Experimental Examination of Feature Emergence in Metaphor Understanding with Consideration for Individual Differences

Asuka Terai (Asuka@Nm.Hum.Titech.Ac.Jp)
Global Edge Institute, Tokyo Institute of Technology, 2-12-1 Ohokayama
Meguro-ku, Tokyo, Japan

Masanori Nakagawa (Nakagawa@Hum.Titech.Ac.Jp)
Graduate School of Decision Science and Technology, Tokyo Institute of Technology, 2-12-1 Ohokayama
Meguro-ku, Tokyo, Japan

Abstract

The purpose of this study is to clarify the mechanism of feature emergence within the process of understanding metaphors of the form "A(target) like B(vehicle)" based on model simulation and psychological experiments. Feature emergence refers to low-salient features of a target and a vehicle being emphasized in the metaphor understanding process((Nueckles & Janetzko, 1997)(Becker, 1997)et al.). However, previous studies have not examined the mechanism of feature emergence in terms of individual differences. In this study, a psychological experiment is conducted using multiple model simulation results obtained with various parameter settings in order to clarify what kinds of metaphor can be understood when numerous features emerge. The experimental results indicate that within the understanding process for metaphors that are low in terms of their conventionality and understandability there is a great deal of feature emergence.

Keywords: metaphor, simile, neural network, feature emergence

Introduction

In this paper, the mechanism of feature emergence, occurring within the process of understanding for metaphors represented in the form of "A(target) like B(vehicle)", is investigated in terms of individual differences using a psychological experiment and simulations of a computational model of metaphor understanding. There are two types of theories within psychology to account for the understanding process of metaphors, where one noun is likened to another. The first type consist of comparison theories, such as the Structure Alignment Theory (Gentner & Wolff, 1997), which the second type are categorization theories, such as the Class Inclusion Theory (Glucksberg & Keysar, 1990). Within comparison theories, metaphor understanding is realized by aligning similar elements of the target and the vehicle (Gentner & Wolff, 1997). For example, in understanding the metaphor "Socrates is like a midwife", the understanding process is realized when similar elements relating to "help to birth" are identified in both "Socrates" and "midwife" and are mutually aligned. However, this theory faces difficulties from the perspective of distinguishing between targets and vehicles. In categorization theories, metaphor understanding is explained in terms of class-inclusion statements, where a target is regarded as being a member of an ad hoc category of which the vehicle is a prototypical member(Glucksberg & Keysar, 1990). For example, in comprehending the metaphor of "Socrates is like a midwife", the target of "Socrates" is considered as belonging to a "helpful" category that could be typically represented by a vehicle like "midwife". Both theories see metaphor understanding as basically requiring a knowledge structure for concepts(targets, vehicles and so on)(Kusumi, 1995), and hold that metaphorical expressions emphasize the high-salient features of a vehicle as the features of the target.

On the other hand, previous studies have shown that the low-salient features of a target and a vehicle are also emphasized within the process of metaphor understanding. And the studies reported that features play an important part in metaphor understanding (Becker, 1997)(Nueckles & Janetzko, 1997)(Gineste, Indurkhya, & Scart, 2000)(Utsumi, 2005). Previous studies have examined the relationships between feature emergence and metaphor characteristics. Participants were asked to respond with high-salient features of a target, a vehicle or a metaphor (that is "target" compared to "vehicle"). It should be noted that emergent features are usually determined as being features that are given as high-salient features of a metaphor but not as high-salient features of a target and a vehicle, when participants are asked to list the high-salient features of a target, of a vehicle and of a metaphor. Furthermore, metaphor characteristics are usually represented as mean evaluation ratings for each metaphor. Hence, those studies have tended to ignore the differences between participants who regard a metaphor as being understandable and thus produce many emergent features, on the one hand, and participants who find the same metaphor as being incomprehensible and do not produce any emergent features, on the other hand. In this paper, this difference is regarded as individual difference. Accordingly, this paper conducts an experiment using model simulation results in order to examine the relationships between feature emergence and metaphor characteristics with consideration for such individual differences.

Previous computational models of metaphor understanding have proposed that feature emergence is due to an interaction among features (Utsumi, 2000), (Terai & Nakagawa, 2007), (Terai & Nakagawa, 2008). However, while the first model (Utsumi, 2000), based on a psychological experiment, does not attempt to represent the dynamic interaction among features, the final two models (Terai & Nakagawa, 2007),(Terai & Nakagawa, 2008) is able to incorporate such dynamic interaction because they are constructed using a recurrent neural network. Furthermore, those two models can adequately cover many kinds of metaphorical expressions because they are based on a statistical language analysis. While Terai & Nakagawa's (2007) model suffers somewhat in its inability to distinguish targets and vehicles, because it is based on a comparison theory, Terai & Nakagawa's (2008) model avoids that problem because it is based on a categorization theory. The later model is also capable of computing feature emergence within the metaphor understanding process, and its psychological validity has been examined in an experiment. A parameter of the model indicates the strength of interaction influence, which makes it possible to simulate the understanding process either with many emergent features or without emergent features. Hence, this paper conduct a psychological experiment using the results of Terai & Nakagawa's (2008) model simulation in order to clarify the mechanism of feature emergence with consideration for individual differences.

The procedure for conducting the experiment is as follows:
Step 1: The model simulates three versions of the metaphor understanding process (with many emergent features, with few emergent features, and without any emergent features) by changing the value of the parameter representing the influence of interaction. Specifically, the parameter value determines of the level of feature emergence within the metaphor understanding process. The model outputs three kinds of feature sets as interpretations of a metaphor in accordance with the parameter values.
Step 2: In order to clarify whether feature emergence occurs as part of the understanding process for a given participant, the participants are asked to evaluate the validity of the various simulation results from Step 1. Thus, the Step 2 results provide indication as to which parameter setting is the most appropriately matches the metaphor understanding process of each participant in question.
Step 3: The same participants from Step 2 evaluate the characteristics of the metaphors (their conventionality, understandability, interestingness, and the similarity between a vehicle and a target).
Step 4: The mechanism of feature emergence is clarified by identifying the relationships between feature emergence and metaphor characteristics obtained at Steps 2 and 3.

Model Simulation of Metaphor Understanding

In order to represent the metaphor understanding processes with many emergent features, with a few emergent features and without any emergent features, Terai & Nakagawa's(2008) model of metaphor understanding is used. It can represent many types of metaphor understanding processes relating to feature emergence by changing the value of the parameter that represents the influences of interaction among features.

The model of Metaphor Understanding

The model (Terai & Nakagawa, 2008) is based on a statistical language analysis and consists of two processes: a categorization process and a dynamic interaction process. Firstly, the knowledge structure of concepts is estimated through statistical language analysis (Kameya & Sato, 2005) employing extracted frequency data for adjective-noun modifications and three types of verb-noun modification in Japanese. The statistical method assumes that the terms n_i(noun) and a_j(adjective or verb) co-occur through latent classes and that the co-occurrence probabilities of these terms, $P(n_i,a_j)$, can be computed using formula(1).

$$P(n_i,a_j) = \Sigma_k P(n_i|c_k)P(a_j|c_k)P(c_k), \qquad (1)$$

where c_k indicates the k-th latent class assumed in the method. The parameters ($P(n_i|c_k)$, $P(a_j|c_k)$, and $P(c_k)$) are estimated as the value that maximizes the log likelihood of the co-occurrence frequency data between n_i and a_j using the EM algorithm. In this paper, these parameters are estimated from extracted data consists of 21,671 noun types and 3,403 adjective types for adjective-noun modifications, 29,745 noun types and 22,832 verb types for verb-noun(object), 26,113 noun types and 21,487 verb types for noun(subject)-verb, and 28,451 noun types and 24,231 verb types for verb-noun(modification). The model deals with the 18,142 noun types (n_i) that are common to all four types of modification data. The conditional probability of the latent class c_k given the noun n_i ($P(c_k|n_i)$) is computed using Bayes' theory. The nouns(concepts) are represented by vectors using the conditional probability ($P(c_k|n_i)$).

In the categorization process model, a vector, representing an assigned target as a member of an ad hoc category for a vehicle, is estimated based on a categorization theory using the meaning vectors of concepts. The algorithm for the categorization process is as follows. First, the semantic neighborhood ($N(n_i)$) of a vehicle of size s is computed on the basis of similarity to the vehicle, which is represented by the cosine of the angles between meanings. Next, L concepts are selected from the semantic neighborhood ($N(n_i)$) of the vehicle on the basis of similarity to the target (L indicates the number of the selected concepts). Finally, a vector is computed for the centroid of the meaning vectors for the target, the vehicle and the selected L concepts as the vector representing the assigned target as a member of the ad hoc category for a vehicle. The computed vector represents the assigned target as a member of an ad hoc category for a vehicle and the vector is indicated using $V(M)$. The strength of relationship between feature a_j(adjectives or verbs) and $V(M)$ is indicated using $P(a_j|M)$.

In the dynamic interaction process model, the meaning of a metaphor is computed using the meaning vectors estimated by the categorization process($P(a_j|M)$) by applying the dynamic interaction process model using a recurrent neural network model(Fig.1). Each node corresponds to a feature and there are connections. These nodes have both inputs and outputs. The dynamics of the network are based on the following system of simultaneous differential equations(2):

$$\frac{dx_q(t)}{dt} = exp(-\alpha t)(-x_q(t) + \beta \Sigma_{q'} w_{qq'} x_{q'}(t) + I_q(M)), \quad (2)$$

where $x_q(t)$ represents the activation strength of the q-th node at time t. The range is between 1 and 0. $exp(-\alpha\ t)$ is the term for convergence which decreases according to time t. When $dx_q/dt = 0$, the node outputs $O_q(M) = x_q(t)$. The vector($O(M)$), which is a set of $O_q(M)$, represents the meaning of the metaphor M. $I_q(M)$ represents the input value of the q-th node related to the metaphor M. The value of $P(a_{j(q)}|M)$ is used as the input value $I_q(M)$ where a_j corresponds to the meaning of the q-th node. $w_{qq'}$ denotes the weight of the connection from the q'-th to the q-th node and is the correlation coefficient among the q-th and q'-th features related to the sibling concepts of the target and the vehicle. β denotes the influences of the dynamic interaction among features. The model can represent many types of metaphor understanding processes from the perspective of feature emergence by changing the value of the parameter β.

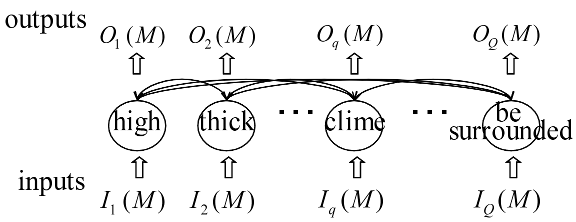

Figure 1: The architecture of the model of metaphor understanding (e.g. ”Difficulty like a wall”)

The Results of the Model Simulations

The model simulated the metaphor understanding processes concerning 16 metaphors in Japanese. In a previous study(Nakamoto & Kusumi, 2004), 120 metaphors were classified into 7 categories based on their characteristics. Three metaphors were selected from two categories consisting of very understandable metaphors and 2 metaphors were extracted from each of the other categories. These 16 metaphors are presented in Table1.

The model was simulated using the fixed parameters of $s = 50$, $L = 3$, $\alpha = ln(10)$. The value of the parameter β was changed $\beta = 0$, $\beta = 0.3$ and $\beta = 0.6$. The simulation results with $\beta = 0$ correspond to metaphor understanding without feature emergence, the simulation results with $\beta = 0.3$ correspond to low occurrences of feature emergence, while the results with $\beta = 0.6$ correspond to metaphor understanding with a considerable level of feature emergence. The simulation results are shown in Table2.

The pilot study was conducted for the purpose of confirming the efficiency of the parameter β. The participants were 85 undergraduates, who were divided into two groups (Group 1: 42 undergraduates; Group 2: 43 undergraduates). Participants in one group were asked to respond with appropriate features for the target and for the vehicle, while participants in the other group were asked to respond with appropriate features for the metaphor. Features given by three or more participants were regarded as being appropriate features for the vehicle, the target, or the metaphor, respectively, while the features that were given as being appropriate for the metaphor but not for the vehicle or the target were regarded as emergent features. The pilot study results indicate that the emergent features of ”difficulty like a wall” are ”run over”, ”get”, and ”block”. The model with $\beta = 0.3$ and $\beta = 0.6$ estimates ”be blocked” (as passive of ”block”) as the 9th most salient feature and as the 7th most salient feature, respectively. The model with $\beta = 0.3$ and $\beta = 0.6$ estimates the passive feature which is defined as the emergent feature through the pilot study. The results also demonstrate that the β value represents the extent of feature emergence within the metaphor understanding process.

Table 1: The model simulated the metaphor understanding processes for the 16 metaphors (”A (target) like B (vehicle)”)

target	vehicle
holiday	directing post
ballpark	bucket
compassion	flurry
time	flood
suspicion	tumor
love	season
blowing snow	muddy stream
eye	lake
demo	avalanche
conversation	gear
romance	fever
music score	cipher
affections	vortex
difficulty	wall
discussion	war
fury	eruption

Experiment using the Simulation Results

In order to clarify the relationships between feature emergence and the characteristics of metaphors, an experiment was conducted.

Method

Participants: 45 undergraduates.
Metaphorical expressions: 16 metaphors, which were used for the model simulation.
Simulation results: Results with $\beta = 0$, $\beta = 0.3$ and $\beta = 0.6$
Characteristics of the metaphors: understandability, conventionality, interestingness, similarity between a vehicle and a target, which are used in the previous study(Nakamoto &

Table 2: Metaphor meanings computed by the model (”difficulty like a wall”). The output values are shown in parentheses. These are the top 10 features. The emergent features are shown in bold type.

	$\beta = 0$	$\beta = 0.3$	$\beta = 0.6$
1	high (0.0431)	high (0.0558)	be surrounded (0.4000)
2	various (0.0293)	be surrounded (0.0438)	be covered (0.3985)
3	get over (0.0267)	be covered (0.0422)	run on (0.3956)
4	thick (0.0261)	run on (0.0393)	encounter (0.3926)
5	white (0.0242)	white (0.0372)	hit (0.3921)
6	be surrounded (0.0173)	encounter (0.0363)	climb (0.3907)
7	breach (0.0162)	hit (0.0358)	**be blocked** (0.3888)
8	be covered (0.0157)	climb (0.0344)	crash into (0.3877)
9	big (0.0157)	**be blocked** (0.0325)	plow into (0.3875)
10	collapse (0.0149)	crash into (0.0314)	be buried (0.3872)

Kusumi, 2004). 7-point scale, from 1 ”Strongly disagree” to 7 ”Strongly agree”.

The experiment consisted of two parts: evaluating the validity of the simulation results and evaluating the characteristics of the metaphors.

Evaluating the Validity of the Simulation Results

In order to clarify whether feature emergence occurs within the understanding process of a given participant, the participants evaluated the validity of the simulation results. The participants were presented with each metaphor and the three simulation results (top 10 features) with $\beta = 0$, $\beta = 0.3$ and $\beta = 0.6$, although the significance of the different output values was hidden with a story about them being the results from three different computers (Computer A, Computer B and Computer C). The participants were asked to evaluate the validity of each interpretation and to choose one computer that had simulates the most appropriate interpretations.

In this paper, β^* indicates the β that was used in the simulation that yielded the results that were selected as being the most appropriate interpretation. The mean and the entropy of β^*, as well as the mean ratings for the simulation results using β^*, are shown in Table3.

On the scale, 5 corresponded to ”slightly agree that the interpretation is appropriate” (with 4 being ”neutral”). The rating means for 14 metaphors are in excess of 5, while the means for the remaining two metaphors are in excess of 4. These results indicate the validity of the simulation results.

The entropy of β^* represents the individual difference in terms of feature emergence. If one third of the participants chose $\beta^* = 0$, one third chose $\beta^* = 0.3$ and the other one third chose $\beta^* = 0.6$, the entropy of β^* would be 1.59. The entropies for 2 metaphors are less than 1. This indicates that there are individual differences relating to feature emergence within each metaphor understanding process.

Evaluating the Characteristic of the Metaphors

In the second part of the experiment, the participants were asked to evaluate the characteristics of the metaphors. The participants were presented with each metaphor and they were asked to evaluate its characteristics. The mean ratings for metaphor characteristics are presented in Table4.

Table 3: Evaluation results concerning the validity of the simulation results.

metaphor	mean	entropy	rating
holiday-directing post	0.49	1.13	4.91
ballpark-bucket	0.37	1.51	5.11
compassion-flurry	0.43	1.37	4.82
time-flood	0.41	1.36	5.56
suspicion-tumor	0.43	1.29	5.22
love-season	0.51	0.92	5.67
blowing snow-muddy stream	1.58	1.09	5.80
eye-lake	0.15	1.26	5.51
demo-avalanche	0.25	1.44	5.87
conversation-gear	0.32	1.57	5.38
romance-fever	0.24	1.50	5.38
music score-cipher	0.45	1.19	5.31
affections-vortex	0.25	1.52	5.84
difficulty-wall	0.29	1.58	5.73
discussion-war	0.45	1.28	6.07
fury-eruption	0.51	0.96	5.13

Relationships between Feature Emergence and the Metaphor Characteristics

In order to clarify the relationships between feature emergence and metaphor characteristics, a number of statistical analyses are conducted using 717 responses (45 participants multiplied by 16 metaphors minus 3 not available responses). A one-way analysis of variance was carried out to examine the differences in the ratings for each characteristic as a function of the β used for the simulation results that were chosen as being the most appropriate interpretation (β^*). The results indicated that while there were no significance differences for conventionality, interestingness and similarity, there was a significantly effect of β^* for understandability

Table 4: Evaluation results for the metaphor characteristics (US: understandability, CV: conventionality, IR: interestingness, SL: similarity)

metaphor	US	CV	IR	SL
holiday-directing post	2.59	2.00	3.57	2.23
ballpark-bucket	2.80	1.91	2.39	2.29
compassion-flurry	3.22	2.36	3.22	2.67
time-flood	3.40	2.47	3.33	2.91
suspicion-tumor	3.62	2.47	3.53	3.22
love-season	3.82	2.36	4.18	3.64
blowing snow-muddy stream	4.13	2.62	3.42	3.51
eye-lake	4.15	2.76	4.42	3.33
demo-avalanche	4.52	3.57	3.86	4.26
conversation-gear	5.11	3.49	4.47	4.07
romance-fever	5.44	3.58	4.73	4.64
music score-cipher	5.44	4.00	4.96	4.60
affections-vortex	5.60	4.11	4.76	4.29
difficulty-wall	5.96	5.04	4.16	5.13
discussion-war	6.00	4.42	4.73	4.69
fury-eruption	6.13	4.93	4.42	5.64

Table 5: The results of two-way analysis of variance testing differences in conventionality as a function of both understandability (low or high) and β^*.

	Sum Sq	Df	F value	Pr(>F)
β^*	4.69	2	1.23	0.29
understandability	673.28	1	352.2	<2e-16
β^* x understandability	10.64	2	2.78	0.06
Residuals	1359.13	711		

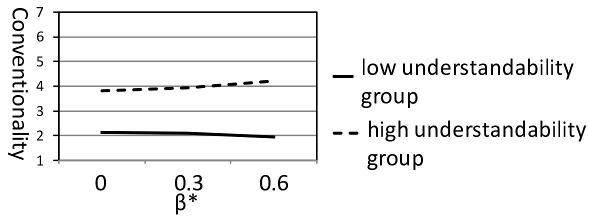

Figure 3: Differences in the ratings for conventionality as a function of β^*

(F(2,714)=3.06, P<.05), as shown in Fig.2. This result suggests that there is a relationship between understandability and feature emergence.

Reponses to each metaphor by each participant were then are divided into two groups according to the ratings for understandability, with understandability ratings of 1- 4 comprising the low-understandability group and a high-understandability group consisting of the remaining responses. A two-way analysis of variance was conducted for the each characteristic apart from understandability as a function of β^*. The results indicated a marginally significant interaction in terms of conventionality (as showing in Table5 and Fig.3). However, no significance differences were observed in terms of interestingness and similarity. These findings indicate that feature emergence (β^*) can be influenced by a combination of the conventionality and understandability characteristics.

Accordingly, all responses were subsequently are divided

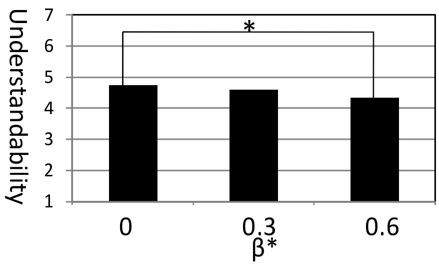

Figure 2: Differences in the ratings for understandability as a function of β^* (* p<.05 (Tukey test))

Table 6: Cross tabulation of the understandability groups and the conventionality groups

		Understandability	
		high	low
Conventionality	high	193	4
	low	243	273

into two groups according to the ratings for conventionality, with conventionality ratings of 1- 4 comprising the low-conventionality group, and a high-conventionality group consisting of the remaining responses. Cross tabulation of the understandability groups and the conventionality groups is shown in Tab.6.

The number of responses in the low-understandability and the high-conventionality cell is only 4. Hence, these 4 cases were ignored, and a one-way analysis of variance was conducted to test for differences between the other three cells (high-understandability and high-conventionality (H-H), high-understandability and low-conventionality (H-L), and low-understandability and low-conventionality (L-L)) as a function of β^*. The results indicated a significant effect of β^* on these three groupings (F(2,710)=5.21, P<.001), as shown in Fig.4.

The results indicate that the value of β^* was highest for the L-L grouping, while the value of β^* was lowest for the H-L grouping. Thus, while feature emergence would seem to occur when a metaphor is neither understandable nor conventional, feature emergence seems to happen less when a metaphor is very understandable but not conventional.

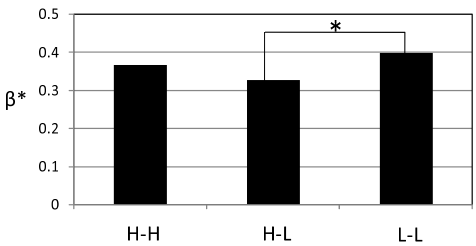

Figure 4: Differences in β* values as a function of the various understandability and conventionality groupings (* p<.05 (Tukey test))

In order to clarify individual differences in terms of feature emergence, the entropy values for β* were analyzed. Similar to the previous analysis, the metaphors were divided into three groups (H-H, H-L, L-L) based on the mean understandability ratings and the mean conventionality ratings. This resulted in 6 metaphors being classified into the H-H group, 5 metaphors for the H-L group, and 5 metaphors for the L-L group. The results of a one-way analysis of variance for the value of β* indicated that there was a significant difference between the L-L group and H-L group. Hence, the difference in entropy between the H-L group and the L-L group was tested using a t-test. The results indicated that the mean of β* for the H-L group (1.47) was significantly higher than the mean for the L-L group (1.26) at 10% level ($t(9)=1.93, p<.1$). This finding indicates that there are greater individual differences in terms of feature emergence for the H-L group metaphors than for the L-L group metaphors.

Discussion

In order to examine the mechanism of feature emergence without ignoring individual differences, a psychological experiment was conducted using simulation results for Terai & Nakagawa's (2008) model. The results of the experiment indicate that feature emergence occurs when a metaphor is neither understandable nor conventional, but that it happens less a metaphor is very understandable but not conventional. The results also indicate that participants tend to be more consistent in their processing of metaphors that are neither understandable nor conventional, but that there is greater individual variation when processing of metaphors that are very understandable but not conventional. These findings suggest that metaphors that are not understandable cannot be comprehended only with the high-salient features of the target and the vehicle, and that they require the additional activation of emergent features. The finding of no significant differences between the high-understandability and high-conventionality (H-H) group and the low-understandability and low-conventionality (L-L) group seems to suggest that, while some metaphors require some level of activation of emergent features to be understood, some metaphors become sufficiently familiar that they can be understood more readily.

In this research, understandability, conventionality, interestingness and similarity between a vehicle and a target were singled out as characteristics of metaphors. However, previous research has also examined that the relationships between feature emergence and the poetic appreciation of metaphors, and argued that there are also individual differences in terms of poetic appreciation((Utsumi, 2005)). With the present experimental method of employing simulation results, it will also be possible to investigate such relationships in greater detail.

Acknowledgments

This research is supported by MEXT's program "Promotion of Environmental Improvement for Independence of Young Researchers" and by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (B)(19330156).

References

Becker, A. (1997). Emergent and common features influence metaphor interpretation. *Metaphor and symbol*, *12*, 243–259.

Gentner, D., & Wolff, P. (1997). Alignment in the processing of metaphor. *J. of memory and language*, *37*, 331–355.

Gineste, M., Indurkhya, B., & Scart, V. (2000). Emergence of features in metaphor comprehension. *Metaphor and symbol*, *15*, 117–135.

Glucksberg, S., & Keysar, B. (1990). Understanding metaphorica comparisons: Beyond similarity. *Psychological review*, *97*, 3–18.

Kameya, Y., & Sato, T. (2005). Computation of probabilistic relationship between concepts and their attributes using a statistical analysis of japanese corpora. In *Proc. of symposium on large-scale knowledge resources* (pp. 65–68).

Kusumi, T. (1995). *Hiyu no syori katei to imikozo*. Kazama syobo.

Nakamoto, K., & Kusumi, T. (2004). Hiyuzairyobun no sinritekitokusei to bunrui: Kijunhyo sakusei no kokoromi. *Dokushokagaku*, *48*, 1–10.

Nueckles, M., & Janetzko, D. (1997). The role of semantic similarity in the comprehension of metaphor. In *Proc. of the 19th annual meeting of the cognitive science society* (pp. 578–583).

Terai, A., & Nakagawa, M. (2007). A neural network model of metaphor understanding with dynamic interaction based on a statistical language analysis; targeting a human-like model. *Int. J. of neural systems*, *17*, 265–274.

Terai, A., & Nakagawa, M. (2008). A corpus-based computational model of metaphor understanding incorporating dynamic interaction. In *V.kurkova et al. (eds): Proc. of icann 2008, part2, lncs 5164* (pp. 443–452). Springer-Verlag Berlin Heidelberg.

Utsumi, A. (2000). Hiyu no ninchi / keisan moderu. *Computer today*, *96*, 34–39.

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