and at most three interpretations (e.g., “gorgeous”)

11 to the given metaphor. Participants were asked to limit their interpretations to verbs, adjectives,

12 and adjectival verbs. For example, a participant answered “poking and prodding” and “sharp”

13 for “her sarcasm is a dagger”. In the listing topic/vehicle features questionnaires, participants

(topic: $N = 51$, vehicle: $N = 50$) were shown topics/vehicles of metaphors used in the metaphor interpretations questionnaire (e.g., topic: smile, vehicle: flower). For example, a participant answered “dirty”, “scary”, and “awful” to “sarcasm”, and another participant answered “short” and “cut” to “a dagger”. Participants were asked to answer their interpretations following the same instructions given in the listing metaphor interpretations questionnaire.

Second, the first author classified the collected features in the three questionnaires into topic-vehicle shared features, topic unique features, and vehicle unique features. This process is needed to account for the feature overlap: even if each questionnaire was designed to collect a specific type of feature (e.g., topic features questionnaires were designed to collect topic-unique feature), sometimes the features obtained were overlapped between questionnaires (e.g., “short” is the property obtained in both the vehicle condition for “a dagger” and the metaphor condition for “her sarcasm is a dagger”). Topic-vehicle shared features were defined as the features mainly obtained in listing the metaphor interpretations questionnaire. For example, in “Her sarcasm is a dagger,” there were 11 responses to “painful” in the metaphor questionnaire but only three responses in the vehicle questionnaire and zero responses in the topic questionnaire. Therefore, we decided to use “painful” as a topic-vehicle shared feature in “Her sarcasm is a dagger.” Similar to this criteria, topic unique features and vehicle unique features were defined as the features mainly obtained in topic/vehicle questionnaires. All the stimuli used in this study (including metaphors, features, and the number of features obtained in three questionnaires) were summarized in the Appendix (Table S1). Thus, topic-vehicle shared features derives from the metaphor, and these were reflective of metaphorical category in nature.

Procedure. We collected the data using the Qualtrics platform (Qualtrics, 2019). Participants could respond to this survey using a personal computer. As a first step, the participants read information about this study and provided informed consent. Next, the

participants read detailed task instructions; then, they were required to read a short sentence that described a topic with one or two topic-vehicle shared feature(s). For example, “Her sarcasm is painful” was presented in a one-feature condition, and “Her sarcasm is painful and harmful” was presented in a two-features condition. Participants were required to rate the aptness of the metaphor (e.g., “Her sarcasm is a dagger”) that was presented beneath the sentence. The aptness of the metaphor was defined as the extent to which the vehicle captured the important feature of the topic. Participants were asked to indicate their perceived metaphor aptness on a 7-point scale (0 = “does not capture the important properties of the topic”; 6 = “captures the important properties of the topic extremely well”). When participants had finished the task, we presented them with a payment code needed to receive credit and concluded the experiment.

Results

Figure 1 shows the mean aptness ratings for the metaphors in each condition in Experiment 1. As predicted, the mean aptness rating was higher in the two-features condition than in the one-feature condition. To further examine this difference, we used a mixed-effects regression model that was executed using the lmer program of the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in the R (version 3.3.2) environment for statistical computing (R Development Core Team, 2016). We entered the number of topic-vehicle shared features as a fixed factor. We coded two shared features as 0.5 and one shared feature as -0.5 to follow the deviation coding scheme. In addition, we specified the random slope of the type of topic-vehicle shared features and the intercepts for participants and items as the random factors. The comparisons between the two features and one feature yielded significant results ($\beta = 0.45$, $SE = 0.12$, $t = 3.87$, $p < .001$). These results supported the hypothesis that two features ($M = 3.67$; e.g., “Her sarcasm is painful and harmful”) will evidence higher aptness ratings than one feature ($M = 3.21$; “Her sarcasm is painful”).

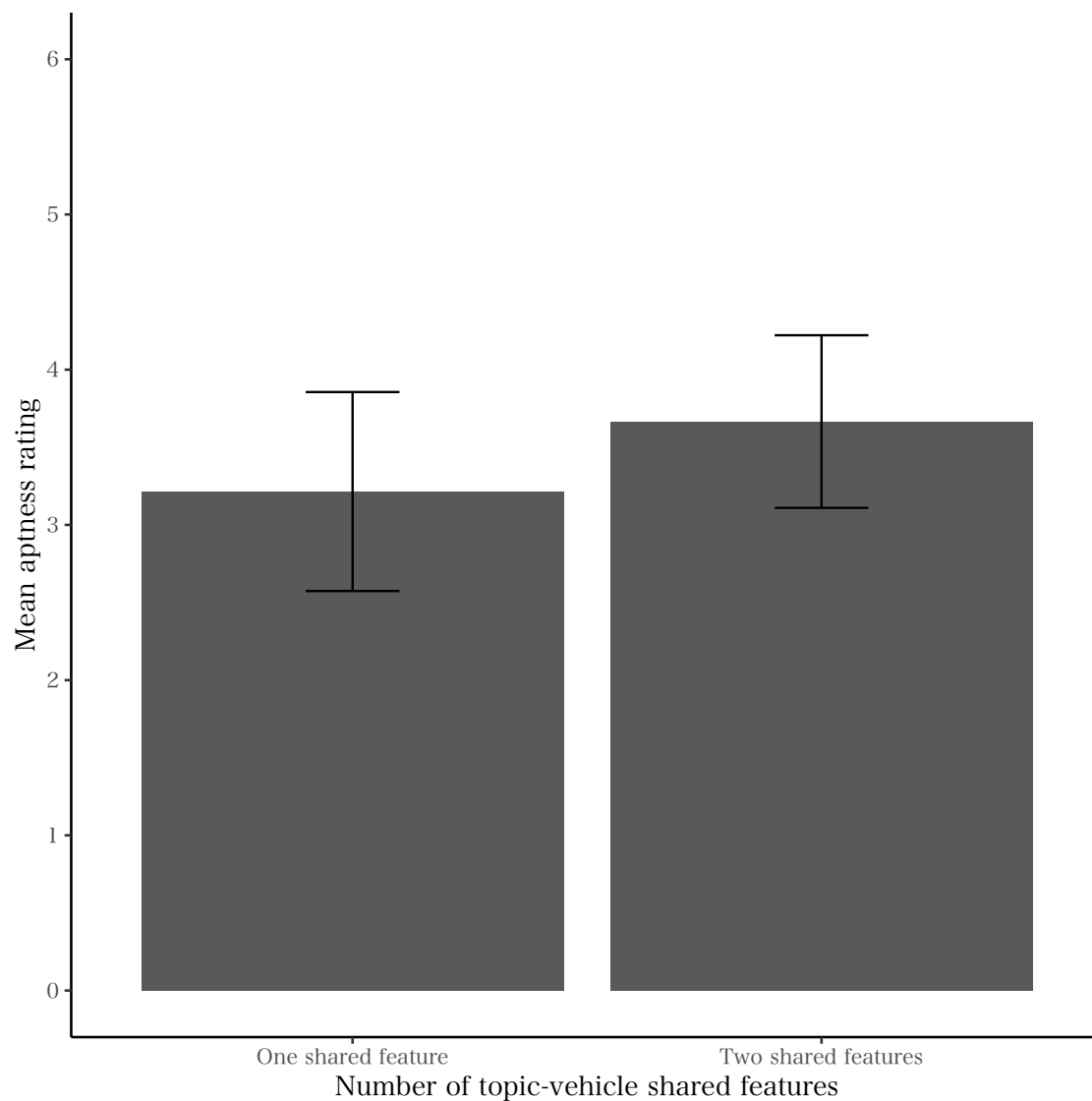


Figure 1. Mean aptness ratings for the metaphor across the two conditions that differed in the number of topic-vehicle shared features (Experiment 1). Error bars represent one standard deviation of the mean.

Discussion

The results of Experiment 1 showed that the main effect of the number of topic-vehicle shared features on a reader's evaluation of metaphor aptness was significant. This suggests that metaphor aptness is influenced by the number of topic-vehicle shared features as predicted by the property attribution model (Glucksberg et al., 1997).

In Experiment 2, we examined whether the type and the number of topic-vehicle shared

features influence the reader's evaluation of metaphor and simile aptness. In this experiment, we compared two topic-vehicle shared features, one topic-vehicle shared feature, one vehicle unique feature, and two vehicle unique features. If the number of topic-vehicle shared features have an effect on a reader's evaluation on metaphor aptness, a reader's evaluation of metaphor aptness should be in accordance with the following descending order: two topic-vehicle shared features, one topic-vehicle shared feature and one vehicle unique feature, and two vehicle unique features.

In Experiment 2, we compared metaphor and simile of a reader's evaluation of metaphor aptness. If a simile does not need to include all the features attributed to the topic (i.e., Chiappe & Kennedy, 2001), the one topic-vehicle shared feature can similarly influence metaphor aptness as the two topic-vehicle shared features.

Experiment 2

Method

Participants. We recruited 168 participants (86 males and 82 females) between the ages of 20 and 59 years ($M = 40.3$, $SD = 8.8$) using Crowdworks. All the participants were recruited anonymously, and each participant was compensated with 50 yen. In this experiment, the sample criteria were defined to include only native speakers of Japanese and those who had not participated in Experiment 1.

Design. Experiment 2 used a 3 (type of topic-attributed features: two topic-vehicle shared features, one topic-vehicle shared feature and one vehicle unique feature, and two vehicle unique features) $\times$ 2 (metaphor type: metaphor and simile) $\times$ 12 (list: lists A, B, C, D, E, F, G, H, I, J, K, and L) design. The type of topic-attributed features served as a within-participants and -items variable; however, the list and the metaphor type were between-participants variables.

Stimuli. Similar to Experiment 1, we used two types of stimuli: metaphors and topic-

attributed features. In Experiment 2, however, we changed three points.

First, we changed the type of topic-attributed features into two topic-vehicle shared features, one topic-vehicle shared feature and one vehicle unique feature, and two vehicle unique features. We used vehicle unique features obtained from the listing vehicle features questionnaire in Experiment 1.

Second, we used both metaphors and similes as stimuli. For example, “Her sarcasm is a dagger” was presented in a metaphor condition, and “Her sarcasm is like a dagger” was presented in a simile condition.

Third, we prepared a counterbalanced list in one topic-vehicle shared feature and a one-vehicle unique feature condition. Because we prepared two types of vehicle unique features in this experiment, we prepared two types for this condition. For example, for the metaphor “Her sarcasm is (like) a dagger,” we prepared the sentence “Her sarcasm is painful and cut” and “Her sarcasm is harmful and short” as a one topic-vehicle shared feature and a one vehicle unique feature condition. Participants were only shown one of these sentences.

Table 2 presents examples of the stimuli that were used in Experiment 2.

1 Table 2. Examples of stimuli that were used in Experiment 2. English translations of the
2 original Japanese text are provided.

Metaphor	Type of Topic-Vehicle Shared Features	Example
Her sarcasm is a dagger (彼女の皮肉は短刀だ)	Two topic- vehicles shared	Her sarcasm is painful and harmful. (彼女の皮肉は痛く、人を傷つける)
	One topic- vehicle shared and one vehicle unique	Her sarcasm is painful and cuts people. (彼女の皮肉は痛く、人を切る)
	Two vehicles unique	Her sarcasm cuts people and is short. (彼女の皮肉は人を切り、短い)
The pearl is a drop of water (この真珠は水滴だ)	Two topic- vehicles shared	This pearl is beautiful and round. (この真珠は美しく、丸い)
	One topic- vehicle shared and one vehicle unique	This pearl is beautiful and cold. (この真珠は美しく、冷たい)
	Two vehicles unique	This pearl is cold and falls. (この真珠は冷たく、落ちる)

3

4

Procedure. The procedure was identical to that of Experiment 1.

Results

Due to the stimuli mistakes, we removed “His love is a season” from the following analysis. Figure 2 shows the mean aptness ratings for the metaphors in each condition in Experiment 2. There were two important results. First, the two topic-vehicle shared features (3.88) showed a higher aptness rating compared to the one topic-vehicle shared feature and one vehicle unique feature (3.58), and the two vehicle unique features (3.32). To clarify these patterns, we again used a mixed-effects regression model. We entered the number of types of topic-attributed features and the metaphor type as fixed factors. To test the effect of topic-vehicle shared features on metaphor aptness, we put two factors: one coded comparison of “two topic-vehicle shared features vs. grand mean (i.e., two topic-vehicle shared features as 0.5; two vehicle unique features as -0.5)”, and the other coded comparison of “one topic-vehicle shared feature and one vehicle unique feature vs. grand mean (one topic-vehicle shared feature and one vehicle unique feature as 0.5; two vehicle unique features as -0.5)”, to follow deviation coding scheme. Here, the grand mean corresponds to the overall mean of the dependent variable. Each number (i.e., 0.5) refers to the dummy number to follow the deviation coding scheme. Because the independent variable has three levels (i.e., two topic-vehicle shared features; one topic-vehicle shared feature and one topic/vehicle feature; and two topic/vehicle features), we need two fixed factors to differently examine the effect of two topic-vehicle shared features, and the one topic-vehicle shared feature and one topic/vehicle feature. To test the effect of the metaphor type on metaphor aptness, we coded metaphor as 0.5 and simile as -0.5. In addition, we specified the random slope of the type of topic-attributed features, type of metaphor, and the intercepts for participants and items as the random factors. The categorical coding of the topic-vehicle shared features and the metaphor type were needed to interpret the regression coefficients of interaction coefficients such as ANOVA (as a reference, see

<http://alexandrecremers.com/docs/UsefulTipsForR.pdf>). Results showed two topic-vehicle shared features vs. grand mean ($\beta = 0.58$, $SE = 0.12$, $t = 4.93$, $p < .001$) were significant. Second, there was a significant effect of the interaction of metaphor type and one topic-vehicle shared feature and one vehicle unique feature vs. grand mean ($\beta = -0.21$, $SE = 0.10$, $t = -2.11$, $p < .05$).

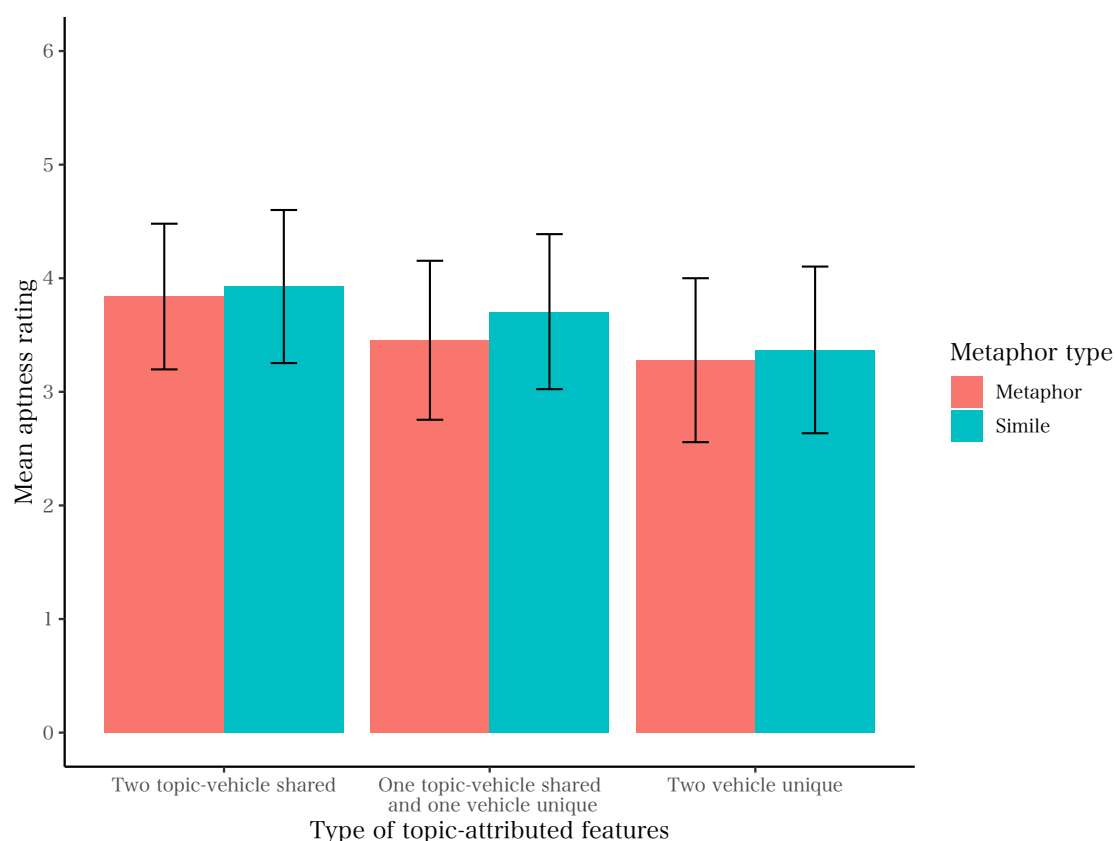


Figure 2. Mean aptness ratings for the metaphor across the three conditions that differed in the type of topic-attributed features (Experiment 2). Error bars represent one standard deviation of the mean.

Discussion

The results of Experiment 2 showed the effect of the type of topic-attributed features on a reader's evaluation of metaphor aptness when there were two topic-vehicle shared features. This finding suggests topic-vehicle shared features affect a reader's evaluation of metaphor aptness than when the vehicle unique features were attributed to the topic. In

addition, there is a difference between a metaphor and a simile in a reader's evaluation on metaphor aptness when the topic-vehicle shared feature and vehicle unique feature were mixed. This suggests that although a metaphor needs to emphasize all the features attributed to the topic, a simile does not need to emphasize all the features, only some features that are derived from the topic-vehicle shared features.

In Experiment 3, we examined the effect of topic unique features. In this experiment, we compared two topic-vehicle shared features, one topic-vehicle shared feature and one topic unique feature, and two topic unique features. Compared to Experiment 2, we anticipated no effect of metaphor type in one topic-vehicle shared feature and one topic unique feature condition, as implied by Chiappe & Kennedy (2001).

Experiment 3

Method

Participants. We recruited 168 participants (91 males and 77 females) between the ages of 20 and 65 years ($M = 38.6$, $SD = 9.5$) using Crowdworks. All the participants were recruited anonymously, and each participant was compensated with 50 yen. In this experiment, the sample criteria were defined to include only native speakers of Japanese and participants who had not participated in Experiments 1 or 2.

Design. Experiment 3 used a 3 (type of topic-attributed features: two topic-vehicle shared features, one topic-vehicle shared feature and one topic unique feature, and two topic unique features) $\times$ 2 (metaphor type: metaphor and simile) $\times$ 12 (list: lists A, B, C, D, E, F, G, H, I, J, K, and L) design. The type of topic-attributed features served as a within-participants and -items variable; however, the list and the metaphor type were the between-participants variables.

Stimuli. Similar to Experiments 1 and 2, we used two types of stimuli: metaphors and topic-attributed features. Compared to Experiment 2, in Experiment 3, we changed two points.

First, we changed the type of topic-attributed features to two topic-vehicle shared features, one topic-vehicle shared feature and one topic unique feature, and two topic unique features. We used topic unique features obtained from the listing topic features questionnaire in Experiment 1.

Second, we prepared a counterbalanced list in the one topic-vehicle shared feature and the one topic unique feature condition. As in Experiment 2, we prepared two types of topic unique features in this experiment; hence, we prepared two types for this condition. For example, for the metaphor “Her sarcasm is (like) a dagger,” we prepared the sentence “Her sarcasm is painful and sarcastic” and “Her sarcasm is harmful and says something” as the one topic-vehicle shared feature and the one topic unique feature condition. Participants were only shown one of these sentences.

Table 3 presents examples of the stimuli that were used in Experiment 3.

Table 3. Examples of stimuli that were used in Experiment 3. English translation of the with original Japanese text are provided.

Metaphor	Type of		Example
	Topic-Vehicle		
	Shared Features		
Her sarcasm is a dagger (彼女の皮肉は短刀だ)	Two	topic-vehicles shared	Her sarcasm is painful and harmful. (彼女の皮肉は痛く、人を傷つける)
	One	topic-vehicle shared	Her sarcasm is harmful and says something.
	and one	topic unique	(彼女の皮肉は人を傷つけ、言う)
	Two	topics	Her sarcasm is sarcastic and says

	unique		something.
			(彼女の皮肉は嫌味で、言う)
The pearl is a drop of water	Two vehicles shared	topic-	This pearl is beautiful and round. (この真珠は美しく、丸い)
(この真珠は水滴だ)			
	One vehicle	topic- shared	This pearl is beautiful and white. (この真珠は美しく、白い)
	and one vehicle		
	unique		
	Two vehicles		This pearl is white and expensive. (この真珠は白く、高価だ)
	unique		

Procedure. The procedure was identical to that of Experiments 1 and 2.

Results

Figure 3 shows the mean aptness ratings for the metaphors in each condition in Experiment 3. There were two important results. First, the two topic-vehicle shared features (3.96) showed a higher aptness rating compared to the one topic-vehicle shared feature and one topic unique feature (3.24), and two topic unique features (2.28). To clarify these patterns, we again used a mixed-effects regression model. We entered the number of types of topic-attributed features and the metaphor type as fixed factors. Regarding the type of topic-vehicle shared features, we put two factors: one coded comparison of two topic-vehicle shared features vs. grand mean (i.e., two topic-vehicle shared features as 0.5 and two topic unique features as -0.5), and the other coded comparison of one topic-vehicle shared feature and one vehicle unique feature vs. grand mean (one topic-vehicle shared feature and one topic unique feature as 0.5 and two topic unique features as -0.5), to follow deviation coding scheme.

Regarding the metaphor type, we coded metaphor as 0.5 and simile as -0.5. In addition, we specified the random slope of the type of topic-attributed features, type of metaphor, and the intercepts for participants and items as the random factors. Results showed two topic-vehicle shared features vs. grand mean ($\beta = 1.60$, $SE = 0.15$, $t = 10.53$, $p < .001$) and one topic-vehicle shared feature and one vehicle unique feature vs. grand mean ($\beta = 0.15$, $SE = 0.06$, $t = 2.75$, $p < .01$) were significant. Second, there was no significant effect of the interaction of metaphor type and one topic-vehicle shared feature and one topic unique feature vs. grand mean ($\beta = 0.10$, $SE = 0.09$, $t = 1.14$, $p = .26$).

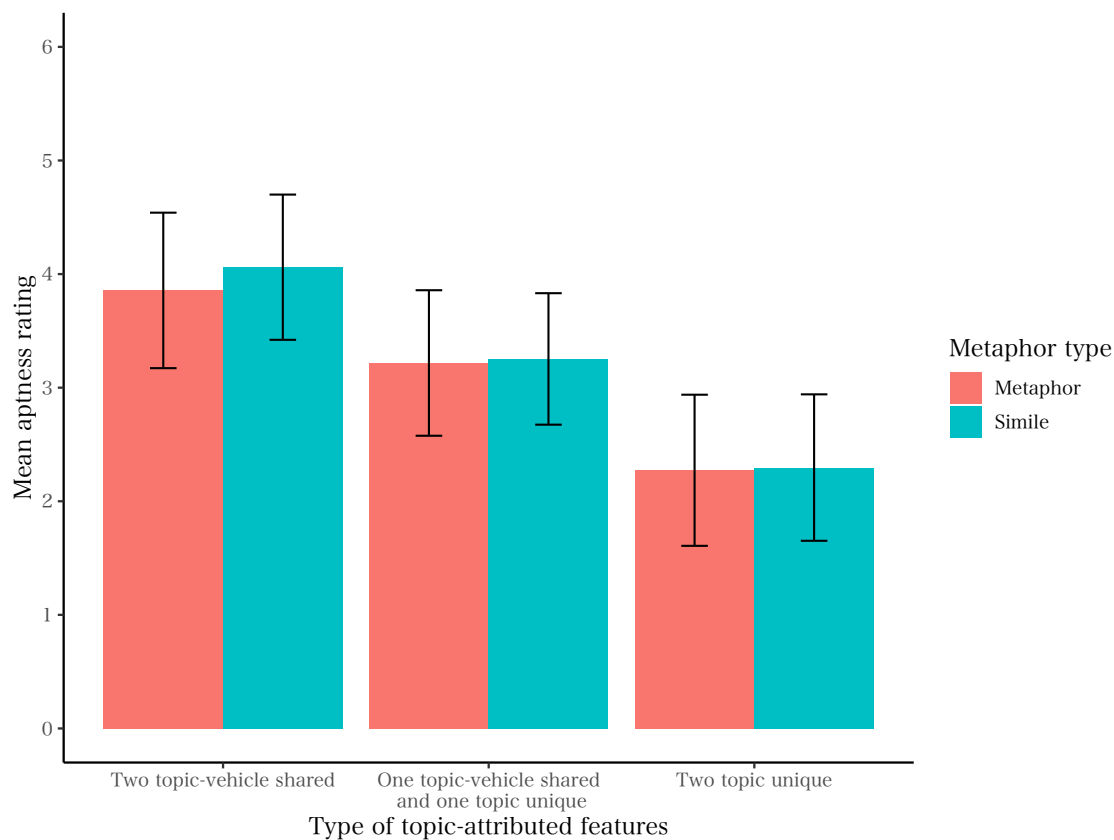


Figure 3. Mean aptness ratings for the metaphor across the three conditions that differed in the type of topic-attributed features (Experiment 3). Error bars represent one standard deviation of the mean.

Discussion

The results of Experiment 3 showed the main effect of the type of topic-attributed features on a reader's evaluation of metaphor aptness. This finding suggests that topic-vehicle shared features affect a reader's evaluation of metaphor aptness than topic unique features. In addition, there was no difference between metaphor and simile in a reader's evaluation of metaphor aptness when the topic-vehicle shared feature and topic unique feature was mixed. Combining the results obtained in Experiment 2 and this experiment, the effect of the topic and the vehicle unique feature was not the same when the topic-vehicle shared features were mixed.

General Discussion

In this study, we examined whether the number of topic-vehicle shared features affects the aptness of metaphor or not. Experiment 1 showed that the two topic-vehicle shared features were more apt than the one topic-vehicle shared feature. This supported our Hypothesis 1 in which the number of topic-vehicle shared features affects the metaphor aptness. The results of Experiment 2 showed that the two topic-vehicle shared features were more apt than the one topic-vehicle shared feature and one vehicle unique feature, and two vehicle unique features. In addition, there is a significant difference between a metaphor and a simile in a reader's evaluation of metaphor aptness when the topic-vehicle shared feature and vehicle unique feature were mixed. The results of Experiment 3 showed that the two topic-vehicle shared features were more apt than the one topic-vehicle shared feature and one topic unique feature, and two vehicle unique features.

Interestingly, there was no significant difference between a metaphor and a simile in a reader's evaluation of metaphor aptness when the topic-vehicle shared feature and topic unique feature were mixed. This supported Hypothesis 2, in which there is an interaction between the type and number of topic-attributed features. These results suggest that the topic-vehicle shared feature determines the metaphor aptness, and the effect of the topic/vehicle

unique feature on the metaphor and the simile was different when the mixed feature was attributed to the topic.

The results obtained in three experiments suggested that the topic-vehicle shared features are important in metaphor aptness. All three experiments showed that the more topic-vehicle shared features attributed to the topic, the more the metaphor/simile became apt. These results were consistent with the results obtained in Glucksberg and Haught (2006) and Haught (2013)—the topic-vehicle shared feature showed a higher aptness rating than the topic unique feature and vehicle unique feature. Though previous studies manipulated features with one feature, our experiments manipulated the number of topic-vehicle shared features (Experiment 1) and the combination of the topic-vehicle shared feature and topic/vehicle unique feature (Experiment 2, Experiment 3). Our results not only conceptually replicated the results of Glucksberg and Haught (2006) and Haught (2013) but also clarified that the topic-vehicle shared feature has more impact than the topic/vehicle unique feature were mixed on the reader's evolution of metaphor/simile aptness.

The results of Experiments 2 and 3 implied that the features highlighted in a simile and a metaphor are different. Our results showed that though the aptness of a simile was higher than that of a metaphor in the one topic-vehicle shared feature and one vehicle unique feature, this pattern was not observed in the one topic-vehicle shared feature and one topic unique feature condition. As suggested by Chiappe and Kennedy (2001), compared to the metaphor, which requires a vehicle to have all the features attributed to the topic, the simile does not require all the features attributed to the topic, only some. The difference of aptness between the metaphor and simile in a one topic-vehicle shared feature, and a one topic/vehicle unique feature condition supported Chiappe and Kennedy's (2001) suggestion.

The pattern of the one topic-vehicle shared feature, the one topic unique feature condition, and the two topic unique features condition was not consistent with the results

obtained in Glucksberg and Haught (2006) and Haught (2013). In Haught's studies, there were differences in aptness between a metaphor and a simile on the topic unique feature condition. One reason that causes the discrepancy between our results and Glucksberg and Haught (2006) and Haught (2013) results is the difference in the definition of aptness. Chiappe et al. (2003), Jones and Estes (2006), and ours defined aptness as the extent to which the vehicle captured the important features of the topic. In contrast, Glucksberg and Haught (2006) and Haught (2013) defined aptness as how good each metaphor or simile is. Differences in definition between Glucksberg and Haught (2006) and Haught (2013) and ours can explain other results obtained in this study. Our results showed that the two vehicle unique features condition (3.33) showed a higher average aptness rating than the two topic unique features condition (2.29) in both a metaphor and a simile. Therefore, our definition can be interpreted as making our participants focus on the vehicle features.

Our results strengthen the relationship between metaphor/simile aptness and the property attribution model. The most straightforward result we obtained is that the more topic-vehicle shared features, the more it becomes apt (Experiment 1). As explained in the introduction, aptness reflects two aspects of the property attribution model: the quality (how strong the association of each feature to the topic) and the quantity (how many features can be attributed to the topic). Many previous studies focused mainly on the quality of the feature attributed to the topic (e.g., Glucksberg & Haught, 2006; Haught, 2013). In contrast, Experiment 1 manipulated the quantity of the feature attributed to the topic and confirmed that the quantity also affects the metaphor aptness. Because metaphors are considered a category assertive sentence, and the aptness reflects how well the vehicle category applies to the topic and its attributed features, our results implied that the number of topic-vehicle shared features was an important candidate factor affecting the categorical aspect of metaphor processing. Moreover, the results obtained in Experiments 2 and 3 preferred not the

comparison-based model account (e.g., Bowdle & Gentner, 2005) but the property attribution model. Comparison-based model account (e.g., Bowdle & Gentner, 2005; Ortony, 1979) supposed that the metaphor as an implicit simile: the meanings of metaphors are derived from the meanings of the similes. Therefore, their model predicts that both metaphor and simile form create the same interpretations; thus, the aptness of the metaphor and simile become the same. In contrast, the property attribution model supposes that metaphors pay more attention to the type of vehicle properties attributed to the topic is more important in a metaphor comprehension than in a simile. Our results showed the interaction between the type of topic-attributed features and the metaphor type. Thus, the results preferred the property attribution model to the comparison-based model.

There are two limitations to this research. First, the evaluation method of metaphor aptness should be correct. As pointed out by Jones and Estes (2006), metaphor aptness is sometimes confounded with other psycholinguistic variables (e.g., conventionality). These confounders can sometimes be removed by experimental manipulation (e.g., vehicle conventionality can be controlled by the repeated presentation of the vehicle words; Bowdle & Gentner, 2005). However, the way to avoid the aptness-related confounder was not explored except collecting the other variables like conventionality (i.e., Jones & Estes, 2006; this issue is also discussed in the direction for future research below). Related to this point, experimental manipulation can also improve the evaluation of metaphor aptness. In this study, all three experiments do not include the filler sentence (e.g., nonsense sentence). Filler sentences might rule out the possibility that participants misunderstand the meaning of aptness in metaphor evaluation.

Second, topic-attributed features should carefully be balanced. Oka and Kusumi (2020) showed that though both “bright” and “gorgeous” are topic-vehicle shared features for the metaphor “her smile is a flower”, “gorgeous” is a more apt topic-vehicle shared feature

because “bright” is shared by many other candidate vehicles (e.g., “her smile is a sun”). This implies that there were variations in aptness even among topic-vehicle sheared features. Though we recruited features from three questionnaires and carefully chose each type (i.e., topic-vehicle shared features, topic unique features, and vehicle unique feature) based on the count from each questionnaire, it still needs to be balanced by other evaluation methods. One possible way to eliminate this issue is to collect the saliency of each topic-attributed features for topic and vehicle separately.

There are two potential directions for future research, which our study did not examine. First, there is a possibility that aptness was confounded by the processing fluency. Thibodeau et al. (2018) collected the data regarding the candidate predictor that affects the metaphor processing (i.e., surprise, comprehensibility, metaphoricity, familiarity, and aptness) and other datasets which collected these factors. All the factors were collected based on the subjective rating as in our study. They conducted a principal component analysis and found two underlying sources of variance: figurativeness and processing fluency. In their study, the factor loading of familiarity and aptness was high in processing fluency but not in figurativeness. This study focused on the aptness itself and examined the factor (i.e., the number and the type of topic-attributed features) that affects aptness. Future research needs to address these two potential factors that affect metaphor aptness and processing fluency, using the sentence comprehension task described in Thibodeau et al. (2018).

Second is the effect of context on metaphor aptness. Bambini et al. (2014) showed that some psycholinguistic variables of a metaphor (e.g., familiarity, concreteness, etc.) rating were affected by the context (i.e., whether the target metaphor is embedded in a context). It could also be true for metaphor aptness. Our study partially supported this view; that is, the number and type of topic-attributed features affect metaphor aptness. However, it is still unclear whether the longer context (i.e., the metaphor embedded in the story context)

surrounded in the target metaphor also affects the metaphor aptness. Further investigation of metaphor aptness on context and context type is needed.

Acknowledgments

The authors declare that they have no conflict of interest. This study was approved by the ethics committee of the institution to which the second author belongs. The work was supported by JSPS KAKENHI Grant Number 19K12121.

References

- Bambini, V., Resta, D., & Grimaldi, M. (2014). A dataset of metaphors from the Italian literature: Exploring psycholinguistic variables and the role of context. *PloS ONE*, 9(9). <https://doi.org/10.1371/journal.pone.0105634>
- Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bowdle, B. F., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, 112(1), 193–216. <https://doi.org/10.1037/0033-295X.112.1.193>
- Chiappe, D. L., & Kennedy, J. M. (1999). Aptness predicts preference for metaphors or similes, as well as recall bias. *Psychonomic Bulletin & Review*, 6(4), 668–676. <https://doi.org/10.3758/BF03212977>
- Chiappe, D. L., & Kennedy, J. M. (2001). Literal bases for metaphor and simile. *Metaphor and Symbol*, 16(3/4), 249–276. <https://doi.org/10.1080/10926488.2001.9678897>
- Chiappe, D. L., Kennedy, J. M., & Chiappe, P. (2003). Aptness is more important than comprehensibility in preference for metaphors and similes. *Poetics*, 31(1), 51–68. [https://doi.org/10.1016/S0304-422X\(03\)00003-2](https://doi.org/10.1016/S0304-422X(03)00003-2)

- 1 Glucksberg, S., & Haught, C. (2006). On the relation between metaphor and simile: When
2 comparison fails. *Mind & Language*, 21(3), 360–378.
3 <https://doi.org/10.1111/j.1468-0017.2006.00282.x>
- 4 Gernsbacher, M. A., Keysar, B., Robertson, R. R., & Werner, N. K. (2001). The role of
5 suppression and enhancement in understanding metaphors. *Journal of Memory and*
6 *Language*, 45(3), 433–450. <https://doi.org/10.1006/jmla.2000.2782>
- 7 Glucksberg, S., McGlone, M., & Manfredi, D. (1997). Property attribution in metaphor
8 comprehension. *Journal of Memory and Language*, 36(1), 50–67.
9 <https://doi.org/10.1006/jmla.1996.2479>
- 10 Haught, C. (2013). A tale of two tropes: How metaphor and simile differ. *Metaphor and*
11 *Symbol*, 28(4), 254–274. <https://doi.org/10.1080/10926488.2013.826555>
- 12 Jones, L., & Estes, Z. (2005). Metaphor comprehension as attributive categorization. *Journal*
13 *of Memory and Language*, 53, 110–124. <https://doi.org/10.1016/j.jml.2005.01.016>
- 14 Jones, L., & Estes, Z. (2006). Roosters, robins, and alarm clocks: Aptness and conventionality
15 in metaphor comprehension. *Journal of Memory and Language*, 55(1), 18–32.
16 <https://doi.org/10.1016/j.jml.2006.02.004>
- 17 Kusumi, T. (1985). Similarity between constituent words or phrases of metaphors: Effects of
18 feature salience in comprehending metaphor. *Japanese Journal of Psychology*.
19 56(5), 269–276. <https://doi.org/10.4992/jjpsy.56.269>.
- 20 Nakamoto, K. (2003). Semantic priming effect of metaphor constituent terms. *Perceptual and*
21 *Motor Skills*, 96(1), 33–42. <https://doi.org/10.2466/pms.2003.96.1.33>
- 22 Nakamoto, K., & Kusumi, T. (2004). A classification of 120 Japanese metaphorical
23 expressions on the basis of four psychological dimensions. *The Science of Reading*,
24 48(1), 1–10. https://doi.org/10.19011/sor.48.1_1
- 25 Oka, R., & Kusumi, T. (2020). Distinctive features influence perceived metaphor aptness and

- 1 preference for metaphor use. *Metaphor and Symbol*, 35(1), 12–22.
- 2 <https://doi.org/10.1080/10926488.2020.1712780>
- 3 Oka, R. & Kusumi, T. (2021). Number of shared topic-vehicle significant features affects
- 4 speakers' preference for metaphorical expressions. *Journal of Cognitive*
- 5 *Psychology*, 33(2), 157–171. <https://doi.org/10.1080/20445911.2021.1876071>
- 6 Ortony, A. (1979). Beyond literal similarity. *Psychological Review*, 86(3), 161–180.
- 7 <https://doi.org/10.1037/0033-295X.86.3.161>
- 8 Qualtrics (2019). Qualtrics [Online survey platform]. Retrieved from
- 9 <https://www.qualtrics.com>
- 10 R Development Core Team (2016). *R: A language and environment for statistical computing*.
- 11 Retrieved from <http://www.r-project.org/>
- 12 Thibodeau, P. H., Sikos, L., & Durgin, F. H. (2018). Are subjective ratings of metaphors a red
- 13 herring? The big two dimensions of metaphoric sentences. *Behavior Research*
- 14 *Methods*, 50(2), 759–772. <https://doi.org/10.3758/s13428-017-0903-9>
- 15 Tversky , A. (1977). Features of similarity. *Psychological Review*, 84(4), 327-352.
- 16 <https://doi.org/10.1037/0033-295X.84.4.327>