Computational Evidence for Two-Stage Categorization as a Process of Adjective Metaphor Comprehension

Akira Utsumi (utsumi@se.uec.ac.jp)
Department of Systems Engineering
The University of Electro-Communications
1-5-1, Chofugaoka, Chofushi, Tokyo 182-8585, Japan

Maki Sakamoto (sakamoto@hc.uec.ac.jp)
Department of Human Communication
The University of Electro-Communications
1-5-1, Chofugaoka, Chofushi, Tokyo 182-8585, Japan

Abstract

Most of existing metaphor studies address comprehension of nominal metaphors like "My job is a jail" and predicative metaphors like "He shot down all of my arguments". However, little attention has been given to how people comprehend adjective metaphors such as "red voice". In this paper, we address adjective metaphors and argue that adjective metaphors are comprehended via a two-stage categorization process. In a two-stage categorization process, the adjective of an adjective metaphor evokes an intermediate category, which in turn evokes an abstract category of property to be mapped onto the target noun, rather than directly creating a category of property as predicted by the categorization theory. We then test our argument by means of computer simulation in which the meanings of adjective metaphors are computed from the representations of the adjective and the noun in a multidimensional semantic space constructed by latent semantic analysis. In the simulation, three algorithms for adjective metaphor comprehension, i.e., two-stage categorization, categorization and comparison, were compared in terms of how well they mimic human interpretation of adjective metaphors. The simulation result was that the two-stage categorization algorithm best mimicked human interpretation of adjective metaphors, thus suggesting that the two-stage categorization theory is a more plausible theory of adjective metaphor comprehension than the categorization theory and the comparison theory.

Keywords: Metaphor comprehension; Computational modeling; Latent semantic analysis (LSA); Adjective metaphor; Two-stage categorization

Introduction

Many studies in the domain of cognitive science have been made on the mechanism of metaphor comprehension. Although they have paid much attention to nominal metaphors such as "*My job is a jail*" (e.g., Bowdle & Gentner, 2005; Gentner, Bowdle, Wolff, & Boronat, 2001; Glucksberg, 2001; Jones & Estes, 2006; Utsumi & Kuwabara, 2005) and predicative metaphors such as "He *shot down* all of my arguments" (e.g., Lakoff & Johnson, 1980; Martin, 1992), little attention has been given to adjective metaphors such as "argumentative melody" and how they are comprehended. Some studies (e.g., Shen & Cohen, 1998; Werning, Fleischhauer, & Beşeoğlu, 2006; Yu, 2003) have focused on a synesthetic metaphor, a kind of adjective metaphor in which an adjective denoting the perception of one sense modality modifies a noun denoting a different modality. However these studies only examine how the acceptability of synesthetic metaphors can be explained by the pairing of adjective's and noun's modalities, rather than exploring the mechanism of adjective metaphor comprehension.

In this paper, we address the problem of how adjective metaphors are comprehended and argue that adjective metaphors are comprehended via a *two-stage categorization* process, which is an extended view of Glucksberg's categorization theory (Glucksberg, 2001; Glucksberg & Keysar, 1990). We then test our argument by means of computer simulation of adjective metaphor comprehension. For this purpose, we use a semantic space constructed by latent semantic analysis (LSA) (Landauer & Dumais, 1997) and provide a computational model of the two-stage categorization process, together with computational models of other possible processes for adjective metaphor comprehension such as categorization and comparison. In the computer simulation, we examine how well a computational model embodying each metaphor theory mimics human comprehension by comparing the interpretations of metaphors obtained by the computer simulation with human interpretations of the same metaphors obtained in a psychological experiment (Sakamoto & Sano, 2004). The metaphor theory that achieves the best simulation performance can be seen as the most plausible theory of adjective metaphor.

Adjective Metaphor Comprehension

Metaphor comprehension can be viewed as the process of finding relevant features (or predicates) that constitute the metaphorical meaning from the interaction between a source concept and a target concept, i.e., the process of generating the modified target concept in which some features or properties are highlighted and some other features are downplayed. In the case of adjective metaphors, the target concept is expressed by the head noun and modified by the source concept expressed by or associated with the adjective. The problem is how people determine which features of the target concept are highlighted or downplayed by the source concept.

One probable theory that can explain the mechanism of adjective metaphor comprehension would be the categorization theory of metaphor proposed by Glucksberg and his colleagues (Glucksberg, 2001; Glucksberg & Keysar, 1990). The categorization theory addresses mainly nominal metaphors and argues that people understand nominal metaphors by seeing the target concept as belonging to the superordinate metaphorical category exemplified by the source concept. Glucksberg (2001) has also argued that predicative metaphors function very much as do nominal metaphors; just as nominal metaphors use vehicles that epitomize certain categories of objects or situations, predicative metaphors use verbs that epitomize certain categories of actions. Some empirical evidence in favor of this view of predicative metaphors was also provided by Torreano, Cacciari, and Glucksberg (2005). Therefore, although they do not explicitly mention adjective metaphors in their works, it is likely that the same argument can be applied to adjective metaphors, that is, adjective metaphors use adjectives that epitomize certain categories of properties. According to this view, an adjective metaphor "red voice", for example, is comprehended so that the source concept *red* evokes an ad hoc category of property like "scary, screaming and dangerous" and such metaphorical property is mapped onto the target concept.

Against the categorization theory of adjective metaphors, we propose a two-stage categorization theory. The intuitive idea behind two-stage categorization is that correspondences between the properties literally expressed by the adjective and the properties to be mapped onto the target concept would be indirect, mediated by an intermediate category, rather than direct as predicted by the categorization theory. In the case of "red voice" metaphor, for example, the adjective *red* first evokes an intermediate category "red things", to which blood, fire, passion, apple, danger typically belong. Then exemplars relevant to the target concept *voice* such as blood, passion and danger are selected and they evoke a final abstract category of property like "scary, screaming and dangerous". [1]

An alternative, but probably less likely, explanation of adjective metaphor comprehension is given by the comparison theory of metaphor (Gentner, 1983; Gentner et al., 2001). This theory argues that metaphors are processed via a comparison process consisting of an initial alignment process between the source and the target concepts followed by a process of projection of aligned features into the target concept. According to the comparison theory, the "red voice" metaphor is comprehended in such a way that two concepts *red (or redness)* and *voice* are aligned, some features such as ones about scariness, scream or danger are found, and they are mapped onto the target noun.

In the rest of this paper, we examine which of these three theories best explains the mechanism of adjective metaphor comprehension by comparing them in terms of how accurately computational models embodying these theories simulate human behavior.

Computational Model

Vector Space Model

A vector space model is the most commonly used geometric model for the meanings of words. The basic idea of a vector space model is that words x are represented by high-dimensional vectors $\boldsymbol{v}(x)$, i.e., *word vectors*, and the degree of semantic similarity $sim(x, y)$ between any two words x and y can be easily computed as the cosine $\cos(\boldsymbol{v}(x), \boldsymbol{v}(y))$ of the angle formed by their vectors.

Word vectors are constructed from the statistical analysis of a huge corpus of written texts in the following way. First, all content words in a corpus are represented as m-dimensional feature vectors, and a matrix A is constructed using n feature vectors as rows. Then the dimension of M's rows is reduced from m to l. A number of methods have been proposed for computing feature vectors and for reducing dimensions (Utsumi & Suzuki, 2006). In this paper, we used an LSA technique (Landauer & Dumais, 1997) for constructing word vectors. LSA uses the term frequency in a paragraph as an element of feature vectors, and singular value decomposition as a method for dimensionality reduction. LSA was originally proposed as a document indexing technique for information retrieval, but several studies (e.g., Kintsch, 2001; Landauer & Dumais, 1997) have shown that LSA successfully mimics many human behaviors associated with semantic processing.

For example, using a semantic space derived from a corpus of Japanese newspaper used in this paper, similarity between *computer* ("*konpyuta*" in Japanese) and *Windows* ("*uindouzu*" in Japanese; Microsoft's OS) is computed as .63, while similarity between *computer* and *window* ("*mado*" in Japanese; glass in the wall) is computed as –.02.

Metaphor Comprehension Algorithms

In the vector space model, a vector representation $\boldsymbol{v}(s)$ of a piece of text s (e.g., phrase, clause, sentence, paragraph) consisting of constituent words $w_1, \cdots, w_n$ can be defined as a function $f(\boldsymbol{v}(w_1), \cdots, \boldsymbol{v}(w_n))$. Therefore, adjective metaphor comprehension is modeled as computation of a vector $\boldsymbol{v}(M) = f(\boldsymbol{v}(w_T), \boldsymbol{v}(w_S))$ which represents the meaning of an adjective metaphor M with the noun w_T (target) and the adjective w_S (source). In the rest of this paper, I use the phrase "n neighbors of a word (or a category) x" to refer to words with n highest cosine similarity to x, and denote a set of n neighbors of x by $N_n(x)$.

Categorization The algorithm of computing a metaphor vector $\boldsymbol{v}(M)$ by the process of categorization is as follows.

1. Compute $N_{m_1}(w_S)$, i.e., m_1 neighbors of the source w_S.
2. Selects k words with the highest similarity to the target noun w_T from $N_{m_1}(w_S)$.
3. Compute a vector $\boldsymbol{v}(M)$ as the centroid of $\boldsymbol{v}(w_T)$, $\boldsymbol{v}(w_S)$ and k vectors of the words selected at Step 2.

This algorithm is identical to Kintsch's (2000) predication algorithm and it is also used as a computational model of the categorization process in Utsumi's (2006) simulation experiment. As Kintsch suggests, this algorithm embodies the categorization view in that a set of k words characterizes an abstract superordinate category exemplified by the vehicle.

Two-stage categorization We propose the algorithm of two-stage categorization as follows.

1. Compute $N_{m_1}(w_S)$, i.e., m_1 neighbors of the source w_S.
2. Selects k words with the highest similarity to the target noun w_T from $N_{m_1}(w_S)$.
3. Compute a vector $\boldsymbol{v}(C)$ of an intermediate category C as the centroid of $\boldsymbol{v}(w_T)$, $\boldsymbol{v}(w_S)$ and the vectors of k words selected at Step 2.
4. Compute $N_{m_2}(C)$, i.e., m_2 neighbors of the intermediate category C.
5. Compute a metaphor vector $\boldsymbol{v}(M)$ as the centroid of $\boldsymbol{v}(w_T)$, $\boldsymbol{v}(w_S)$ and m_2 vectors selected at Step 4.

[1] Our preliminary experiment demonstrated that figurative meanings of adjectival metaphors with color adjectives were not directly associated with adjectives, but could be explained more appropriately by considering intermediate concepts associated with both adjectives and target nouns. This finding may lend support to our view based on two-stage categorization.

The first three steps, which are identical to the original categorization algorithm, correspond to the process of generating an intermediate category. Steps 4 and 5 correspond to the second categorization process.

Comparison The algorithm of computing a metaphor vector $\boldsymbol{v}(M)$ by the process of comparison is as follows.

1. Compute a set of k words (i.e., alignments between the target w_T and the source w_S) by finding the smallest i that satisfies $|N_i(w_T) \cap N_i(w_S)| = k$.
2. Compute a metaphor vector $\boldsymbol{v}(M)$ as the centroid of $\boldsymbol{v}(w_T)$ and k vectors computed at Step 1.

This algorithm is proposed by Utsumi (2006). Step 1 corresponds to the initial alignment process, while Step 2 corresponds to the later projection process.

Besides these three models, for comparison purposes, we also consider a simple combination algorithm by which a metaphor vector $\boldsymbol{v}(M)$ is computed as the centroid of the target vector $\boldsymbol{v}(w_T)$ and the source vector $\boldsymbol{v}(w_S)$.

Simulation Experiment

Method

Human experiment For human interpretation of adjective metaphors, we used the result of the psychological experiment reported in Sakamoto and Sano (2004). The materials used in the experiment were 50 Japanese adjective metaphors. They were created from all possible adjective-noun combinations of five adjectives (red ["*akai*"], blue ["*aoi*"], yellow ["*kiioi*"], white ["*shiroi*"], black ["*kuroi*"]) with 10 nouns (voice ["*koe*"], sound ["*oto*"], mind ["*kokoro*"], feeling ["*kimochi*"], words ["*kotoba*"], atmosphere ["*funiki*"], character ["*seikaku*"], past ["*kako*"], future ["*mirai*"], taste ["*aji*"]).

Thirty-eight undergraduate students of the University of Electro-Communications, who were all native speakers of Japanese, were assigned to all the 50 metaphors. They were asked to choose among 24 perceptual adjectives (i.e., features) appropriate ones for the meaning of each adjective metaphor. For each chosen feature w_i of an adjective metaphor M, the degree of salience $sal(w_i, M)$ is then assessed as the number of participants who chose that adjective. These features were used as landmarks with respect to which model's interpretation and human interpretation were compared for evaluation. Note that any adjective chosen by only one participant was not included in the analysis. For example, as shown in the bar graph of Figure 1, seven adjectives were chosen for the metaphor "black future", and the adjective *dark* had the highest salience, i.e., the number of participants (26 participants) who listed it was largest.

Computer simulation The semantic space used in the simulation experiment was constructed from a Japanese corpus of 251,287 paragraphs containing 53,512 different words, which came from a CD-ROM of Mainichi newspaper articles (4 months) published in 1999. The dimension l of the semantic space was set to 300, and thus all words were represented as 300-dimensional vectors.

In the computer simulation, for each of the 50 adjective metaphors, four kinds of metaphor vectors were computed using the four comprehension algorithms presented in the preceding section, i.e., categorization, two-stage categorization,

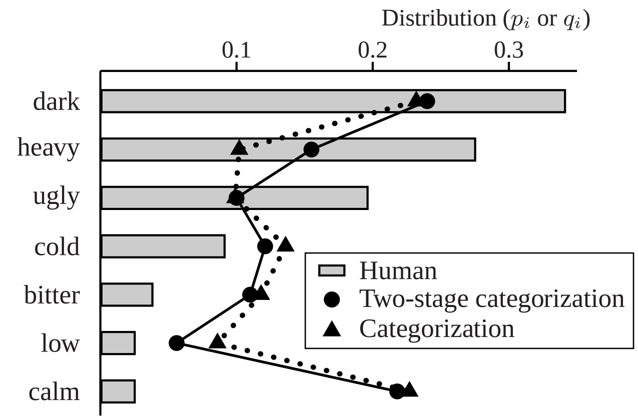


Figure 1: "Black future" metaphor

comparison and simple combination. In computing the metaphor vectors, we varied the parameter m_1 in steps of 50 between 50 and 500, and the parameters k and m_2 from 1 to 10. After that, for all the features $w_i, \cdots, w_n$ chosen for a metaphor M in the human experiment, similarity to the metaphorical meaning $sim(w_i, M)$ was computed separately using the four metaphor vectors. Features with higher similarity to the metaphorical meaning can be seen as more relevant to the interpretation of the metaphor. In Figure 1, for example, the word *dark* has the highest similarity to both the metaphor vectors computed by the categorization algorithm and by the two-stage categorization algorithm, but a least salient word *calm* is also highly similar to the metaphor vectors.

Evaluation measures To evaluate the ability of the model to mimic human interpretations, we use the following measures, which were also used in Utsumi's (2006) simulation experiment for nominal metaphors.

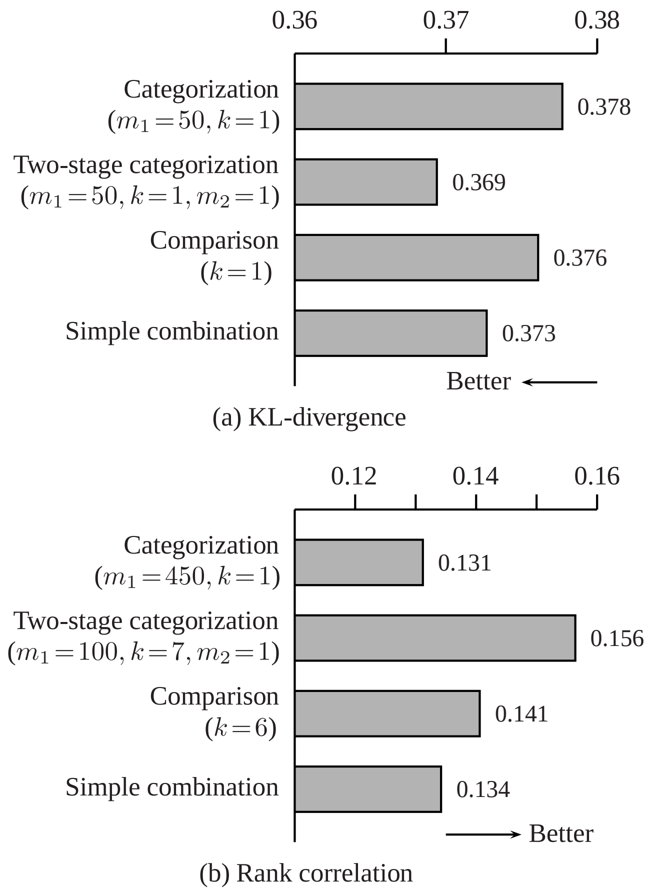
- *Kullback-Leibler divergence (KL-divergence)*:

$$D = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i} \tag{1}$$

$$p_i = \frac{sal(w_i, M)}{\sum_{j=1}^{n} sal(w_j, M)} \tag{2}$$

$$q_i = \frac{sim(w_i, M) - \min_x sim(x, M)}{\sum_{j=1}^{n}\{sim(w_j, M) - \min_x sim(x, M)\}} \tag{3}$$

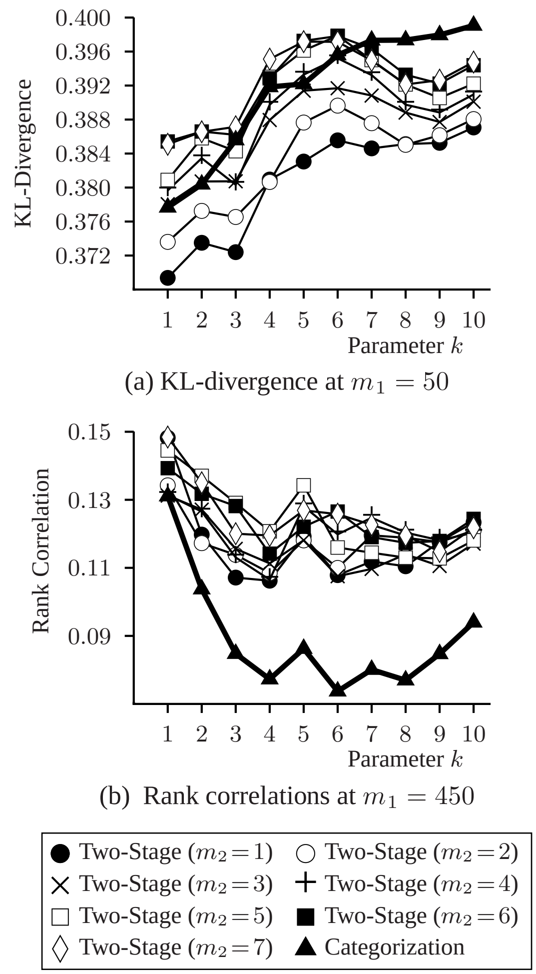
It measures how well a model simulates the salience distribution of features relevant to human interpretation, or in other words, the degree of dissimilarity between human interpretation p_i and computer's interpretation q_i. Hence, *lower divergence means that the model achieves better performance.* In Figure 1, for example, KL-divergence between the salience distribution of human interpretation and the similarity distribution of computer interpretation is 0.546 for the categorization model ($m_1 = 50$, $k = 1$) and 0.396 for the two-stage categorization model ($m_1 = 50$, $k = 1$, $m_2 = 1$). This result suggests that, in this case, the two-stage categorization model better mimics human interpretation than the original categorization model.



(a) KL-divergence

(b) Rank correlation

Figure 2: Simulation results: Comparison among the four comprehension models for adjective metaphors



(a) KL-divergence at $m_1 = 50$

(b) Rank correlations at $m_1 = 450$

Figure 3: Simulation results of the two-stage categorization model and the categorization model obtained with various values of parameters k and m_2

- *Spearman's rank correlation*:

$$r = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n^3 - n} \quad (4)$$

$$d_i = rank(sim(w_i, M)) - rank(sal(w_i, M)) \quad (5)$$

  It measures how strongly the computed similarity of relevant features is correlated with the degree of salience of those features. *A higher correlation means that the model yields better performance.* In Figure 1 the two-stage categorization model yields a higher correlation ($r = .46$) than the categorization model ($r = .28$), which again indicates that the two-stage categorization model is superior to the categorization model.

Result

For each of the 50 metaphors, KL-divergences and rank correlations were computed using the four metaphor vectors. These values were then averaged across metaphors. Concerning KL-divergence, the categorization algorithm achieved the best performance when $m_1 = 50$ and $k = 1$, the two-stage categorization model did the best performance when $m_1 = 50$, $k = 1$ and $m_2 = 1$, and the comparison model did the best performance when $k = 1$. Concerning rank correlation, the combination of $m_1 = 450$ and $k = 1$ was optimal for the categorization model, while the combination of $m_1 = 100$, $k = 7$ and $m_2 = 1$ was optimal for the two-stage categorization model. For the comparison model, $k = 6$ was optimal.

Figure 2 shows mean divergences and correlations calculated using these optimal parameters. The two-stage categorization model outperformed the other three models on both measures. It suggests that *the two-stage categorization theory is the most plausible theory of adjective metaphor comprehension.* Furthermore, in order to demonstrate that this simulation result in favor of the two-stage categorization theory is general, not specific to the particular value of the parameters, we show the simulation results obtained with various values of parameters in Figure 3. Figure 3(a) shows that, when they were compared at the same value of k, the two-stage categorization algorithm had lower divergence (i.e., better performance) than the categorization algorithm at almost all the values of m_2, although it had worse performance at some higher values of m_2 and lower values of k. Similarly, as shown in Figure 3(b), the two-stage categorization algorithm achieved a higher correlation (i.e., better performance) regardless of values of m_2. These results clearly indicate the plausibility of the two-stage categorization model as a cognitive theory of adjective metaphor comprehension.

Discussion

Related Work

Until now there have been some computational studies on metaphor comprehension. For nominal metaphors, Thomas and Mareschal (2001) proposed a connectionist implementation of comprehending nominal metaphors on the basis of the categorization theory, but they did not test the validity of their models in a systematic way, nor did they make a new contribution to the psychological or cognitive theory of metaphor. Kintsch (2000) proposes an LSA-based computational model of metaphor comprehension. His predication algorithm is also used in this study as a model of categorization, but he did not test its psychological validity as a model of metaphor comprehension. In addition, his study does not allow for the fact that some metaphors are comprehended as comparisons. Lemaire and Bianco (2003) also employ LSA to develop a computational model of referential metaphor comprehension. However, they do not address how well it mimics human interpretations; they only showed that it mimics processing time difference between when supporting context is provided and when it is not provided. Moreover, their model is theoretically less well motivated. For adjectival metaphors, Weber (1991) proposed a connectionist model of adjectival metaphors, which can be seen as one computational implementation of the categorization theory. This model uses two methods (direct value transference and scalar correspondence) for establishing semantic correspondences between the properties literally expressed by the adjective and the properties to be mapped onto the target concept. However, her model was not tested in a systematic way, either.

In contrast, our LSA-based computational methodology used in this study tests the validity of competing metaphor theories and predicts which is most plausible. Utsumi (2006) has applied this methodology to nominal metaphors and demonstrated that the interpretive diversity view of metaphor (Utsumi, 2007; Utsumi & Kuwabara, 2005) best explains the mechanism of nominal metaphor comprehension.

Does Two-Stage Categorization Better Explain Nominal Metaphor Comprehension?

In this paper, we have shown that adjective metaphors are comprehended via a two-stage categorization process, rather than via a categorization process or a comparison process. This raises a new interesting question whether or not people also comprehend other types of metaphors, especially nominal metaphors, via a two-stage categorization process.

Recent studies have claimed that people comprehend nominal metaphors as categorizations or comparisons depending on a metaphor property such as vehicle conventionality (Bowdle & Gentner, 2005), metaphor aptness (Jones & Estes, 2006) or interpretive diversity (Utsumi, 2007; Utsumi & Kuwabara, 2005). Especially Utsumi (2007) has demonstrated through a psychological experiment that interpretively diverse metaphors are processed as categorizations but less diverse metaphors are processed as comparisons. Utsumi (2006) also confirmed this finding by means of computer simulation. Therefore, the question mentioned above can be refined as follows: Does the two-stage categorization process better explain comprehension of high-diversity metaphors than the categorization process, and comprehension of low-diversity metaphors than the comparison process?

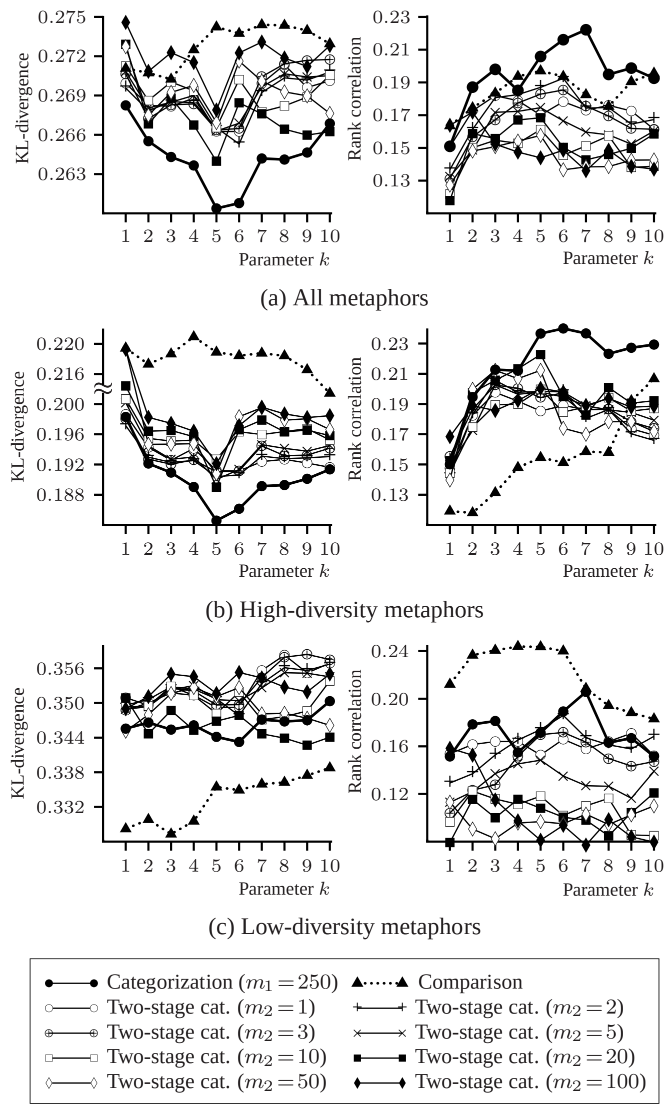


Figure 4: Simulation results of nominal metaphor comprehension ($m_1 = 250$)

In order to tackle this question, we conducted an additional simulation experiment in which the metaphorical meanings of 40 nominal metaphors such as "Life is a game" were computed by the two-stage categorization algorithm, and the results were compared with the results of the categorization algorithm and the comparison algorithm obtained in our preceding study (Utsumi, 2006). The simulation method and evaluation measures used in this additional experiment were identical to those used in the main simulation experiment of this study. For human interpretation of the nominal metaphors, the result obtained in a psychological experiment (Utsumi, 2005) was used. (For further details of the simulation experiment of nominal metaphors, see Utsumi, 2006).

The overall result was that the two-stage categorization algorithm did not achieve better performance than the categorization algorithm and the comparison algorithm. As shown in Figure 4(a), when the scores of all metaphors were averaged, the categorization algorithm had lower divergence and higher correlation (and thus better mimics human interpretation) than the two-stage categorization algorithm at the same value of the parameter k, regardless of value of m_2. Furthermore, Figure 4(b) also shows that the categorization algorithm outperformed the two-stage categorization model even when metaphors were highly diverse, and Figure 4(c) shows that, when metaphors were less diverse, the comparison algorithm outperformed the two-stage categorization model. These findings clearly indicate that people do not comprehend nominal metaphors via the process of two-stage categorization, and the interpretive diversity view (Utsumi, 2007; Utsumi & Kuwabara, 2005) is still the most plausible theory of nominal metaphor comprehension. In other words, the process of adjective metaphor comprehension essentially differs from the process of nominal metaphor comprehension.

Concluding Remarks

Our simulation experiment has shown that adjective metaphors are likely to be comprehended via a two-stage categorization process. We are now trying to confirm this finding by a psychological experiment. It would also be interesting for further work to investigate, both psychologically and computationally, whether people comprehend predicative metaphors via a two-stage categorization process.

Acknowledgments

This research was supported by Grant-in-Aid for Scientific Research(C) (No.17500171) from Japan Society for the Promotion of Science.

References

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

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

Gentner, D., Bowdle, B., Wolff, P., & Boronat, C. (2001). Metaphor is like analogy. In D. Gentner, K. Holyoak, & B. Kokinov (Eds.), *Analogical mind: Perspectives from cognitive science* (pp. 199–253). MIT 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.

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

Kintsch, W. (2000). Metaphor comprehension: A computational theory. *Psychonomic Bulletin & Review*, *7*(2), 257–266.

Kintsch, W. (2001). Predication. *Cognitive Science*, *25*(2), 173–202.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. The University of Chicago Press.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211–240.

Lemaire, B., & Bianco, M. (2003). Contextual effects on metaphor comprehension: Experiment and simulation. In *Proceedings of the 5th International Conference on Cognitive Modeling (ICCM'2003)*.

Martin, J. (1992). Computer understanding of conventional metaphoric language. *Cognitive Science*, *16*, 233–270.

Sakamoto, M., & Sano, M. (2004). Emergent features of color metaphors viewed as a combination of color terms and nouns. In *Proceedings of the 21st Annual Meeting of the Japanese Cognitive Science Society* (pp. 188–189). (in Japanese)

Shen, Y., & Cohen, M. (1998). How come silence is sweet but sweetness is not silent: A cognitive account of directionality in poetic synaesthesia. *Language and Literature*, *7*(2), 123–140.

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

Torreano, L., Cacciari, C., & Glucksberg, S. (2005). When dogs can fly: Level of abstraction as a cue to metaphorical use of verbs. *Metaphor and Symbol*, *20*(4), 259–274.

Utsumi, A. (2005). The role of feature emergence in metaphor appreciation. *Metaphor and Symbol*, *20*(3), 151–172.

Utsumi, A. (2006). Computational exploration of metaphor comprehension processes. In *Proceedings of the 28th Annual Meeting of the Cognitive Science Society* (pp. 2281–2286).

Utsumi, A. (2007). *Interpretive diversity explains metaphor-simile distinction.* (submitted for publication)

Utsumi, A., & Kuwabara, Y. (2005). Interpretive diversity as a source of metaphor-simile distinction. In *Proceedings of the 27th Annual Meeting of the Cognitive Science Society* (pp. 2230–2235).

Utsumi, A., & Suzuki, D. (2006). Word vectors and two kinds of similarity. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th Annual Meeting of the Association for Computational Linguistics (COLING-ACL 2006) Main Conference Poster Sessions* (pp. 858–865).

Weber, S. (1991). A connectionist model of literal and figurative adjective noun combinations. In D. Fass, E. Hinkelman, & J. Martin (Eds.), *Proceedings of the IJCAI Workshop on Computational Approaches to Non-Literal Language: Metaphor, Metonymy, Idioms, Speech Acts, Implicature* (pp. 151–160).

Werning, M., Fleischhauer, J., & Beşeoğlu, H. (2006). The cognitive accessibility of synaesthetic metaphors. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 2365–2370).

Yu, N. (2003). Synesthetic metaphor: A cognitive perspective. *Journal of Literary Semantics*, *32*(1), 19–34.