Title
A mechanism of ontological boundary shifting

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 26(26)

ISSN
1069-7977

Authors
Hidaka, Shohei
Saiki, Jun

Publication Date
2004

Peer reviewed

A mechanism of ontological boundary shifting

Shohei Hidaka (hidaka@cog.ist.i.kyoto-u.ac.jp)
Jun Saiki (saiki@cog.ist.i.kyoto-u.ac.jp)
Graduate School of Informatics, Kyoto University;
Yoshida-Honmachi, Sakyo-ku, Kyoto, 606-8501, JAPAN

Abstract

Past research on children's categorizations has suggested that children use perceptual and conceptual knowledge to generalize object names. Especially, some researches suggested that the relation between ontological categories and linguistic categories is a critical cue to categorize objects. However, this mechanism has not been specified. This paper reports new insights to reveal children's categorizations based on the survey of adults' knowledge. We estimated the English and Japanese ontological spaces from data and used these results to simulate behavioral experiment of previous research. The results show a possibility that linguistic cues help children to attend specific perceptual properties.

Introduction

Categorization is a form of information compression, one solution to handle an almost infinite number of entities efficiently. Where do these categories come from and how do children know which words to map to which categories?

Quine (1960) pointed to the difficulty of word learning without prior category knowledge. If we hear a novel word in an unknown language, how do we infer its meaning ? For example, suppose we heard 'gavagai' while looking at a rabbit in a field. 'Gavagai' might mean *rabbit*, but it could also mean *rabbit's color* or an infinite variety of other possibilities. By Quine's analysis, word learning should be highly problematic for first language learners.

Constraints to acquire word meanings

Markman & Hutchinson (1984) among others proposed that children learn nouns easily because they do have prior knowledge about kinds of categories. Research over the past 20 years has indicated that this knowledge is considerable. Children know, for example, that animal categories are organized by multiple similarities, that artifact categories are organized by shape, and that substance categories are organized by material. Given a single thing of each of these kinds and told its name, children systematically generalize that name to new instances in ways specific to the kind of thing it is (Landau, Smith & Jones, 1988; Soja, Carey & Spelke, 1991, etc.) One hypothesis is that this knowledge is learned, that as children learn common nouns, they learn the correlations between properties specific to different kinds and the similarities relevant to categorizing those kinds - that things with eyes are classified by multiple similarities, things that are solid and rigid are classified by shape, and things that are nonsolid and nonrigid are classified by material (Jones & Smith, 2002; Yoshida & Smith, 2003). The learning hypothesis is plausible because children's differential categorizations of animates, objects and substances emerge only after they have learned some number of names for these different kinds (Samuelson & Smith 1999), The present paper provides a simulation model of how this might be learned.

Linguistic categories and ontological categories

The fact that children distinguish animal, object, and substance categories in noun learning is also interesting because these ontological categories are often related to linguistic individuation, to how different languages quantify nouns. Most of the world's languages treat animates as countable discrete things. Others, like English, also treat inanimate objects as discrete and countable. Few (if any languages) individuate substances in these ways. (Lucy, 1992)

Some have suggested a deep relation between ontological categories and their learning how their language quantifies entities. For example, Quine (1960) hypothesized that children learning English learn to distinguish objects and substances by learning the count-mass distinction. In English, nouns such as "dog" and "cup" that label individuated things are count nouns and mandatorily take the plural if there are more than one instance; in contrast, nouns such as "sand" that label a substance are mass nouns and are not pluralized. Thus this linguistic distinction could teach children that there are two different kinds of categories.

Soja, Carey and Spelke (1991) criticized this idea, because their experiments indicated that 2-year-old children who do not use count-mass syntax nonetheless classify objects by shape and substances by material. Imai & Gentner (1997) reported supporting results in a study comparing English and Japanese speakers. Japanese makes no distinction comparable to the count-mass distinction in its quantification system. Yet Imai and Gentner found that both Japanese- and English-speaking children categorized objects by shape and substances by material.

A boundary shift Imai and Genter also found differences in the range of things treated as objects versus substances by speakers of the two languages. Speakers of English treated complex and simple solids as objects categorized by shape and nonsolid forms as substances categorized by material. In contrast, Japanese speakers treated complex solids as objects classifiable by shape and treated simple solids and nonsolids as substances classifiable by material. Yoshida and Smith interpreted these results in terms a boundary shift, suggesting that count-mass syntax shifted the object boundary in English relative to that in Japanese so that it also included simple but solid shapes. If this interpretation is correct then linguistic contrasts such as count-mass syntax may play a role in the development of ontological categories.

Yoshida and Smith also predicted and found a corresponding boundary shift at the animate-inanimate boundary. They predicted this from an analysis of Japanese which distinguishes animates and inanimates in ways that English does not, through its quantification system and also via the verbs "iru" and "aru" which mean "exists". In locative constructions such as "There is," animates require the use of iru and inmates require the use of aru. Yoshida and Smith hypothesized that this distinction, like the count-mass distinction in English, would perturb the boundary between animates and objects. In support of this idea, they showed that the range of things treated as animates (and classified by multiple similarities) was broader for Japanese-speaking than English-speaking children.

The purpose of the simulations reported here is to explain the mechanism underlying the boundary shift. Following Yoshida and Smith, we propose that ontological categories are the product of learned correlations among the properties such as shape, material and color and also linguistic contrasts such as the count-mass distinction in English and the iru-aru distinction in Japanese.

Experiment 1

We measured statistical structure of common noun category via adult judgments. We studied the statistical structure of 48 nouns that name common categories typically known by 2 year olds (Fenson, Dale, Reznick, Bates, Thal & Pethick, 1994) [1]We did this in two steps. First, we asked adults to judge how a list of 16 adjectives taken from some studies using the Semantic Differential (the SD; Osgood, 1957) described category relevant properties such as shape, material, movement. Second, in vocabulary survey, we asked adults to rate how the 16 adjectives described the 48 noun categories.

Method

Participants In the adjective survey, we recruited 12 volunteers (from 23 to 25 years old) from Kyoto university. In the vocabulary survey, we recruited 104 students (from 18 to 22 years old) from Kyoto Koka women's university who received a class credit for participation.

Stimuli The stimuli consisted of (1) a list of category relevant perceptual properties: shape, material, color, texture, sound, temperature, flavor, movement, smell, and function, (2) 16 adjective pairs: dynamic-static, wet-dry, light-heavy, large-small, complex-simple, slow-quick, quiet-noisy, stable-unstable, cool-warm, natural-artificial, round-square, weak-strong, rough hewn-finely crafted, straight-curved, smooth-bumpy, hard-soft; and (3) 48 nouns commonly known by young children [2] (see also Table 1).

Adjective survey Participants were asked - 'How do you use these words (adjective pairs) to express familiar objects' perceptual features'. Participants made these judgments using an electronic file of the 16 (adjectives) by 10 (properties) cells. The ratings were on a 5 point scale (1: very inappropriate, 2: inappropriate, 3: neither, 4: appropriate, 5 :very appropriate).

Vocabulary survey Participants were presented with one noun at a time and asked to judge the applicability of the 16 adjective pairs on a 5 point scale. For example, if the adjective pair was big-small, participants would be asked the thing labeled by the noun very small, small, ambiguous, big and very big. Five different orders were used across subjects.

Analysis We used Principal Component Analysis (PCA) to analyze the vocabulary with mean linguistic-scale scores of the all participants. PCA is a popular method to compress information by the least loss of data variance.

We used the results to estimate the English and Japanese ontological spaces. We added 1-dimension syntactic cues which was close to ontological categories (Table 1) to raw data (16 dimensions), and analyzed the combined data (17 dimensions). In the English condition, we added count-mass syntax which was encoded as 1-0. In the Japanese condition, we added iru-aru syntax just as in the English condition. In the neutral condition, we added the value 0.5 for all objects. We decided these parameters of syntactic categories based on the dictionaries. We assumed that (1) our ontology space consists of perceptual and linguistic properties, and that (2) the most important factor of these space is the variance of the object's distribution. These assumptions are reasonable, because (1) our goal is to estimate children's ontology space in the context of generalizing novel names and (2) we name entities different labels based on not similar features but different properties.

Our another goal is to estimate perceptual weights in two language conditions. However, principal components consist of weights of linguistic scales, so we can not directly know which perceptual weights the ontology spaces have. Therefore we defined perceptual weights of principal components as the equation(1) to analyze perceptual weights in English and Japanese conditions.

$$W_{dp} = \mid \sum_{l} C_{dl} M_{lp} \mid \qquad (1)$$

[1]This form of the MCDI is a parental checklist of words designed to measure the productive vocabulary of children between 16 and 30 months of age.

[2]The 9 categories are 'animals', 'body parts', 'clothing', 'food and drink', 'furniture and rooms', 'outside things', 'small household items', 'toys', and 'vehicles'.

Table 1: Linguistic categories of 48 nouns in English and Japanese. E=English, J=Japanese, c=count noun, m=mass noun, i=with-'iru' noun, a=with-'aru' noun

	E	J		E	J		E	J
butterfly	c	i	banana	c	a	water	m	a
cat	c	i	egg	c	a	camera	c	a
fish	c	i	ice cream	c	a	cup	c	a
frog	c	i	milk	m	a	key	c	a
horse	c	i	pizza	c	a	money	m	a
monkey	c	i	salt	m	a	paper	m	a
tiger	c	i	toast	c	a	scissors	c	a
arm	c	a	bed	c	a	plant	c	a
eye	c	a	chair	c	a	balloon	c	a
hand	c	a	door	c	a	book	c	a
knee	c	a	refrigerator	c	a	doll	c	a
tongue	c	a	table	c	a	glue	m	a
boots	c	a	rain	m	a	airplane	c	a
gloves	c	a	snow	m	a	train	c	a
jeans	c	a	stone	c	a	car	c	a
shirt	c	a	tree	c	a	bicycle	c	a

d is a dimension of principal components. l is a index of 16 linguistic scales of the SD (see also Method). p is the index of the 10 perceptual properties (see also Method). W_{dp} is the pth perceptual weight of dth principal component. C_{dl} is the loading of lth linguistic scales of dth principal component. M_{lp} is the estimated expressiveness of the pth perception of the lth linguistic scales. C_{d*} is a unit row vector and M_{*p} is a unit column vector, so W_{dp} is the absolute inner product of two vectors, or $\mid cos\theta \mid$ (θ is the angle of two vectors).

Results and Discussion

First three and six principal components respectively accounted for more than 70% and 90 % of the variability in the data.

Estimated ontological spaces The first two principal components of the vocabulary survey data were displayed as a 2-dimensional plot (Figure 1 is the result of neutral condition). In the neutral condition, we found animates and body parts in upper-right area, vehicles in upper-left area, furniture in lower-left area and substance in lower-right area. This distribution of entities leads us the following interpretation of the first two components. The first principal component axis can be interpreted as 'solidity', because solid and non-solid entities are located in the left and right sides, respectively. The second principal component axis can be interpreted as 'animacy', because dynamic and static entities are located in the upper and lower sides, respectively.

There were no clear boundaries in neutral 2-dimensional space, but we found global boundaries in the English and Japanese space. Furthermore, the English and Japanese spaces had a great difference. The English space also had 'solidity' axis as the first principal component, but the Japanese space had 'animacy' axis as the first principal component. Therefore, we analyzed these distributions of entities by clustering.

First three principal components (total 70% over) were enough to analyze global structure of results, so we analyzed this 3-dimensional data by hierarchical clustering(Figures 2 and 3).

The clustering of the neutral condition showed the clusters like MCDI classes, but did not show any global boundaries. On the other hand, the analyses of the English and Japanese conditions showed the global boundary (Figures 2 and 3). There were two global clusters categorized near by the root of the tree. One cluster mainly consisted of 'objects' category members, and another cluster mainly consisted of 'substance' category members. The second branch occurred in the object cluster. There was the 'animates' cluster near substance cluster in the part of objects cluster. That is why, English ontology space seemed defined by 'individuation' or 'solidity'.

In the Japanese condition, there were two global clusters that mainly consisted of 'animates' members and 'inanimates' members. Despite being inanimates, vehicles(e.g. 'airplane', 'car') and body parts(e.g. 'eye', 'hand') were near the animates members. There seemed an 'animacy' boundary in the Japanese ontological space because animates and dynamic objects make cluster and inanimates make another cluster.

Table 2: The estimated perceptual weights. In the Experiment 2, we used the normalized W_{dp} ($\sum_p^{10} W_{dp} = 1$).

	English	Japanese
shape	0.091	0.047
color	0.067	0.194
texture	0.086	0.09

Perceptual weights of the English and Japanese spaces We estimated perceptual weight in the English and Japanese ontological spaces. Tables 2 shows the results of the estimation.

Compared with the Japanese condition, the English condition showed higher weight on shape. Contrary to the English condition, the Japanese condition showed higher weights on color and texture.

Estimating ontological category One potential problem with the present experiment is that the perceptual ontological space was derived from only Japanese speaker's data.

We are currently collecting the English data. Preliminary results indicate that they are extremely similar to those of Japanese speakers. The Pearson's correlation of mean across participants are .79 in vocabulary survey and .80 in adjective survey. In this work, we have assumed that the adjective - noun ratings and the adjectives - properties ratings reflect the perceptual structure of the categories. It could be argued that these rating reflect instead how predicates and nouns co-occur in a language.

Sommers (1963) claimed that knowledge of ontological categories is intimately related to predicability, that is, to the knowledge of which predicates in a language can be combined with which nouns. For example, the predicate 'is asleep' distinguishes animals and non-animals. Furthermore, Keil (1979, 1981) showed that children's judgments of predicability, like those of adults, yield an

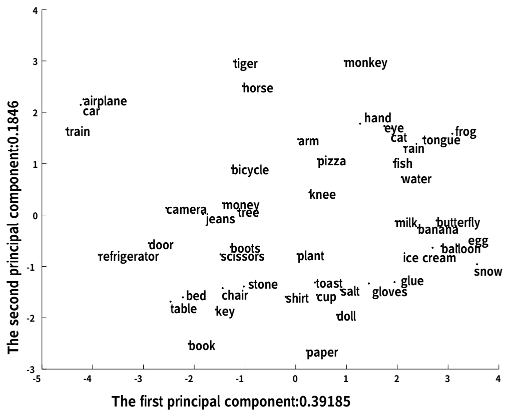

Figure 1: The first two principal components for the neutral condition. The first principal component (x axis) was interpreted as 'solidity' or 'size' of objects. The second principal component(y axis) was interpreted as 'animacy' or 'movement' of objects.

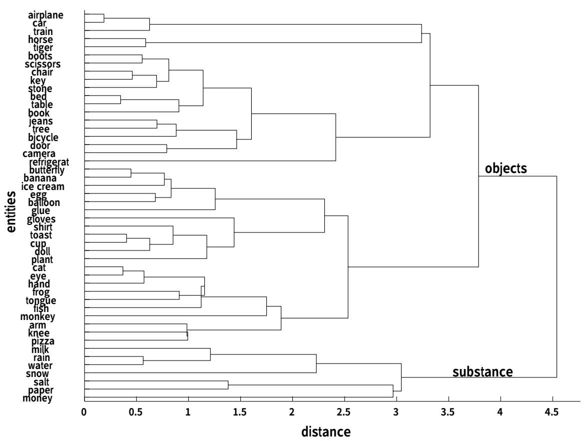

Figure 2: The result of cluster analysis for the English condition. We estimated 'objects' cluster and 'substance' cluster in superior hierarchy.

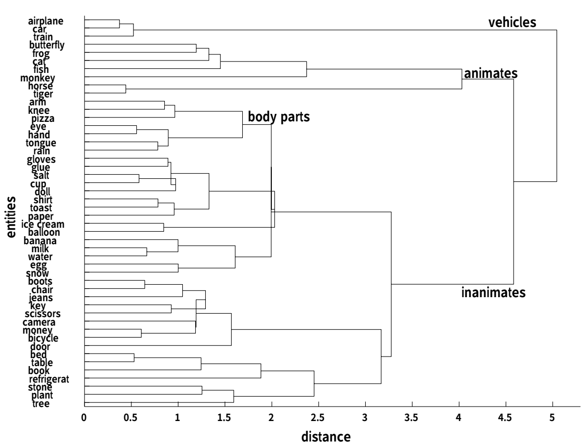

Figure 3: The result of cluster analysis for the Japanese condition. We estimated 'animates' cluster and 'inanimates' cluster in superior hierarchy.

ontological tree, though a less elaborate one than adults. The question of whether our judgments reflect the structure of categories in the world or relations among words is a difficult one. Given our preliminary results from English, if they do reflect relations among words-predicates and nouns,those relations are nearly identical in the two language, a fact one might want to explain by the regularities in the world.

Experiment 2

In Experiment 2 uses the results of Experiment 1 to simulate the boundary shift reported by Yoshida and Smith.

Experiment to be simulated The specific goal is to simulate Yoshida and Smith's second experiment which showed a boundary shift in the animate - inanimate boundary Japanese speakers relative to English speakers. The participants in Yoshida and Smith's experiment were 3-year-old English and Japanese monolingual children. The experimenters presented children with an ambiguous entity that could be seen as depictions of ether animates or inanimates and named it with a novel label (e.g. in Japanese 'Kore-wa teema dayo', in English 'This is a teema.'). Experimenters did not provide any cue such as "iru" or "aru" which might cue these as depictions of animates or inanimates Experimenters then presented children with test objects and asked them whether the test object had the same name Exemplars and test objects matched or did not match in three perceptual features (Table 3).

The results suggested that Japanese speakers treated these ambiguous forms as depictions of animates, extending the name to new instances by multiple similarities. In contrast, English speakers treated them as inanimates objects, extending the name to new instances by shape. Thus, Yoshida and Smith proposed that the Japanese speaking children included a wider range of kinds in the animate category relative to English speakers just as English speakers include a wider range of instances in the objects category than do Japanese speakers. The question for Experiment 2 is whether we can use the adult judgments in Experiment 1 to simulate these results.

Method

Following Yoshida and Smith's method, we assumed that objects categories are defined in terms of shape, color and texture, and that other nonstudied features will have no effect on the similarity of a test object to the exemplar.

We also assume that children's name extensions are based on the psychological distance between a test object and the exemplar. We defined the psychological distance between stimuli by the equation (3). Probability of 'yes' response which means two objects belong to the same category is defined by the equation (2).

$$P_{yes} = exp(-b\delta) \tag{2}$$

$$\delta = (\sum_{i \subset perception} D_i w_{il} \mid (e_i - s_i)^m \mid)^{\frac{1}{m}} \tag{3}$$

Table 3: Experimental conditions of Yoshida & Smith (2003). 'm' means feature match between exemplar and test object, and 'N' means non-match

condition	1	2	3	4	5	6
shape	m	m	m	m	N	N
texture	m	m	N	N	m	N
color	m	N	m	N	N	m
	S+T+C	S+T	S+C	S	T	C

$b > 0$ is the scaling parameter of the transfer between a distance and a yes-response ratio , and $m > 0$ is the metric parameter. $i \subset perception = \{shape(S), color(C), texture(T)\}$ means the population of the perceptual features. e_i represents the ith perceptual dimension of the exemplar, and it is a random value from 0 to 1. s_i represents the ith perceptual dimension of the test stimulus. s_i is a random value from 0 to 1 in case of feature non-match or the same value as the exemplar in case of feature match (see also Table 3). w_{il} is the value of ith perceptual weight of l ($l \subset \{English, Japanese\}$) participant(see also Table 2). D_is are the supplementary terms which represent ith perceptual bias common in English and Japanese. We added these terms to the model because the feature differences of stimuli were not controlled in the behavioral experiment. D_is represent the relative mean difference of perceptual features. The model have four free parameters (b, m, two D_is) ,because D_is are the ratios among three perceptual features.

Results and Discussion

We simulated the second experiment of Yoshida and Smith (Figure 4) by the computational model (Figure 5). We used Monte Carlo simulation to estimate optimal parameters. In the result, we estimated $b = 12$, $m = 0.8$、 $(D_{shape}, D_{texture}, D_{color}) = (7, 1, 0.6)$ ($D_{texture} = 1$ is constant) and $R^2 = 0.916$ between the response patterns (12=2 (language of participants)×6 (feature controlled condition)) of simulation and those of behavior. When we did not add two parameters D_is, the fitness of the model was $R^2 = 0.683$. This suggested the methodological problem of estimation by the equation (2).

In the behavioral experiment, the English speakers categorized the stimuli based on their shape and the Japanese speakers categorized them based on their multiple features. These results suggested that the English speakers categorized ambiguous objects as inanimates whereas Japanese speakers categorize them as animates. Thus our model fitted the behavioral results well, and provides a simple account of the crosslinguistic difference.

General Discussion

Recent studies on early word acquisition have shown that some biases, such as shape bias, are not so universal, but dependent on context and language. For example, Children speaking English show stronger shape bias for

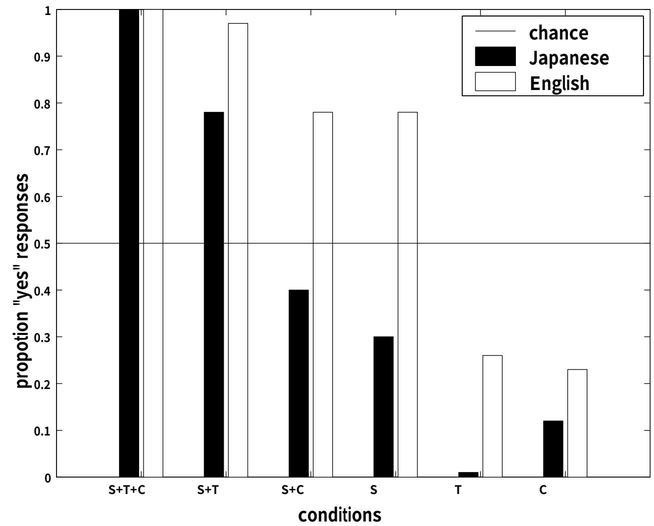

Figure 4: The behavioral data of Yoshida & Smith (2003). The English speakers categorize stimuli based on shape, while the Japanese speakers categorize them based on multiple features.

inanimate objects than those speaking Japanese. These findings are explained by postulating children's linguistic, cultural and category knowledge influences boundaries between ontological categories.

The present simulation offers a mechanism. The results of the simulation suggest that (1) ontological categories may not be a special nor given but an emergent property derived from multidimensional perceptual and linguistic features, and (2) crosslinguistic differences along this ontological continuum can be explained by a difference in the emergent variable due to different statistical structure of linguistic features. Specifically, we assumed that the emergent property can be extracted by information compression of the multidimensional feature space, such as PCA. To evaluate whether we can account for the behavioral findings, we conducted a survey to obtain the multidimensional feature space of objects, and a series of quantitative analyses to obtain the language specific ontological spaces. Without linguistic features, the compressed perceptual space spanned by two principal components was organized by objects' solidity or size. Thus, a solidity-dominant space can be derived from the perceptual feature space, but there was no principal component representing an "individuation continuum" from animates to objects to substances. More interestingly, addition of linguistic features made the ontological space more well-defined, and the estimated language-specific ontological spaces are quite consistent with previous findings. The estimated English ontological space was solidity-dominant and shape-weighted. This is consistent with Colunga & Smith (2000) and Samuelson (2002) showing that American children attended solidity of objects in object categorization. On the other hand, the estimated Japanese ontological space is animacy-dominant and color-and-texture-weighted, which is consistent with Yoshida & Smith (2001, 2003) showing that

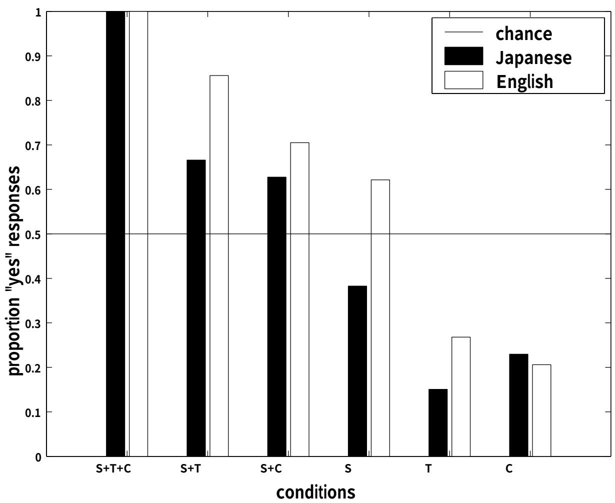

Figure 5: The result of simulation.
The coefficient of determination of 12 responses pattern (R^2) is 0.916

Japanese children attended multiple features of objects.

Furthermore, objects/substance boundary was clearer in the English space than the Japanese space. This results is consistent with Imai & Gentner (1997). In addition to qualitative matches with previous data, our theory make a good quantitative fit to the behavioral data of Yoshida & Smith (2003). With a simple computational model that categorization response is based on similarity derived from a distance on the ontological space, the behavioral data showing difference in shape bias for objects between English and Japanese speaking children with various different stimulus conditions could be simulated quite well.

Beyond the "boundary shift" Our account expands the boundary shift hypothesis in the following senses. First, our theory proposes an underlying mechanism of boundary shift in a quantitative fashion. Second, the individuation continuum is not a separate dimension, but a statistical property embedded in the multidimensional feature space. Ontological features such as animacy and solidity may be extracted from perceptual and linguistic features through statistical learning. This suggests a possibility that more abstract conceptual features are also formed by statistical learning of basic perceptual and linguistic features.

Acknowledgments

This work was supported by Grants-in-Aid for Scientific Research from JMEXT (No. 15650046), and the 21st Century COE Program from JMEXT (D-2 to Kyoto University).

References

Colunga, E. & Smith, L. (2000) Committing to an Ontology: A Connectionist Account, *The Twenty Second Annual Meeting of the Cognitive Science Society*.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59* (5, Serial No. 242) Chicago: University of Chicago Press.

Imai, M. & Gentner, D. (1997). A cross-linguistic study of early word meaning: universal ontology and linguistic influence., *Cognition*, *62*, 169-200

Jones, S. S. & Smith, L. B. (2002). How relevant properties for generalizing object names., *Developmental Science*, *5*, 219-232

Keil, F. (1979). Semantic and conceptual development: An ontological perspective, MA: Harvard University Press.

Keil, F. (1981). Constraints on knowledge and cognitive development, *Psychological Review*, *88*, 197-227.

Landau, B., Smith, L.B. & Jones, S.S. (1988). The importance of shape in early lexical learning, *Cognitive Development*, *3*, 299-321.

Lucy, J.A. (1992). Language diversity and thought: A reformulation of the linguistic relativity hypothesis., Cambridge: Cambridge University Press.

Markman, E.M. & Hutchinson, J.E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations., *Cognitive Psychology*, *16*, 1–27 .

Osgood, C.E., Suci, G.J. & Tannenbaum, P.H. (1957). *The measurement of meaning.*, Univ. of Illinois Press.

Quine, W.V.O. (1960). *Word and Object.*, Cambridge, MA:MIT Press,.

Samuelson, L.K. (2002) Statistical Regularities in Vocabulary Guide Language Acquisition in Connectionist Models and 15-20 Month Olds., *Developmental Psychology*, *38*, 1016–1037.

Samuelson, L. and Smith, L.: Early noun vocabularies: do ontology, category structure and syntax correspond?, *Cognition*, Vol. 73, pp. 1–33 (1999).

Soja, N. N. , Carey, S. & Spelke, E. S. (1991). Ontological categories guide young children's inductions of word meanings: object terms and substance terms., *Cognition*, *38*, 179–211.

Sommers, F. (1963). Types and ontology., *The Phylosophycal Review*, *72*, 327–363.

Spelke, E. S. (1990). Principles of object perception., *Cognitive Science*, *14*, 29–56 .

Yoshida, H. & Smith, L. B. (2001) Early noun lexicons in English and Japanese., *Cognition*, *82*, 63–74.

Yoshida, H. & Smith, L. B. (2003) Shifting ontological boundaries: how Japanese- and English- speaking children generalize names for animals and artifacts., *Developmental Science*, *6*, 1–34.